STUDIES IN STANDARDIZATION - SERUM CHOLESTEROL ANALYSIS PERFORMED FOR EPIDEMIOLOGICAL INVESTIGATIONS

PROEFSCHRIFT

ter verkrijging van de graad van doctor in de geneeskunde
aan de Erasmus Universiteit te Rotterdam
op gezag van de Rector Magnificus
Prof. Dr. B. Leijnse
en volgens besluit van het College van Dekanen.
De openbare verdediging zal plaats vinden op
vrijdag 15 juni des namiddags
te 4.15 uur precies

door

GEERT JURTKO MENNO BOERMA

geboren te Pieterburen (Gr.)

Dr. W. Backhuys, Uitgever
Rotterdam, 1979

Promotor: Prof. Dr. B. Leijnse
Co-referenten: Prof. Dr. H. A. Valkenburg
Dr. A. P. Jansen

Het verschijnen van dit proefschrift is mede mogelijk gemaakt door financiële steun van de Nederlandse Hartstichting.

Aan mijn ouders
Aan Nell en Frank en Robert

CONTENTS

DANKWOORD - ACKNOWLEDGEMENTS

Dit proefschrift is ontstaan uit het werk van zo velen, dat het noemen van hen allen niet wel mogelijk is. Allen komt mijn dank zeker toe, doch ik meen niemand tekort te doen wanneer enkele personen wel bij name worden vermeld.

Mijn promotor, Prof. Dr. B. Leijnse, dank ik zeer voor de stimulerende wijze waarop hij de groei van deze dissertatie vanaf het eerste begin heeft aangemoedigd. Beide co-referenten, Prof. Dr. H. A. Valkenburg en Dr. A. P. Jansen, besteedden veel tijd aan het manuscript waardoor een aantal aspecten duidelijker uit de verf kon komen.

Tienduizenden cholesterolanalyses, waarvan duizenden controles, zijn in de loop van de jaren verricht door de heer N. L. Kranendonk, mevr. J. A. L. Jacobs-de Haas, mevr. F. A. C. de Jong-van Staveren en de heer J. A. de Roo, waarbij een hoge kwaliteit werd gehandhaafd. In het kader van hun HBO-opleiding klinische chemie werkten mej. Liesbeth Boer, mej. Marja Krijnen en de heer Hans Westervoorde mee aan onderzoek over enkele analysemethoden.

In het KRIS project heb ik veel plezierige samenwerking ervaren in de Rotterdamse groep waarbij ik speciaal teamarts Anneliek Dubbeldam-Marree dank zeg en ook graag de statistische adviezen van Drs. Peter Das en Drs. Paul Mulder vermeld.

In het CB Project worden alle werkzaamheden gecoördineerd door de Stuurcommissie en ik ben alle medeleden zeer erkentelijk voor hetgeen ik in vele vergaderingen heb kunnen opsteken. Dr. K. Styblo verzamelde een groot deel van de in hoofdstuk 4 weergegeven resultaten, waarbij hij door zijn echtgenote werd gesteund. Onze blinde kwaliteitsbewaking komt - na het werk in het laboratorium en in de CB's - goeddeels uit hun handen.

De plannen en werkzaamheden van de Stichting KCA zijn voornamelijk in de WAR voorbereid. De medewerking van de deelnemende laboratoria was uiteraard onmisbaar. Drs. Coen van Gent dank ik voor het vele malen praktisch samenwerken

en de collegae van het RIV voor hun tijd en hun serum materialen. Ook het nauwgezette werk aan de referentiemethode, in ons referentielaboratorium, door de heer T. L. Liem, dient hier gememoreerd te worden.

Gaarne zeg ik Dr. Bert Blijenberg dank, die mijn regelmatige afwezigheid in het CKCL collegiaal heeft opgevangen. Voor het laboratorium in het Thoraxcentrum geldt hetzelfde; degenen die er dagelijks werken, hebben mijn monomane bezigheden loyaal verdragen en mogelijk gemaakt, niet in het minst door de voortvarende leiding van mej. Jeannette Veelbehr.

Het ontstaan van het uiteindelijke manuscript dank ik deels aan mevr. E. M. Schuur-Buik en vooral aan mej. Ella de Korte, die vaardig, in betrekkelijk weinig tijd kans zag om vrijwel alle hoofdstukken meer dan eens te typen! De figuren - een tweetal is van de Audiovisuele Dienst - werden bijna alle vervaardigd door de heer Piet Noordeloos, die ook diverse tabellen bewerkte en de drukproeven nazag! Uitgebreide statistische adviezen werden ingewonnen bij Ir. Paul Schmitz, afd. Biostatistica en miss Mariys Stewart controleerde het Engels. Ik ben deze mensen van het laatste uur zeer erkentelijk.

Last, but certainly not the least, I gratefully acknowledge the many things I have learned from Dr. Alan Mather. As a special adviser from CDC, he assisted the KRIS laboratories in their standardization effort. Dr. Adolfas Toleikis in Kaunas played an important role as well. Both of them have contributed a great deal to chapter 3. I also wish to mention the hospitality and cooperation of Dr. Mather and many others at CDC, of Dr. Dusan Grafnetter at the Prague Reference Laboratory and of Dr. Igor Glasunov and Dr. Ed Dowd, WHO, Geneva, who by their support contributed greatly to the writing of this thesis.

Het meest dankbaar ben ik mijn ouders, die mij zeer langdurig vele kansen geboden hebben en mijn gezin, voor alle niet-wetenschappelijke steun die voor een promovendus toch essentieel is.

Cover: the test tube represents the oldest quantitative analytical method for cholesterol (LB-reaction) whereas the mass spectra are results of the latest, definitive, method (IDMS). The two spectra were kindly made available by Dr. B. G. Wolthers, CKCL (Head: Dr. A. Groen), University Hospital, Groningen.

LIST OF ABBREVIATIONS

AAS	atomic absorption spectroscopy
ASCP	American Society of Clinical Pathologists
ASTM	American Society for Testing and Materials
BCR	Bureau Communautaire de Référence; European Community Bureau of Reference
CAP	College of American Pathologists
CB	Consultation Bureau; Chest Clinic for detection and prevention of tuberculosis
CCSP	Cooperative Cholesterol Standardization Program
CDC	Center for Disease Control, Atlanta, Ga. USA
ChOD	cholesterol oxidase
CKCL	Central Clinical Chemistry Laboratory of the Rotterdam University Hospital 'Dijkzigt'
COPIH	Commission on Detection and Prevention of Ischemic Heart disease of the NVAB
CV	coefficient of variation
EDTA	ethylenediamine tetra acetate
FDA	Food and Drug Administration
GLC	gas-liquid chromatography
HDL	high-density lipoproteins (α-lipoproteins)
ICSH	International Commission for Standardization in Hematology
IDSM	Isotope dilution - mass spectrometry
IFCC	International Federation of Clinical Chemistry
IKEM	Institute for Clinical and Experimental Medicine; Prague, Czechoslovakia
ISO	International Standardization Organization
KCA	Foundation for Quality Control of Chemical Analysis in Epidemiological Investigations
KRIS	Kaunas Rotterdam Intervention Study
LB	Liebermann-Burchard (colorimetric analysis)
LDL	low-density lipoproteins (β-lipoproteins)
mg%	milligrams per 100 milliliters
n	number (of persons, or measurements etc.)
NACC	Netherlands Association for Clinical Chemistry
NBS	National Bureau of Standards; Washington DC, USA
NEN	Netherlands Norm
NNI	Netherlands Normalization Institute; The Hague
NVAB	Netherlands Association for Occupational Health
NVKC	Netherlands Association for Clinical Chemistry
p	probability of a result being due to chance

PAP	para aminophenazone
QAP	Quality Assurance Program (of CAP)
QC	quality control
RIV	National Institute of Public Health; Bilthoven, the Netherlands
s	standard deviation
sem	standard error of the mean
SI	International System (of units)
SRM	Standard Reference Material (issued by NBS)
TG	serum triglycerides
TNO	Organization for Applied Research (in the Netherlands)
VLDL	very low-density lipoproteins (pre-β-lipoproteins)
WHO	World Health Organization
$\bar{x}$	mean of a number of results

LIST OF TERMINOLOGY

English (Dutch expression)	
Accuracy (juistheid, zuiverheid)	Agreement between the best estimate of a quantity and its true value.
Analyte	The component to be measured.
Best estimate (meetverwachting)	The mean of a set of replicate measurements in the same material (Population mean).
Bias (onjuistheid, systematische fout)	Numerical difference between the best estimate and the true value (Inaccuracy, Systematic error). Whereas true values are unknown, in this text differences will relate to reference method values.
Calibration (calibratie, ijking)	Process of relating a reading to the quantity required to be measured.
Detection limit (detectiegrens)	The smallest analyte concentration that can be distinguished from a blank, with a stated probability.
Epidemiology	The study of the distribution and frequency of diseases in the population and of the various factors that determine these.
Error, systematic	See: Bias.
Error, random (toevallige fout)	Numerical difference between a result and the best estimate, the sign (positive or negative) being unpredictable.
Interference (storing)	The effect of another component on the accuracy of measurements of the analyte.
Matrix	The milieu in which the analyte is present.
Method, analytical (meetmethode, analysetechniek)	Set of written instructions, describing the procedure, materials, equipment and working conditions which enables different technicians to carry out a measurement in identical fashion.
Method, definitive (definitieve methode)	A method in which, after exhaustive investigation, with our present knowledge and technical capabilities, no source of bias is detectable.
Method, reference (referentiemethode)	A method which after exhaustive investigation has been shown to have negligible bias for stated practical purposes.
Precision (precisie)	Agreement between results of replicate measurements in identical material; the quantitative measure is the standard deviation or coefficient of variation. (see: repeatability; reproducibility)

Normalize, to (normaliseren)	To accept a specified product or method as the only one to be produced or utilized.
Plasma (plasma)	Fluid obtained upon centrifugation of uncoagulated blood. The kind and quantity of the anticoagulant(s) used must be stated.
Quality control (kwaliteitsbewaking)	Procedures used to recognize, study and minimize all errors arising in the laboratory - or outside it - that influence reported analytical results.
Quality control, intralaboratory or internal	All procedures used for quality control without using results from other laboratories.
Quality control, interlaboratory or external	Procedures for quality control utilizing results and/or materials from other laboratories.
Repeatability (dupliceerbaarheid)	Precision obtained within a run. The standard deviation may be calculated from replicate analyses of identical material or from duplicate measurements in several specimens.
Reproducibility (reproduceerbaarheid)	Precision obtained over long periods. It must be stated whether between-run, between-week or other intervals are under consideration.
Result (meetuitkomst, -waarde)	Value obtained in a measurement, subject to usual quality control procedures.
Run (meetserie)	Set of consecutive measurements performed without interruption, results of which are obtained by means of a single calibration procedure.
Sensitivity, chemical (analytical) (gevoeligheid)	The ability of an analytical method to detect small quantities of the analyte (see also: detection limit).
Sensitivity of a test (clinical)	The fraction of positive test results that occurs in patients having the disease; percentage true positives.
Serum	Liquid, obtained upon centrifugation of clotted blood.
Serum calibrator (calibratieserum)	A secundary standard.
Serum, control (controleserum)	Serum or a material made from serum (human or animal) that can be used to check the precision of a method.
Serum, reference (referentieserum)	A serum in which one or more concentrations have been determined with a definitive or a reference method and that may be used to measure bias of another method.
Specificity, chemical	The ability of an analytical method to determine sole-

(analytical) (specificiteit)	ly the analyte it purports to measure.
Specificity of a test (clinical)	The fraction of negative test results that occurs in healthy persons; percentage true negatives.
Standard (standaard)	A material with one or more defined properties that can be used to transfer the value of a measured quantity - chemical, biological or other - between one place and another.
Standard solution, primary (primaire standaard)	Solution in which the concentration is determined solely by weighing an amount of standard into a stated volume of a solvent.
Standard solution, secundary (secundaire standaard)	Solution in which the concentration of one or more components has been determined by an analytical method of stated accuracy.
Standard deviation (standaardafwijking)	The measure of random variations in results. It may be calculated from results obtained in replicate analyses of a single material: $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n - 1}}$ (n results) or from differences between duplicate results in several samples: $s = \sqrt{\frac{\Sigma (\Delta x_i)^2}{2n}}$ (n samples)
Standardize, to (standaardiseren)	To bring into conformity with a standard.
Systematic error	See: Bias.
Value, assigned (toegekende waarde)	Value assigned by some official body subject to defined conditions established by that body.
Value, definitive (definitieve waarde)	Best estimate obtained with a definitive method.
Value, reference method	Best estimate obtained with a reference method.
Value, target or expected	In this text this is identical with assigned value.
Value, true (gezochte, absolute waarde)	Value which a concentration theoretically has, but that can only be approximated by accurate methods, preferably a definitive method.

The problem is simple - there is just too much cholesterol!
J. R. Sabine (1977) (1).

CHAPTER 1

MEASUREMENT OF SERUM CHOLESTEROL - WHY AND HOW?

1.1. Serum cholesterol emerging as a risk indicator

Cardiovascular diseases are the major cause of death in industrialized countries. Death rates from coronary heart disease have increased steadily up to some 30% of the total death toll - they have stabilized in the past ten years and even seem to decline. The relative youth of the victims remains particularly disturbing (2 - 6). The etiology and pathogenesis of the disease are still incompletely known (7).

From a considerable number of epidemiological studies, however, a great many risk indicators (or risk factors) have emerged (8 - 13). The most important among these are elevated blood pressure, elevated serum cholesterol levels and cigarette smoking. The design and objectives of large scale trials have, as a rule, been quite variable. It has been recognized that the conductance of such studies - the recruitment of a representative target population, the data handling, etc. - is an extremely complicated task. Comparability of the several findings is limited (14) and in view of the need to exchange useful scientific information the various study methods must be standardized (15, 16, 17). This need also applies to the measurement of cholesterol. Comparability requires also the compatible analysis and presentation of results and sufficient description of all procedures.

In many countries official and unofficial recommendations have been issued that must contribute to prevention of atherosclerosis. Chiefly this amounts to: a more sober life-style; consumption of less calories, saturated fat and sugar and more poly-unsaturated fats and fiber; consumption of sufficient proteins, vitamins, minerals; no smoking; more physical activity (4, 18 - 24); avoidance of stress. In fact such rules are riminiscent of old medieval wisdom as expressed by the medical tradition in Salerno (± 1200),

'Use three physicians still, first Doctor Quiet
next Doctor Merriman, and Doctor Diet.' (25)

The recommendations also stress the need to continue research into the under-

lying causative mechanisms of the disease (26, 27), as well as regular medical examination of some sort to detect high-risk individuals. One quotation where this desire is expressed reads:

> *'Measurement of the plasma lipidprofile, particularly cholesterol, become a routine part of all health maintenance physical examinations. Such measurements should be made in early adulthood, when coronary heart disease is still rare, and repeated at appropriate intervals. The potential impact of other risk factors should also be periodically assessed.'* (20)

Whether one agrees with such statements or not, it cannot be denied that their impact is huge and that the news media have made large parts of the general population aware of the existence of risk indicators. The subsequent demand for health check-ups has made cholesterol analysis an interesting business in more than one respect. Moreover, government funding of cardiovascular research - in the USA the budget of the National Heart, Lung and Blood Institute was increased from $ 223 million in 1972 to $ 397 million in 1977 - adds to these interests (28). Several industries and companies offer regular medical examinations to their personnel. Whether all this will lay the foundation for an effective future prevention programme depends greatly on the possibilities to compare, and pool, much of the data to give consistent answers to some pertinent questions. One such question might be that of the 'best' or 'correct' or 'optimal' serum cholesterol level (29, 30).

1.2. The difficulty of interpreting serum cholesterol values

It may, at a first glance seem surprising that a well established and common laboratory test like serum cholesterol analysis, should create problems at all. Since the beginning of this century gradually all clinicians and general practitioners have become familiar with cholesterol measurements. The usage of such analytical results, however, is changing in the sense that a new dimension has been added by medical screening. Since the advent of *epidemiological* surveys and screening programmes with large samples of the population, cholesterol analyses and their interpretation have moved away from the traditional hospital situation. There the serum cholesterol values of a patient were - and of course still are - details among many other data that built up a pattern of information on which diagnosis and therapy were based. When in doubt the doctor may gather additional data by requesting more tests to confirm some and rule out other possibilities.

In modern screening examinations - e.g. studies to estimate risk of developing coronary heart disease in an apparently healthy population or periodical medical

check-ups - only a few tests are used to classify the participants into a limited number of categories. No diagnosis in the classical sense is attempted at all. Now a laboratory error, in the absence of counterbalancing mechanisms, will put an individual into the wrong category or will lead to erroneous conclusions when values are compared with those from previous measurements. The quality of laboratory performance therefore must be even better than would be needed in a hospital situation. Such quality requirements are a new challenge to clinical chemistry (31, 32).

Nowadays one realises also that the value of study results is limited when comparisons with similar projects in or outside the country are not possible. The lack of interlaboratory comparability is a major problem that did not attract much attention until clinical chemical investigations became common in epidemiological work. A main cause is the enormous multiplicity of methods, a general phenomenon in clinical chemistry. Schwartz and Hill quote a recent WHO-survey (33) in which 28 methods for cholesterol were found to be in common use in European laboratories. Practically all of these are not analytically specific, meaning that other serum components along with cholesterol will contribute to the overall analytical result by chemical side reactions, light absorption and such. The blood sampling technique is of great importance as well!

In the following chapters a number of problems one meets when performing serum cholesterol analyses for mass examinations, and interpreting their results will be under discussion. Following a literature review, the set-up of our quality control system and the development of standardization procedures are described in detail. A national standardization programme aiming at comparability of all population studies in the Netherlands is discussed also.

In this way several aspects of clinical chemical analysis are studied in order to evaluate whether the laboratory method used in the hospital, or occupational health service, can be employed in epidemiology.

REFERENCES

Chapter 1

1. J. R. SABINE: Cholesterol, Marcel Dekker Inc., New York (1977).
2. Th. F. S. M. VAN SCHAIK: Nota samengesteld door de Commissie 'Meervoudig onverzadigde vetzuren' van de Voedingsraad (1973). Voeding *3*, Supplement (1974).
3. V. K. OVKAROV and V. A. BYSTROVA: Present trends in mortality in the age group 35 - 64 in selected developed countries between 1950 - 1973. World Health Statistics *31*, 208 - 346 (1978).

4. Joint working party of the Royal College of Physicians of London and the British Cardiac Society: Prevention of coronary heart disease. J. Roy. Coll. Phycns. *10*, no. 3 (1976).
5. F. H. EPSTEIN: Coronary heart disease epidemiology revisited. Clinical and community aspects. Circulation *48*, 185 - 194 (1973).
6. R. S. ROSS: Presidential address: The case for prevention of coronary heart disease. Circulation *51*, 57 - 62 (1975).
7. R. J. J. HERMUS: Experimental atherosclerosis in rabbits on diets with milk fat and different proteins. Doctoral thesis Wageningen (1975).
8. W. B. KANNEL, T. R. DAWBER, G. D. FRIEDMAN, W. E. GLENNON and P. M. McNAMARA: Risk factors in coronary heart disease. An evaluation of several serum lipids as predictors of coronary heart disease. The Framingham Study. Ann. Int. Med. *61*, 888 - 899 (1964).
9. A. KEYS: Coronary heart disease in seven countries. Circulation *41, 42*, Supplement I (1970).
10. F. H. EPSTEIN, L. D. OSTRANDER, B. C. JOHNSON, M. W. PAYNE, N. S. HAYNER, J. B. KELLER and T. FRANCIS: Epidemiological studies of cardiovascular disease in a total community. Tecumseh, Michigan. Ann. Int. Med. *62*, 1170 - 1187 (1965).
11. A. ENGLHARDT: Risikofaktoren kardiovaskulärer Erkrankungen. Dtsch. Med. Wschr. *99*, 842 - 847 (1974).
12. K. WESTLUND and R. NICOLAYSEN: Ten-year mortality and morbidity related to serum cholesterol. Scand. J. Clin. Lab. Invest. *30*, Supplement 127 (1972).
13. S. HEYDEN: Risk factors of ischaemic heart disease. Boehringer Mannheim 'Study Monographs' series (1975).
14. L. WERKÖ: Risk factors and coronary heart disease - facts or fancy? Am. Heart. J. *91*, 87 - 98 (1976).
15. G. A. ROSE and H. BLACKBURN: Cardiovascular survey methods. WHO Monograph Series, No. 56, Geneva (1968).
16. I. GLASUNOV, J. DOWD, Z. JAKSIC, B. KESIC, D. RAY, C. STEINBERGER, J. STROMBERG and S. VULETIC: Methodological aspects of the design and conduct of preventive trials in ischaemic heart disease. Int. J. Epid. *2*, 137 - 143 (1973).
17. Mass Health Examinations. Public Health Papers *45*. WHO, Geneva (1971).
18. Mededeling van de Voedingsraad: Advies inzake de vraag: Zijn er adviezen en/of maatregelen gewenst ten aanzien van de hoeveelheid en (of) aard der vetten in de voeding van de Nederlandse bevolking? Ned. T. Geneesk. *117*, 1984 - 1987 (1973).
19. A. KEYS: Official collective recommendation on diet in the Scandinavian countries. Nutr. Rev. *26*, 259 - 265 (1968).

20. A Joint Statement of the Food and Nutrition Board, Division of Biology and Agriculture, National Academy of Sciences - National Research Council and the Council on Foods and Nutrition, American Medical Association: Diet and coronary heart disease. Am. J. Clin. Nutr. *26*, 53 - 54 (1973).
21. J. H. REISEL: Klinische atherosclerose. Een studie over primaire interventie. Bohn, Scheltema en Holkema, Utrecht (1976).
22. S. DAYTON and M. L. PEARCE: Prevention of coronary heart disease and other complications of atherosclerosis by modified diet. Amer. J. Med. *46*, 751 - 762 (1969).
23. Inter-Society Commission on Heart Disease Resources: Atherosclerosis study group and epidemiology study group: Primary prevention of the atherosclerotic diseases. Circulation *42*, A55 - A95 (1970).
24. J. H. DE HAAS: Primary prevention of coronary heart disease. A socio-pediatric problem. Hart Bull. *4*, 3 - 11 (1973).
25. Salernitan Guide to Health, quoted in Encyclopaedia Britannica 15th ed. Macropaedia, Vol. 11, p. 828, University of Chicago (1974).
26. G. B. KOLATA and J. L. MARX: Epidemiology of heart disease: searches for causes. Science *194*, 509 - 512 (1976).
27. G. B. KOLATA: Atherosclerotic plaques: competing theories guide research. Science *194*, 592 - 594 (1976).
28. R. B. CONN: Clinical laboratories. Profit center, production industry or patient-care resource. N. Engl. J. Med. *298*, 422 - 427 (1978).
29. I. S. WRIGHT: Correct levels of serum cholesterol. Average vs. normal vs. optimal. J. Am. Med. Assoc. *236*, 261 - 262 (1976).
30. E. CHERASKIN and W. M. RINGSDORF: Predictive Medicine V. Linear versus curvilinear functions. J. Am. Geriatrics Soc. *19*, 721 - 728 (1971).
Predictive Medicine VI. Physiologic versus normal values. J. Am. Geriatrics Soc. *19*, 729 - 736 (1971).
Predictive Medicine VII. The specificity of tests. J. Am. Geriatrics Soc. *19*, 802 - 806 (1971).
31. IFCC; Provisional Recommendations on quality control in clinical chemistry. Part 6. Quality requirements from the point of view of health care. Clin. Chem. *23*, 1066 - 1071 (1977).
32. R. J. EILERS: Quality assurance in health care: missions, goals, activities. Clin. Chem. *21*, 1357 - 1367 (1975).
33. M. K. SCHWARTZ and P. HILL: Problems in the interpretation of serum cholesterol values. Prev. Med. *1*, 167 - 177 (1972).

There is also too much cholesterol in the literature. The index to Biological Abstracts for a recent 6-month period (1) *lists 487 items under the heading Cholesterol - and my guess is that two-thirds of these either flatly contradict, or at least provide an alternative observation to or explanation for, the corresponding number of the previous six months.*

J. R. Sabine (1977) (2).

CHAPTER 2

SERUM CHOLESTEROL - A REVIEW OF ANALYSIS, BIOLOGICAL VARIABILITY AND STANDARDIZATION

Introduction

The serum cholesterol level measured in two blood samples taken from one individual will rarely be the same. The results, presented to us as a number of mmol/l or mg% or else (!), are greatly influenced by biological and analytical variability. In this chapter a short review is presented of the factors that contribute most to such variations and thus hamper the interpretation of what we have measured. Of course we can only try to control the various analytical errors by attaining sufficient proficiency in the laboratory. For this, adequate control and standardization procedures are needed. The subjects to be discussed below are:

2.1. Multiplicity of analytical techniques
2.2. Biological variations
2.3. Precision and accuracy
2.4. Definitive and reference methods
2.5. Quality control
2.6. Reference laboratories
2.7. Standardization and comparability
2.8. Examples of standardization: CB Heart Project and KRIS
2.9. Standardization in the Netherlands: KCA

2.1. Multiplicity of analytical techniques

Since the discovery of cholesterol in gallstones by Poulletier de la Salle (± 1769)

the same substance was found to exist in all animal tissues. Until 1853 over a hundred references had appeared in the literature (3). During the past hundred years, since Salkowski described a reaction of cholesterol with sulfuric acid (4) literally hundreds of papers have been published on analytical methods - including many modifications and refinements - for serum cholesterol and cholesterol esters. A few important developments appear in table 2.1. Several authors have classified and reviewed cholesterol methodology (7 - 11), the latest of which includes the new enzymatic procedures (11).

Table 2.1.
Some landmarks in cholesterol research (3, 4).

1769	Poulletier de la Salle	: extraction of gallstones yields white crystalline compound.
1815	Chevreuil	: name 'cholesterine' from the Greek: chole = bile and stereos = solid.
1859	Berthelot	: with fatty acids synthetised esters.
1872	Salkowski	: reaction with sulfuric acid in chloroforum yields red colour.
1885	Liebermann	: same reaction including acetic anhydride yields green-blue colour.
1890	Burchard	: quantitative colorimetric analytical method.
1909	Windaus	: gravimetric analysis with digitonin.
1932	Wieland, Windaus and others	: elucidation of steroid formula.
1952	Abell et al.	: reference method (LB-reaction).
1953	Bloch, Cornforth and others	: biosynthesis described (in rat liver).
1973	Flegg, Richmond (5,6)	: enzymatic analysis in serum.

The Liebermann-Burchard reaction has become the most widely applied procedure. This colour reaction is a part of several versions of 'indirect' or 'direct' analytical methods. In the first, cholesterol and other lipids are extracted from the serum, esters may or may not be hydrolyzed with a solution of potassium hydroxide in alcohol, and then the LB-reaction is carried out with reagents in which the mixture of sulfuric acid, acetic acid and acetic anhydride has a composition adapted to the preferences and priorities of the authors. Many mechanized and partly auto-

mated modifications were developed (12). 'Indirect' methods gradually have been replaced by 'direct' methods, their advantage being the simplicity of manipulation. The serum sample is mixed without extraction and/or hydrolysis with a stabilized LB-reagent. The simplicity of these methods does make them attractive although a loss of specificity can not be denied and the water in the sample, the esters and other components influence the results.

In the Netherlands a modification of Huang's procedure is used (13) according to a recommendation of the 'Buitengewone Normcommissie' of the NACC and RIV. This method has been published under the auspices of the Netherlands Normalization Institute as method NEN 2415 (14).

Because the LB-method in many versions is used so often, a short summary of relevant aspects is given here, emphasizing the many pitfalls and trouble that may be encountered. The last paragraph mentions enzymatic methods also.

2.1.1. Mechanism of the Liebermann-Burchard reaction

A few studies have contributed observations that led to the first proposed mechanism. In concentrated chloroform solution cholesterol yields red or green colours (15) with strong acid, depending on the amount of acetic anhydride added (16). Molecules like $C_{54}H_{88}$ and $C_{54}H_{86}$ were identified (17) and a mechanism based on dimerisation of cholesterol with formation of sulfonic acid groups (9, 18) was suggested. Tietz inserted it in his authoritative textbook (1st edition) (19). Recent studies at NBS employing spectroscopy and mass spectrometry have shown various cations to be the more likely coloured products. Thus the red Salkowski product is the cholestatetraenylic cation (λ_{max} = 563 nm) and the Liebermann-Burchard colour stems from the cholestapentaenylic cation (λ_{max} = 620 nm) (20, 21). Still this reaction sequence does not explain very well the different colour development of cholesterol esters (fig. 2.1.) nor the colour reaction of some sterols with structures not closely resembling that of cholesterol (22).

2.1.2. Analytical specificity

Blood serum samples that are *lipemic, jaundiced* or *hemolytic* create a great number of problems in clinical chemical analysis. In the LB-reaction bilirubin causes erroneously elevated values (23), because its oxidation product biliverdin (λ_{max} = 619 nm) increases the optical density at 620 nm. Double wavelength measurements may correct for this (24). In our Huang method 10 μmol/l of bilirubin increases cholesterol results by some 0.05 mmol/l. Fortunately in screening studies hemolytic and jaundiced serum is rarely seen. Lipemia will as a rule lead to additional and/or repeated lipid analyses. Sterols other than cholesterol are said to

interfere: various contributions to the final values are reported: 0.4 - 1.3% from Δ^7-cholesterol (25, 26), 1 - 2% from cholestanol (27) and 1 - 3% from dehydrocholesterol (28). Total percentages of 10% are published (29 - 32) but must be attributed to unreliable analytical techniques (33).

The number of *drugs* known to influence cholesterol levels, either physiologically or in the analysis, is huge (well over a hundred) and discouraging. Such interferences are often small and usually ignored (34, 35). Again, in screening of the healthy population most of the drugs are probably not important, oral contraceptives being a major exception (36).

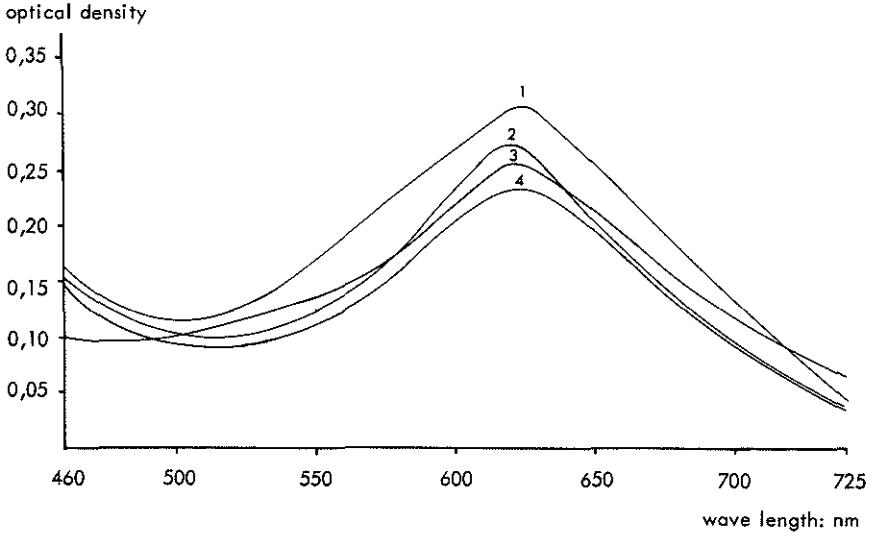
There exists considerable confusion in the literature on the increase or decrease of analytical results by the various *cholesterol esters*. Early data were collected by Brown (37) who cites several positive and negative errors caused by long chain fatty acid esters and makes the gloomy statement that:

> *'apparently accurate values might be due to a favourable combination of errors'.*

Since most of this, and more recent information (38, 39), relates to chloroform solutions, we decided to perform some measurements ourselves in the aqueous situation. Solutions of cholesterol esters were prepared in water, employing a detergent (40). Pure cholesterol linolate, cholesterol oleate and cholesterol palmitate were purchased in 100 mg vials (Merck, analytical grade). The exact amount was found by weighing the vial before and after washing the contents out with chloroform or hexane. The organic solvent was evaporated under nitrogen and the ester dissolved in 3.0 ml of TWEEN-40 (polyoxyethylen sorbitan monopalmitate, Sigma Chemical Co.). After adding 1.5 mg of anti-oxidant NDGA (nordihydro guaiaretic acid) the flask was heated in boiling water and a hot 0.1% sodium azide solution (preservative) was added dropwise with constant swirling of the flask. The total volume (± 15 ml) was such that concentrations of approximately 10 mmol/l were obtained. For comparison the same esters and cholesterol itself were dissolved into acetic acid, chloroform and isopropanol.

To these solutions (0.1 ml) the LB-reagent (acetic acid: acetic anhydride : sulfuric acid = 30 : 60 : 10 (v/v); with 2% anhydrous sodium phosphate (w/v), Merck analytical grade reagents) was added (5.0 ml, temperature 25° C). The colour development of the solutions was compared to that of corresponding cholesterol solutions after the maximum was reached (measurement after 16 minutes, λ 618 nm, Aminco-DW-2 spectrophotometer). The palmitate dissolved with difficulty and precipitated partly in the reaction which precluded an exact evaluation. Table 2.2. presents the ratio of optical densities of esters as compared with cholesterol itself. Spectrophotometric scans are shown in fig. 2.1. Our observations confirm the

Figure 2.1. Maximum colour development in direct LB-reaction; cholesterol and some esters in 10 mmol/l aqueous solution.



1: cholesterol linolate; 2: cholesterol palmitate, 5 mmol/l, O.D.scale x 2; 3: cholesterol; 4: cholesterol oleate.
Five ml of reagent mixed with 0.1 ml of sample; reaction time 16 - 20 minutes.

Table 2.2.
Relative colour development of cholesterol esters.[1]

	isopropanol	acetic acid	chloroform	water + TWEEN	percentage of serum total cholesterol
Cholesterol	1.00	1.00	1.00	1.00	25
Cholesterol linolate	1.06	1.24	1.25	1.30	38
Cholesterol oleate	0.80	0.98	0.84	0.94	15
Cholesterol palmitate	1.21	-	(1.07)[2]	(1.07)[2]	10
Number of analyses	2	2	4	12	

[1]) The coefficient of variation is approximately 5%.
[2]) Precipitation of some material during the reaction.

general conclusions one may draw from the literature regarding elevated colour development of the most abundant serum cholesterol esters. The aqueous solutions demonstrate that such esters indeed account in part for the positive bias of the direct LB-procedure, although we could not accurately predict it in mixed samples.

When a small serum sample is added to a few ml of LB-reagent its *water content* causes instantaneous heat generation and this greatly speeds up the reaction. The course of this reaction is determined in the mixing procedure for this reason. As soon as maximum extinction is reached (at 620 nm) it gradually diminishes while increasing at $\lambda = 410$ nm. Combined readings have been suggested (41). The temperature initially reached, is lower when cold reagent is used. But the final result, when employing a 25° C water bath, is the same (42). We have measured the following temperature elevations (Table 2.3.).

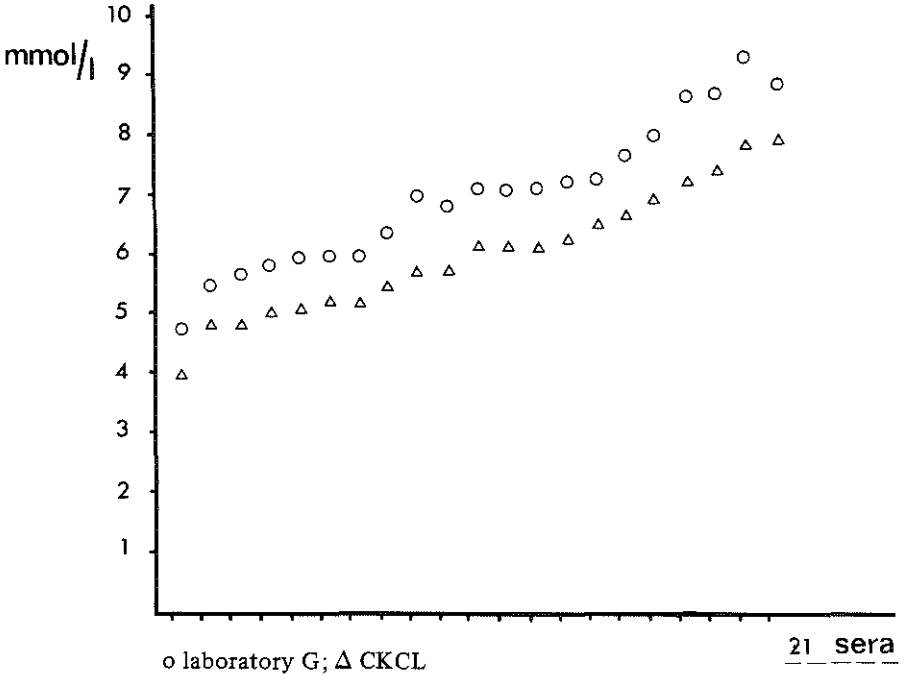
Table 2.3.
Temperature rise in direct LB-procedure.

Five ml of LB-reagent (22° C) mixed with:	Temp. 10 sec. after addition	Temperature change
0.1 ml distilled water	67°	45
0.1 ml aqueous cholesterol standard (10 mmol/l)	59°	37
0.1 ml cholesterol standard in isopropanol „	33°	10
0.1 ml cholesterol standard in chloroform „	25°	2
0.1 ml cholesterol standard in acetic acid „	22°	0

Standards of cholesterol in acetic acid or isopropanol produce little or no heat and the early stage of the LB-reaction will proceed calmly. For safety reasons it may be wise to add the sample to the test tube before the reagent - this can prevent spraying (44). Fading of the green-blue colour occurs faster when the reaction temperature is increased (44, 43, 8). The best way to standardize temperature control is to place the reagent tubes in a thermostatted bath as soon as possible after mixing of the sample and the reagent. Without this precaution a bad precision may occur. In connection with this temperature effect encountered in the direct LB-reaction two remarks must be made in addition: the reagent to sample volume ratio (or the amount of water added) should be such that the reaction does not slow down too much (44, 8) nor must the heat create brown colorations (45).

The importance of mixing is illustrated by our comparison of two pipetting sequences. Figure 2.2. shows duplicate results obtained on 22 patient specimens. In our laboratory (CKCL) a Hook and Tucker diluter/dispenser, type III, was used and in the other laboratory (G) the serum was pipetted manually (0.1 ml), followed by the reagent (5.0 ml): standards, reagents and further reaction conditions were the same (46). The systematic difference between our methods must be attributed mainly to the different early stage of the process: the mixing procedure. This is one example of slight modifications with serious consequences. The speed of placing the test tube in the thermostatted water bath is a part of this early stage (47).

Figure 2.2. Serum cholesterol - results of two pipetting procedures.



2.1.3. Interference of light, wavelength and temperature

The influence of light on the LB-reaction products is strongest at low wavelength (e.g. 430 nm) and at the usual 620 nm the routine laboratory situation with subdued daylight does not have a measurable effect. The lowered extinction - the maximum absorbance may be a few percent lowered in light - is compensated for

by the same effect on standards (43). When the maximum optical density is reached, a slow fading of the colour sets in while a yellow product (λ_{max} = 410 nm) appears. Measurements at this wavelength, however, suffer from extraneous light and the serum water content (41, 47). Colour stability is best at 540 nm but sensitivity is less and hemoglobin interference enhanced (42, 7). Bilirubin interference may be corrected by measurement at two wavelengths: 618 and 730 nm (24) (with electronic calculation). A water bath of 25° C is necessary to cool down the heated reaction mixture of samples.

Each sample must be mixed, cooled and measured in exactly the same way (8) because response rates for standards and serum are not parallel (see also Appendix 1) and timing errors will lead to poor precision.

2.1.4. The Liebermann-Burchard reagent

Several reagent compositions have been proposed (13, 48 - 51) and increased stability, sensitivity, speed and optimal behaviour towards esters are claimed. Kinetic measurement is possible also (52).

In any composition the preparation of LB-reagent is a hazardous job and care must be taken to avoid accidental heat production and spilling. The reagent and its components are dangerously corrosive and with regard to acetic acid vapours a well ventilated hood or fume-cupboard is needed. After the analysis the reagent mixture may be collected so as to minimise pollution.

Of course, these properties are very disadvantageous and have become one argument in favour of enzymatic reagents. Commercially available LB-reagents may be employed for convenience but their composition should be known and should not change between lot numbers. A properly made reagent mixture is always colourless.

2.1.5. Calibration and quality control

The particular problems with calibration and control of cholesterol measurements are the main subjects of this dissertation. Calibration is usually done with a solution of sufficiently pure cholesterol (several commercially available preparations) in an organic solvent like ethanol, propanol-2, acetic acid or in aqueous mixtures. All of these have a different matrix (physico-chemical surrounding of the analyte) and do not contain esters. Anyone will see at once the difference between the blue colour obtained with the standard and the green-blue colour obtained with serum. Viscosity differences may create pipetting errors (53, 54, 55). Temperature, timing and chemical interferences were discussed already. A proper blanking procedure is virtually impossible in this reaction; a reagent blank is used for establishing the base

line absorbance.

Quality control is concerned with checks on precision and on accuracy. The first is usually carried out by insertion of reconstituted, lyophilized control serum into the analytical runs. This is a useful, necessary and legitimate procedure. It must be kept in mind that homogeneity of a batch is a necessity (56). Accuracy, however, cannot be estimated in this way because the particular matrix of each control preparation precludes a general conclusion (52, 57). One example of matrix effects is the following: in a comparison between two groups of laboratories employing two continuous-flow-systems some lyophilized serum control (Quality Assurance Programme, ASCP) showed a 24% difference for both high and low values (proportional). A comparison of simultaneously analyzed patient sera revealed a *constant* difference of 1 mmol/l (58).

In other surveys similar problems are seen (59). Also, when human serum is used to make control materials or human lipids are isolated to spike animal serum one should be aware of matrix effects (60, chapter 5). And the stated label values should not be used in calibration before they have been confirmed by the reference method (61). Calibration with serum (see chapter 5) is a practical way to reduce the positive bias of the LB-method, as was proven in CDC-coordinated trials. The hazard of using only one single serum calibrator was also experienced (62, 63). Long term stability and vial-to-vial homogeneity are absolutely required (64-67).

2.1.6. Enzymatic cholesterol analysis

The appearance of the enzymatic reagents has in the past 6 years already resulted in a wide application. The advocated advantages are specificity, simplicity ('direct' procedure), mild reaction conditions and the small sample volume needed. Many procedures, several of them adaptations to various instruments, have been published (11). The reaction proceeds as follows:

a. cholesterol esters are split with an esterase.
b. cholesterol is oxidised with cholesterol oxidase (ChOD) and hydrogen peroxide is formed with cholestenone
c. hydrogen peroxide concentration is measured.

When the first step is omitted only free cholesterol is measured and omission of the second step allows for an approach to blank determinations.

The first enzyme does not have to be specific, but it must completely split all cholesterol esters within the necessary time period - short chain fatty acids in esters sometimes added to control serum are dissociated too slowly. And some of the long chain esters may behave differently too (68, 32). The second enzyme must be specific and also ensure a complete conversion of all cholesterol. Although several other

sterols are oxidized as well. They are not quantitatively important. The third step - the quantitative measurement itself - may be done in several ways: enzymatic conversion of peroxide with katalase, colour reaction with chromogens, electrochemical measurement with electrodes. In the index of recent Clinical Chemistry volumes alone one finds some 60 articles (1976: 22, 1977: 18, 1978: 22) dealing with enzymatic modifications, including HDL-cholesterol procedures. Interferences by lipemia, bilirubin, hemoglobin, drugs and other serum constituents and the balancing of optimal reaction conditions with regard to enzyme activities, buffer and buffer strenght, pH, detergents and chromogens are factors of importance (69 - 71). Calibration, or rather the inconsiderate selection of the standard, may lead to considerable bias (72, 73, 74).

Interlaboratory differences among enzymatic procedures are as large as in other methods. If this analysis is to be used as a reference procedure, standardized enzyme preparations and the employment of fool-proof detection techniques are required. One may think of UV measurement of cholestenone (75), oxygen electrodes (76) or calorimetry (77).

2.2. Biological variations

As a result of screening for presence of hyperlipidemia usually group mean values are studied and calculated averages of results are presented in tables and graphs. But before all this is possible the individual has been examined and classed into 'normal' or 'at risk' categories. In what way is a serum cholesterol measurement used to determine risk? The serum cholesterol level in human beings may fluctuate over a wide range of values. Screening studies have not sought to find 'elevated *mean* cholesterol levels' by averaging several analytical results. Instead our risk indicator is the transient elevated serum cholesterol at the time a single medical examination happens to be scheduled. Clearly at that moment the analysis must be precise.

Without good precision no valid distinction between low, medium or high values is possible. An estimate of the rate of erroneous classification is made in section 6. 5. When participating subjects have been divided into groups with different risk, re-examinations occur. Changes in the cholesterol are then evaluated. Here, in addition to laboratory error, the biological variations play an important role - and knowledge of their dimension is necessary.

The analysis, if not quite accurate, must have at least a known bias. Without this information cut-off levels have no meaning. The blood collection procedure must be taken into account too as an important source of variation (see section 3.4.1.3.).

Several values for 'personal biological variations' are available. Some are derived

from duplicate measurements in several individuals (averaged individual s) and some from serial examinations in the same individual (intra-individual s). Examples from a number of sources are shown below, in table 2.4.

Many more, older data are reviewed in ref. 83, but do not contribute to our insight.

Considering this normal, physiological variability of the cholesterol concentration, what is a 'statistically significant change' that we might ascribe to intervention? The probability of such a change being due to chance is great. Blackburn calculated some examples (84) based on an estimated biological standard deviation of 0.40 mmol/l (15 mg%): two single measurements must differ more than 1.1 mmol/l to make the change significant ($p < 0.05$). If these measurements are averages of three analyses 0.6 mmol/l is the smallest significant change. Thus we have good arguments to consider groups of people only and to calculate means of many measurements.

Seasonal variations, if they exist, should have a regular pattern returning every year and they would have to be known in order to interpret changes in cholesterol levels. In the 1961 literature review of Paloheimo (83) there were three accounts of high values during the summer season and two in winter, whereas in seven cases no trends were found. More recently contradictory results have continued to appear.

There are practically no studies in which laboratory fluctuations, although undoubtedly existing, have been taken into account. Some findings are tabulated in table 2.5. As in the intra-individual variations some seasonal data were derived from serial examinations in one group and others from analysis of serum specimens obtained from several population samples large enough to be comparable. In the absence of detailed information about the laboratory error it can only be concluded - particularly from well-controlled methods (95) - that rather irregular fluctuations occur, the cause of which is unknown, and the extent of which is unpredictable. A discussion appears in section 6.6.

2.3. Precision and accuracy: how to select requirements?

From the point of view of the analytical chemist precision and accuracy must be of the highest attainable level. In daily practice the statement of the IFCC Expert Panel on Nomenclature and Principles of Quality Control in Clinical Chemistry is more realistic:

> *'consumer needs - i.e. health care requirements as determined by benefit to patients, clinical practice and cost to the community - must be taken into account, in order to avoid wrong management decisions which might result from reliance on internal laboratory criteria alone; for example, an analytical*

Table 2.4.
Intra-individual biological variation of the serum cholesterol level.

Author ref.	examined persons	type of persons	tests per person	time interval	averaged individuals		intra-individuals	
					mg%	mmol/l	mg%	mmol/l
Anderson (78)	22	healthy soldiers	2	1 week	19.51	0.50		
and	13	healthy obese young men	2	4 days	23.19	0.60		
Keys	24	ambulant coronary patients	2	3 moths	29.61	0.77		
Harris and Cotlove (79)[1)]	68	healthy men and women	10 - 13	1 week	13.1	0.34		
Statland (80)	11	healthy young students	2	1/2 hour	6.8	0.18		
		" " "	3	6 hours	7.6	0.20		
		" " "	5	2 - 4 days	10.6	0.27		
Boerma, (81) KRIS study	23	middle aged men	3	6 months	19.7	0.51		
Thomas (82)[1)]	25	male young prisoners	10 - 12	1 month			10.8 - 49.6	0.28-1.28
Paloheimo (83)	45	policemen	10 -12	1 month			7 - 38	0.18-0.98
	37	prisoners	10 -12	1 month			11 - 36	0.28-0.93

[1]) In addition women over 30 are shown to have a significantly higher variation than average (79) and high cholesterol varies stronger than low levels (82).

Table 2.5.
Seasonal variations of serum cholesterol.

Author	Ref.	Location	Period of time	Examined persons	Same group always or different groups each examination	Maximum [4]	Minimum
Paul	(85)	Chicago	4 years	1982 workers, aged 40 - 55	partly the same	autumn + winter	spring + summer[1]
Pincherle	(86)	London	2.5 years	7133 business men	different	winter	summer
Thomas	(82)	Baltimore	1 year	24 young male prisoners	same	winter	summer + October
Fuller	(87)	London	1 x spring	80 healthy persons	same	no trend	
			1 x autumn		same	no trend	
			1 x summer	73 healthy persons	same	no trend	
Fyfe	(88)	Glasgow	2 years	3701 ambulant and in-patients	different	spring	autumn
Bleiler	(89)	Iowa City	1 year	43 healthy persons	same	no trend[2]	
Paloheimo	(83)	Helsinki	1 year	45 policemen	same	no trend[3]	
				37 prisoners	same	spring	autumn
Boerma	(90)	Rotterdam	1 year	3360 middle aged men	different	no trend[3]	
	(91)	The Netherlands	4 years	27271 persons, aged 20 - 49	different	no trend[3]	
Doyle	(92)	Albany, N.Y.	1 year	53 men	same	autumn + winter	spring + summer

Notes: [1]) at least one year had a reversed pattern; [2]) women and men had different, though slight trends; [3]) monthly, unpredictable, fluctuations are observed; [4])maximum never in summer.

method may be the best available for a given component, but it may still not be good enough for medical application; conversely, a method may be more sensitive, specific or costly than justifiable by its use in a particular clinical situation.' (93)

Of course, it does seem reasonable that health care requirements dictate the analytical needs, but the answer to the question regarding requirements has yet to be given. Several propositions have been made concerning the precision and accuracy limits of serum cholesterol analysis. They are based either on comparisons of laboratory errors with normal ranges in the population, or with individual biological variability or on comparability between laboratories.

The Subcommittee on Criteria of Medical Usefulness of the CAP has given several guidelines to define such requirements (94). Some are of little help, because they perpetuate the present situation:

'Desirable accuracy should be such that the method will create no substantial divergence from generally accepted values for normal and disease states,'

whereas others have gained wide application:

'Desirable precision should be such that errors induced by the measurement process do not significantly widen the range of values for the normal population.'

It is thereby advised that the laboratory standard deviation be within one-twelfth or even one-twentieth of the population range. In these cases the laboratory will add 5.4% or 2.0% respectively, to the normal biological range. In our laboratory the observed serum cholesterol range for 95% of a group of 110 blood donors and laboratory technicians is 3.8 - 8.0 mmol/l (5.9 ± 2.1 mmol/l) and the two criteria mentioned allow for standard deviations of 0.35 and 0.22 mmol/l respectively. Several criteria as proposed in the literature are collected in table 2.6.

Although it is reasonable to consider intra-individual biological variations, these are quite unpredictable and differ widely from person to person. Individual reference values are not easily established. Tonks (95) is mainly concerned with distinguishing the 'healthy' 95% of a population from diseased patients. Cotlove (96) and Barnett (94) are likewise interested in distinguishing a normal physiological variation from pathological changes outside the normal range of homeostasis. Boone (98) and Gilbert (99) have evaluated numerous results from surveys among hundreds (CDC) or thousands of laboratories (QAP - CAP); Eilers' (100) recommendation is the only one that reckons with individual human biologic variations. Very recently it was, in more general terms, recommended that intralaboratory CV should be within

Table 2.6.
Guidelines for laboratory precision and accuracy; cholesterol analysis.

Author	Ref.	Criterium for analytical s	Cholesterol level	Recommendations			
				Accuracy	CV	s intralaboratory	s interlaboratory
Tonks	(95)	allowable limits of error ± 1/4 of normal range or else within 10%			5%		
Barnett	(94)	medical significance	6.5 mmol/l		8%	0.52 mmol/l (20 mg%)	
Cotlove	(96)	1/2 of biological s				0.44 mmol/l (17 mg%)	
Steele	(97)	same				0.34 mmol/l	
Boone	(98)	ref. lab. median ± 1/4 norm. range 'clinical requirement'	e.g. 3 - 5 mmol/l			0.41 mmol/l	
Gilbert	(99)	goal for 1980 in QAP of CAP		4%	6%		
Eilers	(100)	goal for 1980		4 mg% 0.1 mmol/l	3 mg% 0.08 mmol/l		6 mg% (0.16 mmol/l)
Aspen Conference	(101)	s < 1/2 intra-individual s					

1/2 $CV_{intra\text{-}individual}$ (101) without further specification.

In interlaboratory surveys comparability is in the first place strived at and the group mean values must be approached to within ± 2s (99, 102 - 104). In the case that such a group mean by chance coincides with the true value, accuracy is in fact judged in addition to comparability. Sometimes the mean obtained from results of reference laboratories is used as a target value for trials (105). Scoring systems for laboratory proficiency have been suggested in which analytical techniques are judged one by one (106) or as a whole package (107, 108).

In the international cholesterol standardization programmes of CDC and IKEM the narrow limits of the WHO may not be exceeded. In epidemiology, laboratories taking part in population trials are now beginning to make efforts to fulfil these WHO requirements for interlaboratory comparability (table 2.7.). These limits are meant to serve the purpose of pooling data from several studies and do not reflect the requirements derived from 'normal ranges'.

Cholesterol values in large-scale examinations are used often to classify all subjects into risk categories like 'normal', 'borderline' and 'elevated'. This division may be based upon more than one examination and at least the classifications 'borderline' and 'elevated' are as a rule confirmed by repeated screening. For these classifications to be exact the analytical method must be very precise. In the ideal situation the cholesterol measurement must distinguish between 7.2 mmol/l (highest 'borderline' value in CB-Project) and 7.3 mmol/l (lowest 'elevated' value in CB-Project) (chapter 4). Knowing the actual precision of the analysis - the coefficient of variation is 3 - 4% - we can calculate the number of misclassifications. Such calculations are performed in chapter 6. Although 1.5% of all individuals in the CB Heart Study are misclassified the overall group results do not suffer from limited laboratory precision.

Secondly the values are subsequently used to determine whether intervention has been successful or whether 'normal' groups or individuals show a changing cholesterol level over time. In this case biological variations play a more pronounced role and, especially for the follow-up of a group of participants, the analytical variability of the measurement will play a minor role.

When criteria for *precision* have been set it is the task of the laboratory staff to fulfil them. This is entirely possible within the laboratory by selecting a good method and by standardization of its execution. All details of the procedure must be carried out in as much an identical manner day after day as possible. Quality control is necessary to ensure that the level of performance stays within limits (section 2.4.).

Accuracy largely, but not entirely, depends on the specificity of the selected

Table 2.7.
Recommended WHO allowable lipid standardization limits for accuracy and precision (on mean and overall standard deviation).[1)]

CHOLESTEROL

Expected value (EV)		Limit on overall standard deviation		CV	Limit on mean (tolerated deviation from EV)
mg%	mmol/l	mg%	mmol/l	%	EV ± 5%
100	2.59	6.5	0.1681	6.5	EV ± (0.05 x EV)
150	3.88	7.0	0.1810	4.7	-" -
200	5.17	7.5	0.1940	3.8	-" -
250	6.47	8.0	0.2069	3.2	-" -
300	7.76	8.5	0.2198	2.8	-" -
350	9.05	9.0	0.2328	2.6	-" -
400	10.35	9.5	0.2457	2.4	-" -
450	11.64	10.0	0.2586	2.2	-" -
500	12.93	10.5	0.2716	2.1	-" -

TRIGLYCERIDES

		[2)]			mmol/l	± %
35	0.4	5.3	0.06	15.0	EV ± 0.07	17.5
88	1.0	7.1	0.08	8.0	EV ± 0.09	9.0
142	1.6	8.8	0.10	6.3	EV ± 0.11	6.9
195	2.2	10.6	0.12	5.5	EV ± 0.13	5.9
248	2.8	12.4	0.14	5.0	EV ± 0.15	5.4
301	3.4	14.2	0.16	4.7	EV ± 0.17	5.0
354	4.0	15.9	0.18	4.5	EV ± 0.19	4.8
407	4.6	17.7	0.20	4.3	EV ± 0.21	4.6

[1)]Kindly provided by Dr. D. Grafnetter, Prague.
[2)]Expressed as mg% of triolein.

method. The modification of a technique and the way it is performed, especially the calibration procedure, greatly influence the outcome (109). It appears that accuracy - the production of true values - is generally impossible to attain for the average clinical chemistry laboratory. This holds true for almost all analyses in clinical chemistry as a result of the complexity of biological materials and the difficulty of obtaining valid standards for calibration.

2.4. Definitive and reference methods

The basis for evaluation of accuracy is the availability of:

a. pure standard materials

b. accurate methods

The need for pure standards in clinical chemistry is met by a number of 'standard reference materials' of NBS. Among these is SRM-911: cholesterol of more than 99.4% purity (110). The same quality is also commercially available from several sources (111, 112). Serum preparations with assigned values for calibration purposes will be discussed in section 2.5. Usually it is impossible to measure accurately under routine conditions, because our methods have insufficient specificity. Bias can, however, be assessed by comparison with a superior method, carried out at high cost with sophisticated equipment. Such a method, according to the very workable IFCC definition, should have no source of bias. Such a method must also be given a name and this does not stop to initiate many a lengthy discussion and even controversy (101 and section 6.2.). The term definitive method or absolute method do not necessarily imply the exclusion of any improvement nor do they block scientific progress, as is sometimes feared. They too mark a state of the art at the present time. And besides, the development of possibilities for confirmation of the accuracy of a far-away decimal is not of any importance in clinical chemistry. Accuracy of serum cholesterol analysis within ± 0.05 mmol/l is certainly sufficient for the purposes in epidemiological investigations.

The first example of a definitive method in clinical chemistry is the IDMS procedure for calcium in serum (113). A reference method, based on AAS followed (114) as well as an immediate discussion about who assignes such titles as 'reference' and 'definitive' methods (115, 116, 117). The calcium reference method must yield values within ± 2% from the true value (118) and the exhaustive research in the method already has reveiled up to 7.6% instrument dependent bias (119). Remarkably enough a serum calibrator improved this error significantly.

For cholesterol the IDMS technique has also provided the means to obtain accurate results (120 - 123) as well as for numerous other body fluid constituents (124, 125). Research in this area was stimulated by federal regulations of the FDA

in the United States.

Reference methods, that are generally usable and for a long time have been proposed as such, owe specificity to the isolation of cholesterol before it is measured. Such isolation may be performed by precipitation (126) (Sperry and Webb method) or extraction (127) (Abell method) after hydrolysis of the cholesterol esters. Several colour reactions or GLC may then be employed to obtain reasonably accurate results (13). The Abell-Kendall method (127) has for a long time been generally recognized as the most attractive reference method. Alternatives like enzymatic (128) and GLC procedures (129) are under study by several groups of specialists at CDC and in European countries. It must be noted that according to the defined terminology a reference method can only really be considered as such, after verifying its accuracy with the definitive method. Disadvantages of the Abell-Kendall method are that several steps are involved. The shaking device and reagent tube size for instance determine the efficiency of extraction as do timing and temperature. Slight modifications alter the final results as is demonstrated in interlaboratory comparisons and the CDC standardization programme. Surely not everyone's Abell method is the Abell-method (108). It is of interest to note that recent investigations - comparisons between the IDMS, the Abell and enzymatic procedures - led to the finding that Abell results were approximately 1% too low. This was corrected in the Lipid Reference Center by optimising the extraction (74). A remarkable finding indeed after some 20 years of experience with the method. It is doubtful whether this 1% correction, that was introduced with little publicity, is of any benefit at this moment.

In this study the enzymatic method was carried out, with and without blanking procedures, with very much the same precision and a negative bias of a few percent. Further work is in progress at CDC.

2.5. Quality control

Any laboratory issuing analytical results to serve as indicators of healthy or pathological states in a patient must control the reliability thereof. Although the results of clinical chemical investigations are expressed in quantitative terms, they are by no means absolutely correct. The numerous steps occurring between taking a blood sample and logging a final result into the patients record all are potential sources of error and must be guarded. The best insurance for quality is to have dedicated, qualified personnel working with care, paying attention to all necessary details. A quality control system is needed to identify the several sources of error and to determine whether repeatability, reproducibility and, as good as possible, accuracy are such, that useful information is obtained. The literature on

quality control is extensive. Most current textbooks on clinical chemistry contain chapters describing fundamental techniques and lately some monographs have appeared (130 - 133).

Proceedings of conferences, some devoted entirely to the subject of quality control, contain much detailed and recent information (101, 134, 135, 136). The hallmark article of Whitby et al. in Advances in Clinical Chemistry (137) has been recently updated by Whitehead (138). His discussion on terminology, although not unambiguous, and his systematic approach towards limitation of errors are very valuable.

In the following chapters control procedures for cholesterol analysis are described in detail, providing internal and external control schemes. The importance of a normalized blood sampling procedure is stated in section 3.4.1.3 and Appendix I. In our country the clinical chemist is free to organize quality control as he deems fit. Several other countries have legislation that demands specific qualifications of the laboratory director and the personnel and/or require a quality control system. One of the first laws of this kind was the 'Pennsylvania Analytical - Biochemical - Biological Laboratory Act' (act 389) of 1951 (amended 1961, 1972). A laboratory may only be in operation after a permit has been granted. Some regulations in this State now read as follows (139):

'Quality control procedures in chemistry, microbiology, hematology, and other laboratory specialties shall be those approved by the Department.

A degree of accuracy, specificity, and precision satisfactory to the Department shall be shown in quality control records at all times.

Results of proficiency tests shall be maintained in the acceptable ranges statistically determined for each evaluation. Failure of a laboratory to satisfactorily perform in a proficiency test may be cause for revocation of approval of the specific tests involved.

A clinical laboratory shall be prepared at all times during the normal working hours to accept, perform, and report promptly on all specimens submitted by the Department for purpose of testing the adequacy and accuracy of its procedures.'

The federal 'Clinical Laboratories Improvement Act' (1967) now makes similar requirements effective in the entire USA. Proficiency testing is carried out by CDC, including shipments of control serum as well as on-site inspections. European coun-

tries like France, the United Kingdom and the German Federal Republic require participation in proficiency testing or specify qualifications of the laboratory director. It must be added that in these cases law was intended to put an end to undesirable situations unknown in the Netherlands. Analytical chemical techniques do appear of course in some other Dutch laws. The Food and Drug Act (Warenwet) specifies exactly which techniques must be used by Food Inspection Departments. In contrast with this the Plant Protection Law (Plantenziektewet) authorizes the Plant Protection Service (Plantenziektekundige Dienst) to select her own methods to detect diseases. Thus the presence of malignant virus infections, when it is detected by enzyme immuno assay, will automatically prohibit the commerce of the entire lot of checked plants.

2.6. Reference laboratories

2.6.1. CDC

In July 1946, the Communicable Disease Center began to coordinate the attack on contageous and vector-borne diseases in the USA. It had already played an important role in the eradication of malaria in Florida and was, for that purpose, headquartered in Atlanta, Ga. Today it has grown into one of the large US Federal Health Service Departments. The name was changed to Center for Disease Control in 1970 to emphasize its position as leading institution for State Health Departments, clinical laboratories, prevention programmes, nutritional studies and many more action lines could be named (140). Its Bureau of Laboratories coordinates work in all possible kinds of health oriented research, from occupational safety problems to a search for exotic microorganisms from outer space.

In the Clinical Chemistry Division, the Lipid Standardization Laboratory is carrying out the WHO Cooperative Cholesterol and Triglycerides Standardization Programme. Since an experimental programme for cholesterol was initiated in 1957 the facilities were continuously enlarged and in the past 20 years over 800 laboratories took part in this CCSP (108, 141). The programme begins with 'part 1', which is a self-evaluation phase. All participants check their performance and improve upon it when necessary. Samples with known assigned values are sent to them. In 'part 2' proficiency is checked with samples having unknown lipid levels.

Advice and additional samples are provided whenever problems should arise. 'Part 3' is a continuing control phase to ensure sufficient long term quality. A certificate may be granted after a period of time has elapsed in which good proficiency has been demonstrated. This certification is therefore a retrospective one.

Today the programme is called Cooperative Cholesterol and Triglycerides

Standardization Program and it is, among other international standardization efforts, an important example.

2.6.2. IKEM

The World Health Organization has been coordinating and initiating research on cardiovascular diseases in several countries which has been one of the important drives for standardization - WHO requested participating laboratories to become standardized in much the same way the US Government did in Federal health projects. A WHO Regional Lipid Reference Centre for Europe was established in 1973 in Prague, Czechoslovakia, to meet the growing demands for reference work in Europe (142). It is situated in the Institute for Clinical and Experimental Medicine (IKEM). The programme is very similar to that of CDC, but is not strictly divided into three parts. Another difference is the use of lyophilized control pools, commercially obtained or own made. In addition to the IKEM reference method, the identical manual Abell procedure at CDC is used to establish assigned values. Today, over 120 laboratories in 47 countries are cooperating in this Regional Programme.

2.6.3. The Netherlands Lipid Reference Laboratory

The growing number of laboratories seeking assistance with standardization means an ever increasing burden upon the two WHO Reference Laboratories. This includes the shipping of expensive serum samples without payment from the customers. Recently CDC has recommended that national coordination be improved in order to channel interlaboratory contacts between a particular country and CDC's Lipid Reference Section through one national centre (141).

In the Netherlands a discussion about pros and cons of one or more reference laboratories has hardly begun. But to serve the needs for our cholesterol standardization plans a reference laboratory is in the founding stage in the Rotterdam University Hospital Department of Clinical Chemistry. Its position fits in with the reference task which the RIV considers its duty by a mutual agreement reached in consultation with the KCA foundation.

This reference laboratory has now duplicated the manual Abell-Kendall method as it is employed at CDC, without changes. In the near future our reference method will be used to establish assigned values to serve as our basis for checking accuracy of cholesterol determinations. Our draft proposal to set up a reference system in the European Community awaits approval of the BCR in Brussels (143).

2.7. Standardization and comparability

2.7.1. The cholesterol standardization programmes currently in existence differ from other interlaboratory surveys in several ways:

1. The values obtained by the labeling procedure in the reference laboratory may for all practical purposes be considered as true values.
   Participants therefore can establish bias, provided that samples are used with chemical and physical matrix properties closely resembling those of fresh human serum (57).
2. The two reference laboratories will in case of difficulties always be prepared to give advice and to make additional control samples available. They are actively engaged in helping the participants to solve their problems. (142).
3. The surveys are focusing on only one or two analytes and participants must devote a relatively large amount of time to these techniques.
4. The allowable limits for bias and standard deviation are fixed and used uniformly throughout the world (table 2.6.).

The results of these programmes and other interlaboratory trials are similar in that the introduction of better analytical techniques and effective quality control procedures is stimulated. This has led to improved comparability among laboratories. All publications about external quality control schemes report improved interlaboratory precision over the past decade, but improvement with cholesterol is usually small (144 - 147).

2.7.2. Comparability of published data on cholesterol analysis is severly limited by lack of standardization. Only few papers report having employed a method that took part in WHO international standardization programmes (examples are: 148 - 151). Some have made intralaboratory comparisons with reference methods but such results form a less reliable basis (108) because of bias among the Abell modifications. For instance, in a paper by Keys et al. (152) cholesterol levels in different studies were compared (table 2.8.).
It is then stated:

'Essentially identical methods were used in all of these studies and there is no reason to ascribe differences in the results to the methods used.'

There may well be differences between Chicago and Minnesota men, but direct comparisons were not made and this statement is not valid. It would be better to suspect between-method differences unless their absence is proven (153). Even if mean values of two methods do seem to be alike, the correctness of it must be con-

Table 2.8. Serum cholesterol levels found in some studies.

	Percentage of population with specified serum cholesterol level:	
	< 220 mg% < 5.69 mmol/l	> 260 mg% > 6.72 mmol/l
Minnesota men, aged 45 - 55	47%	20%
Chicago men, aged 40 - 59	26%	47%
etc.		

firmed in comparative measurements (154).

Sometimes repeated measurements *within* the same study are not comparable because of method changes: e.g. Busselton, Australia (155).

2.7.3. Are the WHO narrow limits a sufficient guarantee that comparing or pooling of data is possible? The answer must be: no; and it may be illustrated by an example.

Suppose laboratory A and laboratory B participate in a screening programme. Their bias, employing the same method, is + 4% and - 4% respectively and the mean values obtained in two control pools are compared (table 2.9.):

Table 2.9. Analytical results of two laboratories with + 4% and - 4% bias.

Serum	Laboratory A	Reference laboratory	Laboratory B
pool X	7.28 mmol/l	7.00 mmol/l	6.72 mmol/l
pool Y	6.10 mmol/l	5.86 mmol/l	5.67 mmol/l

If laboratory A uses cut-off levels of 6.15 and 7.25 mmol/l, as is done in the CB Heart Project, the corresponding cut-off levels in laboratory B must be 5.65 and 6.75 mmol/l respectively. But both laboratories work *within* WHO narrow limits. In the CB Heart Project during a certain time period 6006 men and women were screened and classified as 'normal', 'borderline' and 'elevated' by laboratory A (Table 6.6). Had this been done by laboratory B a different impression would have been obtained altogether: table 2.10.

Table 2.10. Sample of 6006 persons grouped by serum cholesterol.

laboratory	number of persons in three categories (cholesterol level: x)		
	$x < 6.2$ mmol/l 'normal'	$6.3 \leqslant x \leqslant 7.3$ 'borderline'	$x > 7.3$ mmol/l 'elevated'
A	4398 (73.0%)	1177 (19.5%)	431 (7.5%)
B	5076 (84.5%)	720 (12.0%)	210 (3.5%)

It may be concluded that for interlaboratory comparison the bias of the laboratories must be known, and included in published results.

If the bias is unknown, both laboratories may consider to designated the upper 5%, or 10%, of the group as 'elevated'. Compare section 6.7.7. in which the need for standardization is discussed for different kinds of study set-ups.

2.8. Examples of standardization: CB Heart Project and KRIS

In the next two chapters the standardization and quality control procedures will be described, that have been parts of an international and a national cardiovascular epidemiological study: Chapter 3: The Kaunas Rotterdam Intervention Study, (KRIS) and Chapter 4: The CB Heart Project.

2.8.1. In the KRIS comparability between the two participating laboratories was ensured by keeping all procedures and equipment identical in Rotterdam and Kaunas. From the blood sampling to the reporting of results all methods and materials were exactly alike. The laboratory results were compared by employing common serum controls and by exchanging frozen patient specimens.

To estimate the accuracy of our method a special series of surveys was organized and during the 2.5 years of the KRIS-project eleven survey-sets of a dozen samples of human reference serum were analyzed and the results compared with those from the Lipid Reference Laboratory in the Center for Disease Control (Dr. A. Mather). A part of the results has been published elsewhere (156, 157).

2.8.2. The CB Heart Project has only one central laboratory, the Central Clinical Chemistry Laboratory of the University Hospital-Dijkzigt in Rotterdam. To make the study results comparable with others, the survey samples from the WHO Regional Lipid Reference Centre in Prague (Dr. D. Grafnetter) were analyzed from 1973 onwards. Results have been published, in part, elsewhere (158, 159, 160).

2.8.3. In both projects several problems were encountered, some could be solved easily and some, unexpectedly, demonstrated that our cholesterol analysis remains a difficultly controllable process. Its sensitivity to slight changes in the analytical procedure and to the sources of the serum pools for example, makes the measurement of bias a hard task.

2.9. Standardization in the Netherlands: KCA

In the Netherlands numerous epidemiological studies have been and are carried out (e.g. 161 - 167).

Screening of population samples occurred in studies carried out through university research projects, or by professional organisations, the Government and the food-industry. The laboratories of the coordinated study by the Commission for Detection and Prevention of Ischaemic Heart Disease (COPIH) of the NVAB have been participating in comparative measurements. These surveys were set up by the Central Laboratory TNO in Delft (168) and comparability among the group - all employing the same Huang modification (14) - was thus assessed. Acceptability criteria were not established. It was decided to pursue the desired standardization by an organization especially devoted to the task of making the analytical results obtained in Dutch epidemiological studies comparable. This Foundation for Quality Control of Chemical Analysis in Epidemiological Investigations (KCA) aims at increasing precision and accuracy by offering a standardization programme set up like the programmes of CDC, Atlanta, USA and IKEM, Prague, Czechoslovakia. In addition to this the KCA plans (in collaboration with the RIV) to make serum calibrators available to the participants in an attempt to decrease the positive bias of direct LB-reactions. A preliminary experiment has yielded rather satisfactory results. Chapter 5 describes in more detail what has been accomplished so far in this national standardization project.

REFERENCES

Chapter 2

1. Biol. Abstr. *59*, (Jan. - June) 1975.
2. J. R. SABINE: Cholesterol, Marcel Dekker Inc., New York (1977).
3. C. E. BILLS: Physiology of the sterols, including vitamin D. Physiol. Rev. *15*, 1 - 97 (1935).
4. L. F. FIESER and M. FIESER: Steroids. 2nd. ed., 403 - 419. Reinhold Publishing Corp., New York (1967).

5. H. W. FLEGG: An investigation of the determination of serum cholesterol by an enzymatic method. Ann. Clin. Biochem. *10*, 79 - 84 (1973).
6. W. RICHMOND: Preparation and properties of a cholesterol oxidase from Nocardia sp. and its application to the enzymatic assay of total cholesterol in serum. Clin. Chem. *19*, 1350 - 1356 (1973).
7. B. ZAK and N. RESSLER: Methodology in determination of cholesterol; a review. Am. J. Clin. Path. *25*, 433 - 446 (1955).
8. R. CROCKETT, M. DALLOCHIO, J.-P. BROUSTET, N. SAGARDILUZ and H. BRICAUD: Comparaison des résultats du dosage du cholestérol total par diverses méthodes appliquées directement au sérum et après extraction. Intérêt de la méthode de Ferro et Ham. Arch. Mal. du Coeur *56*, Suppl. 2, 55 - 87 (1963).
9. G. VANZETTI: Méthodes photométriques de dosage du cholestérol dans le sérum. Clin. Chim. Acta *10*, 389 - 405 (1964).
10. D. B. TONKS: The estimation of cholesterol in serum: a classification and critical review of methods. Clin. Biochem. *1*, 12 - 29 (1967).
11. B. ZAK: Cholesterol methodologies: a review. Clin. Chem. *23*, 1201 - 1214 (1977).
12. F. L. MITCHELL: Present and future trends of automation in clinical chemistry. 7th International Congress of Clinical Chemistry, Vol. I. Methods in clinical chemistry. M. Roth ed. S. Karger, Basel (1970).
13. T. C. HUANG, C. P. CHEN, V. WEFLER and A. RAFTERY: A stable reagent for the Liebermann-Burchard reaction. Anal. Chem. *33*, 1405 - 1407 (1961).
14. NEDERLANDS NORMALISATIE-INSTITUUT: Procedure NEN 2415, Klinisch-chemisch onderzoek. Cholesterol (totaal) in serum, fotometrisch zonder extractie (1975).
15. R. DULOU, J. CHOPIN and Y. RAOUL: Structure de deux bischolestadiènes isolés de la réaction de Salkowski sur le cholestérol. Bull. Soc. Chim. France *18*, 616 - 621 (1951).
16. C. H. BRIESKORN and L. CAPUANO: Der Chemismus der Farbreaktionen nach Liebermann und Salkowsky bei Triterpenen und Sterinen. Chem. Berichte *86*, 866 - 873 (1953).
17. T. WATANABE: Coloured products and intermediates of the Liebermann-Burchard reaction and its mechanism. Chem. Abstr. *55*, 10503^{c} (1961).
18. W. LANGE, R. G. FOLZENBOGEN and D. G. KOLP: Colorless crystalline perchloric and hexafluorophosphoric acid salts of sterols. The color formation between sterols and perchloric acid. J. Am. Chem. Soc. *71*, 1733 - 1737 (1949).
19. N. W. TIETZ: Fundamentals of Clinical Chemistry. W. B. Saunders Co. Philadelphia (1970). In the 2nd edition (1976) the newer mechanism is given.
20. R. W. BURKE, B. I. DIAMONDSTONE, R. A. VELAPOLDI and O. MENIS:

Mechanisms of the Liebermann-Burchard and Zak color reactions for cholesterol. Clin. Chem. *20*, 794 - 801 (1974).
21. R. A. VELAPOLDI, B. I. DIAMONDSTONE and R. W. BURKE: Spectral interpretation and kinetic studies of the Fe^{3+}-H_2SO_4 (Zak) procedure for determination of cholesterol. Clin. Chem. *20*, 802 - 811 (1974).
22. A. C. PAREKH, C. SIMS, R. W. FONG, M. CARANTO and D. H. JUNG: Steroid interference in iron-cholesterol reactions: a comparative study. Steroids *25*, 525 - 533 (1975).
23. B. ZAK, L. M. WEINER and B. WELSH: Spectrophotometric study of bilirubin interference in the Huang reaction for cholesterol. Clin. Chim. Acta *30*, 697 - 705 (1970).
24. P. B. SOMMERS, P. I. JATLOW and D. SELIGSON: Direct determination of total serum cholesterol by use of double wavelength spectrophotometry. Clin. Chem. *21*, 703 - 707 (1975).
25. K. NAKANISHI, B. K. BHATTACHARYYA and L. F. FIESER: Cholesterol and companions. V. Microdetermination of Δ^7-sterols. J. Am. Chem. Soc. *75*, 4415 - 4417 (1953).
26. L. F. FIESER: A companion of cholesterol. J. Am. Chem. Soc. *73*, 5007 (1951).
27. F. R. FRANEY and E. ARMADOR: Serum cholesterol measurement based on ethanol extraction and ferric chloride - sulfuric acid. Clin. Chim. Acta *21*, 255 - 263 (1968).
28. R. RICHTERICH: Clinical Chemistry, Theory and Practice. S. Karger, Basel (1969).
29. W. RICK: Klinische Chemie und Mikroscopie. 5th rev. ed. Springer Verlag Berlin (1977).
30. R. J. HENRY, D. C. CANNON and J. W. WINKELMAN: Clinical Chemistry, Principles and Technics. 2nd. ed. Harper & Row, New York (1974).
31. L. LILIENBERG and A. SVANBORG: Determination of plasma cholesterol: comparison of gas-liquid chromatographic, colorimetric and enzymatic analyses. Clin. Chim. Acta *68*, 223 - 233 (1976).
32. D. J. MUNSTER, M. LEVER and R. W. CARRELL: Contributions of other sterols to the estimation of cholesterol. Clin. Chim. Acta *68*, 167 - 175 (1976).
33. A. E. KOEHLER and E. HILL: 7-Dehydrocholesterol in human serum. Fed. Proc. *12*, 232 - 233 (1953).
34. D. S. YOUNG, L. C. PESTANER and V. GIBBERMAN: Effects of drugs in clinical laboratory tests. Clin. Chem. *21*, (5), 1D - 432D (1975).
35. N. V. CONSTANTINO and H. F. KABAT: Drug-induced modifications of laboratory test values. Am. J. Hosp. Pharm. *30*, 24 - 71 (1973).
36. A. C. ARNTZENIUS, C. M. VAN GENT, H. VAN DER VOORT, C. I. STE-

GERHOEK and K. STYBLO: Reduced high-density lipoprotein in women, aged 40 - 41, using oral contraceptives. Lancet *1*, 1221 - 1223 (1978).
37. W. D. BROWN: Errors in the determination of serum cholesterol. Aust. J. Exp. Biol. *39*, 209 - 222 (1961).
38. G. KURZWEG and W. MASSMANN: About the differences of extinction of the free and esterified cholesterol using a few simple methods of estimation. Clin. Chim. Acta *7*, 515 - 521 (1962).
39. D. WEBSTER: The direct estimation of esterified cholesterol with a number of commonly used cholesterol colour reagents. Clin. Chim. Acta *8*, 731 - 743 (1963).
40. L. VASS: Wässrige primäre Cholesterinstandardlösung mit Cholesteringehalt im oberen pathologischen Bereich. Zubereitung und Bedeutung. Ärztl. Lab. *20*, 148 - 151 (1974).
41. A. P. KENNY: The determination of cholesterol by the Liebermann-Burchard reaction. Biochem. J. *52*, 611 - 619 (1952).
42. A. T. NESS, J. V. PASTEWKA and A. C. PEACOCK: Evaluation of a recently reported stable Liebermann-Burchard reagent and its use for the direct determination of serum total cholesterol. Clin. Chim. Acta *10*, 229 - 237 (1964).
43. E. H. MARTENSSON: Investigation of factors affecting the Liebermann-Burchard cholesterol reaction. Scand. J. clin. Lab. Invest. Suppl. *69*, 164 - 180 (1963).
44. R. F. WITTER and G. R. COOPER: Cholesterol. ACSP Commission on Continuing Education: Clinical Chemistry No. CC 73, 8 - 12 (1972).
45. P. N. M. DEMACKER: Hyperlipoproteinemias; a study on diagnostic methods. Doctoral thesis, 65 - 66, Nijmegen (1978).
46. C. M. VAN GENT: Gaubius Institute, Leiden, personal communication (1975).
47. W. M. SPERRY and F. C. BRAND: The colorimetric determination of cholesterol. J. Biol. Chem. *150*, 315 - 324 (1943).
48. S. PEARSON, S. STERN and T. H. McGAVACK: A rapid, accurate method for the determination of total cholesterol in serum. Anal. Chem. *25*, 813 - 814 (1953).
49. A. RAPPOPORT and R. EICHHORN: Sulfosalicylic acid as a substitute for paratoluene sulfonic acid. A. In estimation of cholesterol. B. In the diagnostic test for systemic lupus erythematosus. Clin. Chim. Acta *5*, 161 - 163 (1960).
50. D. R. WYBENGA, V. J. PILEGGI, P. H. DIRSTINE and J. DiGEORGIO: Direct manual determination of serum total cholesterol with a single stable reagent. Clin. Chem. *16*, 980 - 984 (1970).
51. J. LEVINE, S. MORGANSTERN and D. VLASTELICA: A direct Liebermann-Burchard method for serum cholesterol. Automation in Analytical Chem-

istry; Technicon Symposia 1967, Vol. 1., 25 - 28, Mediad Inc., White Plains, N.Y. (1968).
52. T. E. HEWITT and H. L. PARDUE: Kinetics of the cholesterol-sulfuric acid reaction: A fast kinetic method for serum cholesterol. Clin. Chem. *19*, 1128 1128 - 1134 (1973).
53. H. L. VADER and C. L. J. VINK: The influence of viscosity on dilution methods. Its problem in the determination of serum sodium. Clin. Chim. Acta *65*, 379 - 388 (1976).
54. G. PALME, M. KOCHEN and H. LEONHARDT: Veränderungen der Fliesseigenschaft des Blutes bei chronischen Lebererkrankungen. Clin. Chim. Acta *61*, 163 - 169 (1975).
55. M. R. ESSEVELD: Viscositeit van bloedplasma en -serum. Doctoral thesis, Rotterdam (1972).
56. A. P. JANSEN, E. J. VAN KAMPEN, C. A. M. MEYERS, P. J. J. VAN MUNSTER and G. J. M. BOERMA: Quality control and the quality of commercial test sera. Clin. Chim. Acta *84*, 255 - 258 (1978).
57. IFCC: Provisional recommendation on quality control in clinical chemistry. Part 3: Calibration and control materials. Clin. Chim. Acta *75*, F11 - F20 following p. 180 (1977).
58. G. F. GRANNIS and J. A. LOTT: An interlaboratory comparison of analyses of clinical specimens. Am. J. Clin. Pathol. *70*, 567 - 576 (1978).
59. S. S. BROWN: Notes on the quality of performance of serum cholesterol assays. Ann. Clin. Biochem. *10*, 146 - 154 (1973).
60. J. H. WILLIAMS, L. TAYLOR, M. KUCHMAK and R. F. WITTER: Preparation of hypercholesteremic and/or hypertriglyceridemic sera for lipid determinations. Clin. Chim. Acta *28*, 247 - 253 (1970).
61. R. F. WITTER, M. KUCHMAK, J. H. WILLIAMS, V. S. WHITNER and C. L. WINN: Lipids of commercial serum products offered as controls or 'standards' for cholesterol and triglyceride determinations. Clin. Chem. *16*, 743 - 748 (1970).
62. Annual quality control report, Hypertension Detection and Follow-up Project, May 1977. Kindly made available by Dr. G. L. Payne, Nat. Heart, Blood and Lung Institute, Bethesda, Md. (1977).
63. L. ELDJARN and C. C. GILHUUS-MOE: On the need for serum-matrixed reference solutions stating absolute values in the strive towards accuracy in clinical chemistry. Lecture for the Japanese Association for Clin. Chem. (1978). Kindly made available by Dr. A. P. Jansen.
64. D. C. HOHNADEL, F. W. SUNDERMAN, P. TERHUNE, F. H. REID and I. H. POMPER: Comparisons of the precision of replicate analyses of frozen and lyophilized quality control serums. Ann. Clin. Lab. Sci. *3*, 335 - 340 (1973).
65. G. J. PROKSCH and D. P. BONDERMAN: Use of cholesterol-rich bovine lipo-

protein to enhance cholesterol concentration in the preparation of serum control materials. Clin. Chem. *22*, 1302 - 1305 (1976).
66. W. F. FRAJOLA and J. MAURUKAS: A stable liquid human reference serum. Hosp. Lab. Sci. *13*, 25 - 33 (1976).
67. College of American Pathologists: Standards for an Interlaboratory (proficiency) testing programme. Am. J. Clin. Pathol. *66*, 276 - 278 (1976).
68. B. ZAK: A paradox clarified: set-point values for cholesterol, direct and enzymatic. Clin. Chem. *23*, 1940 (1977).
69. A. NOMO and K. NAKAYAMA: Comparative studies on cholesterol oxidases from different sources. Clin. Chim. Acta *73*, 487 - 496 (1976).
70. W. H. COPELAND, J. P. BRETAUDIÈRE and R. REJ: Measurement of cholesterol in serum: an interlaboratory study of factors influencing accuracy and precision. Clin. Chem. *24*, 1046 (1978).
71. M. A. PESCE: Investigation of interfering substances with several enzymatic systems for measuring cholesterol. Clin. Chem. *22*, 1194 (1976).
72. A. M. DE BRUIJN, F. KLEIN, C. M. VAN GENT and H. A. VAN DER VOORT: On the calibration of automated enzymatic cholesterol determination. Clin. Chim. Acta *85*, 215 - 218 (1978).
73. G. R. COOPER, J. S. HAZLEHURST and P. H. DUNCAN: Evaluation of cholesterol standards used for enzymatic determinations. Clin. Chem. *23*, 1153 (1977).
74. G. R. COOPER, J. R. AKINS, J. BIEGELEISEN, J. S. HAZLEHURST, G. M. MYERS and D. D. BAYSE: Comparison of selected cholesterol colorimetric and enzymatic reference methods. Clin. Chem. *24*, 1046 (1978).
75. This method is usually employed for ChOD assay: C. C. ALLAIN, L. S. POON, C. S. G. CHAN, W. RICHMOND and R. C. FU: Enzymatic determination of total serum cholesterol. Clin. Chem. *20*, 470 - 475 (1974).
76. J. M. DIETSCHY, L.E. WECKS and J. J. DELENTE: Enzymatic measurement of free and esterified cholesterol levels in plasma and other biological preparations using the oxygen electrode in a modified glucose analyzer. Clin. Chim. Acta *73*, 407 - 414 (1976).
77. N. N. REHAK and D. S. YOUNG: Enzymatic determination of serum free cholesterol and cholesterol ester by kinetic colorimetry. Clin. Chem. *23*, 1153 (1977).
78. J. T. ANDERSON and A. KEYS: Cholesterol in serum and lipoprotein fractions. Its measurement and stabilities. Clin. Chem. *2*, 145 - 159 (1956).
79. E. K. HARRIS, P. KANOFSKY, G. SHAKARJI and E. COTLOVE: Biological and analytical components of variation in long-term studies of serum constituents in normal subjects. II. Estimating biological components of variation. Clin. Chem. *16*, 1022 - 1027 (1970).
80. B. E. STATLAND and P. WINKEL: Variations of cholesterol and total lipid

concentrations in sera of healthy young men. Differentiating analytical error from biologic variability. Am. J. Clin. Pathol. *66*, 935 - 943 (1976).
81. Experiments carried out in 1976, in cooperation with Mrs. A. L. Dubbeldam-Marree, MD. To be published in KRIS Final Report; see section 3.5.1.
82. C. B. THOMAS, H. W. D. HOLLJES and F. F. EISENBERG: Observation on seasonal variations in total serum cholesterol level among healthy young prisoners. Ann. Int. Med. *54*, 413 - 430 (1961).
83. J. PALOHEIMO: Seasonal variations of serum lipids in healthy men. Ann. Med. Exp. Biol. Fenn. *39*, Suppl. 8 (1961). Doctoral thesis, Helsinki 1961.
84. H. BLACKBURN: Coronary risk factors. How to evaluate and manage them. Eur. J. Card. *1*, 249 - 283 (1975).
85. D. PAUL, M. H. LEPPER, W. H. PHELAN, G. W. DUPERTUIS, A. MACMILLAN, H. McKEAN and H. PARK: A longitudinal study of coronary heart disease. Circulation *28*, 20 - 31 (1963).
86. G. PINCHERLE: Factors affecting the mean serum cholesterol level. J. Chron. Dis. *24*, 289 - 297 (1971).
87. J. H. FULLER, S. L. GRAINGER, R. J. JARRETT and H. KEEN: Possible seasonal variation of plasma lipids in a healthy population. Clin. Chim. Acta *52*, 305 - 310 (1974).
88. T. FYFE, M. G. DUNNINGHAM, E. HAMILTON and R. J. RAE: Seasonal variation in serum lipids, and incidence and mortality of ischaemic heart disease. J. Atheroscler. Res. *8*, 591 - 596 (1968).
89. R. E. BLEILER, E. S. YEARICK, S. S. SCHNUR, I. S. SINGSON and M. A. OHLSON: Seasonal variation of cholesterol in serum of men and women. Am. J. Clin. Nutr. *12*, 12 - 16 (1963).
90. See KRIS results, section 3.5.2., figure 3.3.
91. See CB Heart Project, section 4.6., figure 4.4.
92. J. T. DOYLE, S. H. KINCH and D. F. BROWN: Seasonal variation in serum cholesterol concentration. J. Chron. Dis. *18*, 657 - 664 (1965).
93. IFCC Provisional recommendations on quality control in clinical chemistry. Part 6. Quality requirements from the point of view of health care. Clin. Chem. *23*, 1066 - 1071 (1977).
94. R. N. BARNETT: Clinical laboratory statistics. Little, Brown and Co. Boston (1971).
95. a. D. B. TONKS: A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. Clin. Chem. *9*, 217 - 233 (1963).
    b. Letter to the Ed. Clin. Chem. *10*, 628 - 630 (1964).
96. E. COTLOVE, E. K. HARRIS and G. Z. WILLIAMS: Biological and analytical components of variation in long-term studies of serum constituents in normal subjects. III. Physiological and medical implications. Clin. Chem. *16*,

1028 - 1032 (1970).
97. B.W. STEELE, M. K. SCHAUBLE, J. M. BECKTEL and J. E. BEARMAN: Evaluation of clinical chemistry laboratory performance in twenty Veterans Administration Hospitals. Am. J. Clin. Pathol. *67*, 594 - 602 (1977).
98. D. J. BOONE: Proficiency testing summary analysis. Clinical Chemistry 1977 II. US Dept. HEW, CDC, Atlanta Ga. 30333 (1977).
99. R. K. GILBERT: Progress and analytical goals in clinical chemistry. Am. J. Clin. Pathol. *63*, 960 - 973 (1975).
100. R. J. EILERS: Quality assurance in health care: missions, goals, activities. Clin. Chem. *21*, 1357 - 1367 (1975).
101. J. H. BOUTWELL ed.: A national understanding for the development of reference materials and methods for clinical chemistry. Proceedings of a conference. AACC, Washington D.C. (1978).
102. IFCC: Provisional recommendation on quality control in clinical chemistry. Part 5. External quality control. Clin. Chim. Acta *83*, 189F - 202F, following p. 188 (1978).
103. B. E. COPELAND and R. M. ROSENBAUM: Organization, planning and results of the Massachusetts Society of Pathologists Regional Quality Control Program. Am. J. Clin. Pathol. *57*, 676 - 688 (1972).
104. G. F. GRANNIS, H.-D. GRÜMER, J. A. LOTT, J. A. EDISON and W. C. McCABE: Proficiency evaluations of clinical chemistry laboratories. Clin. Chem. *18*, 222 - 236 (1972).
105. G. RÖHLE, H. BREUER and G. OBERHOFFER: Collaborative surveys of the German Society for Clinical Chemistry. Mitt. Dtsch. Ges. Klin. Chem., *5*, 43 - 56 (1974).
106. A. P. JANSEN, E. J. VAN KAMPEN, B. LEIJNSE, C. A. M. MEIJERS and P. J. J. VAN MUNSTER: Experience in the Netherlands with an external quality control and scoring system for clinical chemistry laboratories. Clin. Chim. Acta *74*, 191 - 201 (1977).
107. T. P. WHITEHEAD, D. M. BROWNING and A. GREGORY: A comparative survey of the results of analyses of blood serum in clinical chemistry laboratories in the United Kingdom. J. Clin. Path. *26*, 435 - 445 (1973).
108. D. C. H. LEY and S. EZER: The development of an interlaboratory proficiency testing programme for the province of Ontario. I. A preliminary survey of clinical chemistry. II. Design of performance graphs. Clin. Biochem. *7*, 223 - 250 (1974).
109. G. R. COOPER: The WHO-CDC Lipid Standardization Programme in: Quality Control in Clinical Chemistry. Transactions of the VIth International Symposium. G. Anido, E. J. van Kampen, S. B. Rosalki, E. Rubin, eds. W. de Gruyter, New York (1975).
110. F. L. MITCHELL: Selected, reference and definitive methods for clinical chem-

istry. Ann. Biol. Clin. *33*, 168 - 171 (1975).
111. J. H. WILLIAMS, M. KUCHMAK and R. F. WITTER: Evaluation of the purity of cholesterol primary standards. Clin. Chem. *16*, 423 - 426 (1970).
112. A. J. BARNARD, E. F. JOY and F. W. MICHELOTTI: Characterization of high-purity chemicals for use as clinical laboratory standards. Clin. Chem. *17*, 841 - 850 (1971).
113. L. J. MOORE and L. A. MACHLAN: High accuracy determination of calcium in blood serum by isotope dilution mass spectrometry. Anal. Chem. *44*, 2291 - 2296 (1972).
114. J. P. CALI, G. N. BOWERS and D. S. YOUNG et al.: A reference method for the determination of total calcium in serum. Clin. Chem. *19*, 1207 - 1213 (1973).
115. D. W. NEILL and J. R. DUGGART: When is a reference method a reference method. Clin. Chem. *22*, 285 (1976).
116. S. S. BROWN, J. BUTTNER, F. L. MITCHELL, M. RUBIN and G. R. COOPER: Collective response. Clin. Chem. *22*, 285 - 287 (1976).
117. J. P. CALI, R. SCHAFFER and F. L. MITCHELL: Suggested modification of the definition for definitive method. Clin. Chem. *23*, 1943 - 1944 (1977).
118. J. P. CALI, G. N. BOWERS, D. S. YOUNG: A reference method for the determination of total calcium in serum. In: Standard methods of clinical chemistry. Vol. 8., G. R. Cooper, ed. AACC, Washington D. C. (1977).
119. B. E. COPELAND, D. W. GRISLEY, J. CASELLA and H. BAILEY: Comparison of serum calcium measurements with respect to five models of atomic absorption spectrophotometers using the NBS-AACC calcium reference method and isotope-dilution mass spectrometry as the definitive method. Am. J. Clin. Pathol. *66*, 619 - 633 (1976).
120. L. BJORKHEM, R. BLOMSTRAND and L. SVENSON: Serum cholesterol determination by mass fragmentography. Clin. Chim. Acta *54*, 185 - 193 (1974).
121. L. SIEKMAN, K. P. HUSKES and H. BREUER: Determination of cholesterol in serum using mass fragmentography - A reference method in clinical chemistry. Z. Anal. Chem. *279*, 145 - 146 (1976).
122. R. SCHAFFER, E. WHITE, L. T. SNIEGOSKI and A. COHEN: A definitive method for the determination of total cholesterol in serum. Proceedings of the 9th national research symposium, NBS (1978). Preprint kindly made available by Dr. R. Schaffer.
123. A. RIGERO (ed.): Advances in Mass Spectrometry in Biochemistry and Medicine, Vol. II, Spectrum Publ. Inc., New York (1976).
124. I. BJORKHEM, R. BLOMSTRAND, O. LANTLO and G. OHMAN: Toward absolute methods in clinical chemistry: application of mass fragmentography to high-accuracy analysis. Clin. Chem. *22*, 1789 - 1801 (1976).

125. I. R. POLITZER, B. J. DOWTY and J. L. LASETER: Use of gas chromatography and mass spectrometry to analyze underivatized volatile human or animal constituents of clinical interest. Clin. Chem. *22*, 1775 - 1788 (1976).
126. W. M. SPERRY and M. WEBB: A revision of the Schoenheimer-Sperry method for cholesterol determination. J. Biol. Chem. *187*, 97 - 106 (1950).
127. L. L. ABELL, B. B. LEVY, B. B. BRODIE and F. E. KENDALL: A simplified method for the estimation of total cholesterol in serum and demonstration of its specificity. J. Biol. Chem. *195*, 357 - 366 (1952).
128. D. D. BAYSE and G. R. COOPER : personal communication 1978. See also ref 74.
129. J. P. AMBERT, A. CAHOUR and L. HARTMANN: Dosage en deux temps au cholesterol total et du cholesterol libre plasmatiques par chromatographie gas-liquide. Proposition d'une méthode de reference. Clin. Chim. Acta *68*, 31 - 41 (1976).
130. T. P. WHITEHEAD: Quality control of clinical chemical analysis. Wiley, New York (1976).
131. M. DHARAN: Total quality control in the clinical laboratory. C. V. Mosby Co., St. Louis (1977).
132. P. J. OTTAVIANO and A. F. DISALVO: Quality control in the clinical laboratory: a procedural text. University Park Press, Baltimore (1977).
133. D. STAMM: Qualitätskontrolle klinisch-chemischer Analysen. Ein Lernprogram in 5 Lektionen für medizinisch-technische Assistentinnen, Studenten, Ärzte. Georg Thieme Verlag, Stuttgart (1972).
134. J. B. HENRY and J. L. GIEGEL (eds.): Quality control in laboratory medicine. Transactions of the first Inter American Symposium, Miami, Florida, Masson Publ. (1976).
135. G. ANIDO, E. J. VAN KAMPEN, S. B. ROSALKI and M. RUBIN (eds.): Quality control in clinical chemistry. Transactions of the VIth International Symposium, Geneva, W. de Gruyter, Berlin (1975).
136. B. LEIJNSE: Kwaliteitscontrole in de klinische chemie. Pericula in mora. Lecture at the occasion of the 25th anniversary of the Netherlands Association for Clinical Chemistry, November 21st, Medical Faculty, Rotterdam (1973).
137. L. G. WHITBY, F. L. MITCHELL and D. W. MOSS: Quality control in routine clinical chemistry, in: Advances in clinical chemistry, Vol. X, O. Bodansky, C. P. Stewart eds. Acad. Press, New York (1967).
138. T. P. WHITEHEAD: Advances in quality control. In: Advances in clinical chemistry, Vol. XIX, O. Bodansky, A. L. Latner, eds. Acad. Press, New York (1977).
139. Dr. L. SIDEMAN: Director Clinical Chemistry Division, Pennsylvania Department of Health, Bureau of Laboratories, kindly made this information available.

140. D. J. SENCER: a. CDC Coordinating Malnutrition Survey. US Medicine *1972*, Jan. 15, 12 - 14.
b. CDC's expanded role described. US Medicine *1974*, Jan. 15, 19 - 24.
141. A. HAINLINE, C. L. WINN, G. R. COOPER, D. T. MILLER and D. D. BAYSE: The CDC Cooperative Cholesterol and Tryglyceride Standardization Programme - Twenty years experience. Clin. Chem. *24*, 1020 (1978).
142. D. GRAFNETTER: World Health Organization (WHO) coordinated quality control in the lipid laboratory. Giorn. Arterioscl. *2*, 113 - 128 (1977).
143. G. J. M. BOERMA and B. LEIJNSE: Draft proposal to set up a standardization programme for the analysis of serum cholesterol (eventually including other lipids). Submitted upon request of Dr. K. F. LAUER; Community Bureau of Reference, Directorate-General for Research, Science and Education, Commission of the European Community, Rue de la Loi 200, Brussels (1976).
144. R. E. WENK: Cholesterol in a regional programme - reduction of interlaboratory variability by correcting systematic bias. Am. J. Clin. Pathol. *55*, 311 - 317 (1971).
145. N. S. LAWSON and G. T. HAVEN: The role of regional quality control programmes in the practice of laboratory medicine in the United States. Am. J. Clin. Path. *66*, 268 - 275 (1976).
146. T. P. WHITEHEAD, D. M. BROWNING and A. GREGORY: The role of external quality control schemes in improving the quality of laboratory results in: Quality control in clinical chemistry. Transactions of the VIth international symposium G. Anido, E. J. van Kampen, S. B. Rosalki, M. Rubin, eds. W. de Gruyter, New York (1975).
147. R. K. GILBERT: Survey data 1975; Chemistry. Publication of the Survey Programme of the CAP, Skokie, III (1977).
148. The Coronary Drug Project Research Group: Clofibrate and Niac in coronary heart disease. J. Am. Med. Assoc. *231*, 360 - 381 (1975).
149. R. R. FRERICHS, S. R. SRINIVASAN, L. S. WEBBER and G. S. BERENSON: Serum cholesterol and triglyceride levels in 3446 children from a biracial community. The Bogalusa Heart Study. Circulation *54*, 302 - 309 (1976).
150. D. KROMHOUT, F. VAN DER HAAR and J. G. A. J. HOUTVAST: Coronary heart disease risk factors in Dutch schoolchildren. Results of a pilot study. Prev. Med. *6*, 510 - 513 (1977).
151. P. F. COHN, S. I. GABBAY and W. B. WEGLICKI: Serum lipid levels in angiographically defined coronary artery disease. Ann. Int. Med. *84*, 241 - 245 (1976).
152. A. KEYS, H. L. TAYLOR, H. BLACKBURN, J. BROZEK, J. T. ANDERSON and E. SIMONSON: Coronary heart disease among Minnesota business and professional men followed fifteen years. Circulation *28*, 381 - 395 (1963).

153. J. F. DE WIJN: Serum-cholesterol en lipidenspiegels bij mannen in Nederland. Overzicht van bevindingen bij screening van groepen in de jaren rond 1970. Hart Bull. *3*, 82 - 88 (1972).
154. J. F. MAY, J. BUREMA, W. NUMAN and D. SIESWERDA: Screenings-onderzoek naar cardiovasculaire risicofactoren met berekening van risico-scores. Ned. T. Geneesk. *122*, 780 - 784 (1978).
155. R. E. BOWYER, D. H. CURNOW and N. S. STENHOUSE: The second Busselton adult population survey (1969): Serum cholesterol. Pathology *6*, 147 - 152 (1974).
156. G. J. M. BOERMA, A. TOLEIKIS, A. MATHER, C. T. BARTELS, L. MARGEVICIENE, I. S. GLASUNOV and B. LEIJNSE: Laboratory standardization in an international study: The Kaunas-Rotterdam Intervention Study (KRIS). Clin. Chem. *24*, 1126 - 1130 (1978).
157. F. STURMANS, L. BUREMA, H. A. VALKENBURG, P. DAS, A. L. DUBBELDAM-MARREE, P. G. H. MULDER and G. J. M. BOERMA: De Kaunas-Rotterdam Intervention Study (KRIS) VI T. Soc. Geneesk. *55*, *suppl. 1* 15 - 24 (1977) (contains refs. of articles I - V).
158. K. STYBLO, G. J. M. BOERMA and A. C. ARNTZENIUS: CB Heart Project in The Netherlands: Cholesterol. Hart Bull. *7*, 47 - 54 (1976).
159. G. J. M. BOERMA and K. STYBLO: CB Heart Project in the Netherlands. Serum cholesterol analysis. Intra- and interlaboratory quality control. Hart Bull. *8*, 66 - 67 (1977).
160. K. STYBLO, G. J. M. BOERMA and B. LEIJNSE: Serum cholesterol analysis in the CB Heart Project. Intra- and interlaboratory continuous quality control. Clin. Chim. Acta *83*, 211 - 222 (1978).
161. J. H. REISEL: Klinisch atherosclerose. Een studie over primaire interventie. Bohn, Scheltema en Holkema, Utrecht 1976(cites many Dutch studies).
162. J. F. MAY: Epidemiological cardiology: Ischaemic heart disease study in Vlagtwedde 1970. Doctoral thesis, Groningen (1974).
163. S. C. UPPAL: Coronary heart disease. Risk pattern in Dutch youth. A pilot study in Westland schoolchildren. Doctoral thesis, Leiden (1974).
164. J. BAART: De oudere havenwerker - Aspecten van zijn cardio-respiratoire conditie. Doctoral thesis, Rotterdam (1973).
165. COPIH Symposium: 'Wie krijgt een hartinfarct?'. Hart Bull. *4*, 57 - 76; 89 - 110 (1973).
166. H. A. VALKENBURG et al.: Epidemiological preventive examination Zoetermeer (EPOZ). Fourth Progress Report, Department of Epidemiology Erasmus University Rotterdam (1977 - 1978).
167. J. F. DE WIJN and N. A. PIKAAR: Zesde oriënterend onderzoek omtrent de voeding en voedingstoestand van achtjarige schoolkinderen in Nederland (1973/1974). Voeding *37*, 127 - 139 (1976).

168. C. PRIES, R. DEEN and W. BROUWER: Evaluatie van de kwaliteit van cholesterol-, totaallipiden- en triglyceridenbepalingen in menselijk bloedserum (COPIH-onderzoek III). Rapport CL 76:100, 8-11-1976. Centraal Laboratorium TNO, Delft.

By standardization is meant the process resulting in meaningful measurement. This process, to be described in some detail, is complex and highly interactive, and it calls for the highest degree of cooperation, enthusiasm, and hard work on the part of scientists, technicians, and manufacturers if success is to be achieved. Furthermore, the scope of the work is so broad that only if international resources are applied to the problem and full cooperation between various national groups and societies is assured will a solution come about in a reasonable time.

J. Paul Cali (1973) (1).

CHAPTER 3

PLASMA CHOLESTEROL AND GLUCOSE ANALYSIS IN THE KAUNAS ROTTERDAM INTERVENTION STUDY (KRIS); INTERNATIONAL INTERLABORATORY STANDARDIZATION (2)

3.1. Introduction

The 'KRIS' was designed to investigate the feasibility of intervention - such as health education or the administration of drugs - in order to prevent the occurrence of atherosclerotic complications in high risk individuals. The amount of risk for myocardial infarction or cerebral stroke was estimated by measurement of three risk factors, namely, elevated blood pressure, hypercholesterolemia and impaired glucose tolerance. The main objectives of the KRIS were: to examine recruitment and complicance of subjects, to study the interaction between the medical community and the research team and to compare the effectiveness of several treatment approaches (3).

This programme was carried out in two cities with contrasting health delivery systems: Rotterdam, The Netherlands and Kaunas, Lithuania, U.S.S.R. The target population consisted of males, aged 45 - 59, and a baseline survey among 4000 men was performed. Finally screening examinations were started in Kaunas in March 1972 with 3561 men and in Rotterdam in September 1972 with 3882 men. A borderline group with mildly elevated risk factors was recruited after two screening

examinations for a double-blind drug treatment trial. A pathological group with highly elevated risk factors was partly referred to their general practitioners or specialists and partly treated by the study team. A 10% sample of the normal group was re-examined for control purposes at the end of the trial, e.g. to detect the influence of screening on risk factor levels. This chapter will deal with standardization and quality control of cholesterol and glucose analysis. The entire laboratory protocol is reproduced in Appendix I.

3.2. Laboratory implementation and standardization

The KRIS programme was developed by the Division of Health Statistical Methodology, WHO, Geneva (Dr. I. S. Glasunov: Project Leader), the Rotterdam Municipal Health Department (Prof. Dr. L. Burema: Director) and the Kaunas Medical Institute (Prof. Dr. Z. I. Januskevichius: Director). As early as October 1970 a preliminary meeting was held in Kaunas to discuss all initial preparations for the study. The methodologies used in Rotterdam and Kaunas for cholesterol and glucose determination were different. In Rotterdam modifications of Turners method (4) and Hofmans ferric cyanide reductions respectively (5) were in use, whereas in Kaunas a version of the Zlatkis reaction (6) and the Nelson-Somogyi method (7) were applied.

It was strongly suggested in the Biochemical group meeting by the chairman (Prof. Dr. B. Leijnse) to use identical examination procedures and SI-units. For cholesterol the Liebermann-Burchard reaction was suggested while for glucose an enzymatic oxidase procedure was discussed. After consulting the special adviser from the CDC (Dr. A. Mather, Associate director; Clinical Chemistry Division) the final proposals were to use a modified Huang method for cholesterol as used in the Netherlands as the normalized procedure NEN 2415 and an o-toluidine method for glucose (8, 9).

Although from an analytical point of view one central laboratory would have been desirable, this was virtually impossible with the large distance between the cities. Further it seemed advantageous to gain experience with standardization procedures since a much larger study in at least ten cities might follow this relatively small feasibility study. Thus it was decided that all chemical analyses would be performed locally in laboratories especially implemented for this task. This operation was performed by the Central Clinical Chemistry Laboratory at the University Hospital-Dijkzigt and in Kaunas by the Department of Cardiology. With distance and transit times involved in the Kaunas-Rotterdam studies maintenance problems and materials supply presented difficulties that made automation undesirable. Complex manual procedures are at the same time difficult to maintain in control and close

interlaboratory comparability requires frequent communication. On this basis the two manual analytical methods were selected, both being performed in plasma samples without deproteinization or extraction. It was considered essential to acceptable standardization that all supplies, equipment and reagents had identical specifications and these were then, through WHO, obtained from common sources.

After selection of the chemical methods, these were studied in both laboratories. Both methods proved to be quite sensitive to slight variations in analytical technique. The support of WHO in ensuring that the two biochemists (Dr. A. Toleikis and Dr. C. T. Bartels) spent a week in each others laboratory and the intensive efforts of the adviser from the American CDC laid the groundwork for ironing out many problems.

A great amount of attention was paid to the elaboration of a detailed protocol. Our experience has shown that procedures without detailed description become poorly controllable and several details may be performed in different ways in different laboratories. The very aim pursued by cooperative epidemiological investigations - to obtain comparable results - is thus defeated. The development of the protocol was a continuous process of refining and detailing of methodology aiming at the elimination of the possibility of different interpretation and performance. The protocol describes why and how units, common sources of materials and methods were selected. It states exactly how specimens for analysis must be obtained, handled and stored. All details of the analytical method are presented, together with procedures for calibration, computation and quality control (Appendix I).

Materials:

Equipment

Colorimeter UC 200-S with flow-through cuvet from Vitatron (Holland), Mettler H 10 TW digital analytical balance (Switzerland), all-glass reagent dispensers from Salm and Kipp (Holland), Vacutainers type 3200PS (containing 25 mg of sodium fluoride and 20 mg of potassium oxalate, size 10 ml) for blood collection from Becton-Dickinson (France). All other materials like MacLean pipettes and test tubes were identical in both laboratories by supplying them from common sources.

Reagents

Propanol-2, sulfuric acid, acetic acid, acetic anhydride, benzoic acid, anhydrous sodium sulfate, thiourea, D-glucose, quality 'pro analyse' from Merck, Darmstadt, West-Germany; o-toluidine and cholesterol from British Drug Houses Ltd.; a

common lot of control serum (Metrix Normal) was obtained from VWR Scientific International (USA) in December 1972. Before that time Monitrol and Labtrol (Dade) were used.

3.3. Standardization and quality control procedures

A manual 'direct' Liebermann-Burchard procedure, well-known in the Netherlands as method NEN 2415 was used. This modification from Huang (9) was issued by the Netherlands Normalization Institute (NNI) on recommendation of the Committee on Methods of the Netherlands Association for Clinical Chemistry (NVKC) and the National Institute of Public Health (RIV). Several problems connected with this method were discussed in chapter 2.

When the staff of both laboratories during 1971 had familiarized themselves with the analysis they were able to exchange plasma or serum samples or standard solutions and obtain results that were quite close together. The programme that was finally adopted to ensure good quality and adequate comparability, consisted of the following essential procedures.

A. Collection of specimens (Protocol page 11, 12)

Collection of plasma samples according to the protocol.

B. Intralaboratory quality control (Protocol page 13 - 28)

Every step of the analytical procedure carried out *strictly* according to the protocol.

Calibration with common standard material, including a check on linearity and drift.

Measurement in duplicate of all samples.

Measurement in all runs of two serum controls (normal and elevated levels).

Application of predetermined quality criteria.

Measurement in each run of at least one patient specimen from a previous run - preferably 'blind'.

C. Interlaboratory quality control (Protocol page 3, 7, 8)

Periodical analysis of reference serum from CDC (external surveillance programme).

Analysis of quantities of patient specimens stored continuously in both Rotterdam and Kaunas and shipped frozen by air for comparability evaluation.

3.3.1. ad A. Specimens of venous blood were collected by venipuncture employing 10 ml Vacutainers with 25 mg of sodium fluoride (preservative for glucose) and 20 mg of potassium oxalate (anticoagulant). The subjects had fasted for two hours (initial screening) or overnight (later examinations). The tubes were mixed by in-

version ten times and centrifuged after an hour at 1000 g for 10 minutes.
Comments.

a. An investigation into the effectiveness of the preservative led us to conclude that glycolysis continues at a slow pace (table 3.1.) and that centrifugation should occur as soon as possible - we compromised on 1 hour after blood sampling. The cholesterol level practically was not influenced by standing of the blood. Kaunas findings were in fair agreement.

Table 3.1. Time dependence of plasma cholesterol and glucose levels.

Sample nr.	GLUCOSE (mg%)				CHOLESTEROL (mg%)			
	T_1 [1]	T_2	T_3	T_4	T_1	T_2	T_3	T_4 [1]
7925	105	100	100	100	190	190	191	191
7922	91	85	83	81	179	179	179	180
8626	113	110	108	108	187	188	184	193
7279	170	170	169	169	149	146	146	146
7516	194	194	190	186	204	201	205	203
7689	60	57	56	55	266	266	271	270
6935	72	71	71	68	211	212	209	212
6564	64	62	62	60	210	207	211	214
8470	146	144	142	141	211	211	212	214
9301	103	96	95	94	155	155	157	160
($\bar{x}$)	111.1	108.9	107.6	106.2	196.2	195.5	196.5	198.3
	0%	- 2.6%	- 3.8%	- 5.0%	0%	- 0.4%	+0.2%	+1.1%

[1] T_1 = 5 min after sampling; T_2 = 30 min after sampling; T_3 = 60 min after sampling; T_4 = 120 min after sampling.

b. Usually the plasma samples were frozen overnight until transportation to the laboratory. Upon thawing clot formation was observed frequently and thus for short term storage a refrigerator was preferred.

c. Oxalate and fluoride may cause hemolysis - plasma samples occasionally appeared slightly reddish.

d. Both laboratories at the beginning of the study had been supplied with oxalate-fluoride Vacutainers, although originally EDTA was chosen instead of oxalate. That specific type of tube was not available in sufficient quantity at the appropriate time. Unfortunately afterwards logistical difficulties in Kaunas precluded

the replacement of oxalate tubes by EDTA tubes. The plasma used in this study therefore had a lower cholesterol level as compared to corresponding serum samples due to an osmotic water shift from the blood cells towards the oxalate containing plasma. The systematic error caused by this water shift is discussed below under 'Bias'. Practically all anticoagulants cause water shifts to some extent, but unfortunately oxalate holds the record (10).

3.3.2. ad B. Intralaboratory Quality Control

The internal quality control programme included four chief procedures:

1. Measurement in duplicate of all patient and control samples, to rule out large errors.
2. Measurement of pure standard solutions at several concentration levels at the beginning and end of each run for constant checking of linearity and drift.
3. Measurement in each run of at least two serum controls (normal and elevated levels).
4. Measurement in each run of one or more patient samples taken from a previous series (sample resubmission programme). This was intended to provide both centres with a 'blind' means of measuring between-run precision. In daily practice, however, the selection of resubmitted samples took place within the laboratory.

In the development of the laboratory manual it was decided to include fixed limits for quality acceptability. We considered adherence to such imposed limits to be a vital prerequisite to interlaboratory comparability. The rejection limits used were:

1. Any duplicate pair of optical density readings deviating by more than 3% of the higher value called for reanalysis of the specimen, or rejection of a standard reading (an exception is noted below).
2. Standard solutions in duplicate at three levels (for glucose) and four levels (for cholesterol - later reduced for practical reasons to two levels) were included in each run of about 20 specimens. The slope of the calibration curve calculated for each point was required to fall within ± 3% of the mean slope. One of the duplicate values at one concentration level of three or four levels could be discarded for deviations of slope or duplication greater than 3%. When only two standard levels were used, one single value deviating by more than 3% but less than 6% could be deleted, provided that the other duplicate and all other values fell within the limit. Otherwise the run was rejected. Note that the limits on duplicate deviation between the initial and final readings of the standard provided control for excessive instrument or operational drift as well as for linearity. For an extensive description see Appendix I.

3. The two serum controls were analysed in duplicate in 20 runs and the average value then remained the target control value as long as a lot was used; the mean of duplicates for both controls in each run had to lie within ± 6% of this target level, based upon a maximum allowable coefficient of variation preset at 3%. Initially a number of runs had to be discarded, but with growing experience the precision was ordinarily kept well within the limits of the protocol.
4. The duplication of plasma samples from previous runs was meant to be an additional check upon the repeatability but was not formalized and not reported to Geneva on special sheets as originally scheduled. However, it provided the laboratory directors with an extra tool in evaluating analytical performance.

Comments

a. In the SI-system the kg is the unit of mass, and so g/l may be used. However, the mmol/l is the unit to express quantities (e.g. in a solution) and it is gradually becoming more generally employed, or even legally required. In 1971 the practical situation made its usage almost impossible. Neither the USSR nor the USA had chosen to adopt this unit. It was agreed that results expressed in both units should be retrievable from the data base. The use of the mmol/l must be strongly recommended of course in future studies as a prerequisite for real standardization.
b. Calibration with pure standards often leads to bias in clinical chemistry. Usage of serum calibrators as a means of interlaboratory standardization may now be recommended (Chapter 6), but this was much less certain at that time.
c. The predetermined quality criteria contributed a great deal to the good performance of the two laboratories. It forced us to discard sample or run results that might otherwise have been accepted. When experience grew it proved to become easier to remain within the imposed limits. In several situations the decision to discard a series of measurements is often hard to make and borderline control results may be accepted too soon. The maximum difference between duplicates, however, (3%) now seems to be a too severe limit for low concentrations, because it required reanalysis when the glucose values were e.g. 80 and 83 mg % (4.44 and 4.61 mmol/l). The maximum acceptable coefficient of variation (3%) has always been a good compromise with the state of the art.
d. Blind replication of human plasma samples is a superior way to establish the true precision of the laboratory. This was also demonstrated in the C.B. Heart Project (Chapter 4). The procedure ought to be included in all population studies and perhaps in all laboratory procedures.
e. Quality control also included checks on the calibration of the volumes of rea-

gent dispensers, temperature of thermostatted water baths, exact timing of each individual reaction, comparison of each freshly prepared standard solution against the older one etc. all in accordance with the protocol.

3.3.3. ad C. Interlaboratory Quality Control

In order to measure the bias and monitor drift in each laboratory, an external surveillance system was required, although logistic problems precluded completely blind evaluation. Periodically a set of 12 unknown specimens from human serum pools was sent out from the Center for Disease Control, Atlanta, for the evaluation of the current proficiency of the Kaunas and Rotterdam Laboratories. Two sets of six samples each were included in each shipment: the six vials in one set were analyzed in duplicate, the second set of six vials was similarly analyzed in a separate analytical run. All results were recorded on a standard form supplied with each shipment and promptly returned to Atlanta. Evaluations of the survey were sent to WHO and a summary of performance (blinded as to individual samples and reported only as % bias) was sent to each laboratory.

The early survey results demonstrated interlaboratory bias between Kaunas and Rotterdam and measures were taken to make standardization more rigid.

A second comparison, made possible through the exchange of large numbers of frozen specimens, nicely coincided with the survey results (section 3.4.1.2.).

Comments:

a. The external surveillance programme might have been more frequent through 1974. As a whole it furnished the analyses with an indispensable check on accuracy and stability.
b. The first large shipment of some 450 frozen plasma samples from Kaunas to Rotterdam created analytical problems when several showed clot formations upon thawing. More than a hundred plasmas for that reason were discarded. Accurate pipetting was hardly possible. When more attention was paid to proper mixing of the oxalate fluoride and blood this kind of problem was much reduced in later exchanges.

3.4. Results and discussion

3.4.1. Accuracy

Common standards and specified requirements for evaluating accuracy for the two chemical analytes do not yet exist. They are being developed by concerted national and international efforts and are under evaluation (11, 12). Comparability

between any two or more quantitative data bases can be judged only on a case-by-case comparison of all the procedures that may contribute to systematic errors in measurements of biological characteristics: patient preparation, specimen collection and preservation, analytical procedures, etc. Thus, unless a carefully documented common basis exists for these factors and reliable estimates of bias components are available, between-study comparisons of quantitative data will necessarily be of limited value.

The design of the Laboratory Protocol attempted to provide mechanisms for the control and measurement of quality parameters. Not all of these factors (e.g. blood collection) were under laboratory control.

3.4.1.1. Analytical bias

The reference methods selected as bases for analytical accuracy were the method of Abell, Levy, Brodie and Kendall (13) for cholesterol, as performed in the Lipid Reference Laboratory and for glucose an o-toluidine reaction on a protein free filtrate, cross checked with a glucose oxidase manual method and an automated oxidase procedure employing a Beckmann glucose analyser with an electrode sensing the oxygen consumption (14). Recent work on definitive and reference methods for both analytes indicates that these methods do indeed have no substantial bias (11, 15). The bias of the analytical procedure as performed in each laboratory was estimated by a programme of monitoring results periodically throughout the study by the external quality surveillance programme, using common tests samples of human serum, analyzed in Atlanta. A panel of 16 sample pools of related composition, being obtained from 5 analyzed human serum pools in various mixtures, provided sufficient variety to maintain blindness of test sample identity in the laboratories; blindness was maintained in the reports of the survey results regarding deviations from target values. Within-run and between-run duplicates were included. In addition, quality assurance materials (Metrix) were provided to both laboratories as a common reference point for estimating and detecting out-of-control conditions. Owing to selection and procurement problems these common materials became available only in early 1973. Since they represent modified and lyophilized serum, their value in measuring bias of results with subject specimens is limited, but they did assist in the evaluation of between-laboratory bias as well as in short term drift and between-run variability. Unfortunately the 'normal' level consisted of two different lots - only the 'elevated' serum was identical in both laboratories.

Results:

The samples of survey 1 - a glucose survey only - arrived in Rotterdam one month later than in Kaunas, which may account for low values in some samples. The average ratios (R/CDC and K/CDC) were 0.917 and 0.953. At the time, October 1971, linearity, temperature of reagents and inter observer variability were still under investigation. Later results showed great improvement.

During 1972 when survey 3 and 4 revealed considerable discrepancies in cholesterol values - a mean difference between R and K of 5% was seen - the possibility of correction factors or serum calibrators was considered. None of these were applied as in Kaunas after the summer holidays a downward shift occurred and the two sets of results became quite close together.

Use of serum calibrators may indeed offer solutions when comparability problems arise.

Example:

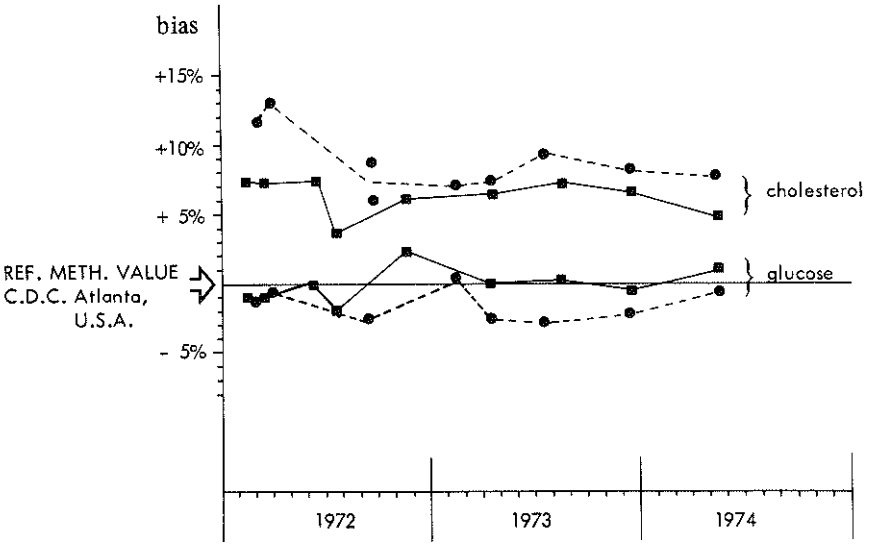
In survey 3 two reference samples had reference method values of 250 mg% (6.46 mmol/l) and of 282 mg% (7.03 mmol/l). These samples were inserted two times among the 12 samples as blind duplicates. If the mean of each duplicate result were to be used in calibration, the following results can be calculated.

	Kaunas bias	Rotterdam bias	Difference K minus R
survey 3	+ 13.8%	+ 8.7%	+ 5.1%
'standard 250'	- 0.9%	+ 1.6%	- 2.5%
'standard 282'	+ 1.3%	- 1.8%	+ 3.1%

This calculation shows how both participating laboratories approach the CDC values a lot better, but between them interlaboratory bias is dependent upon the chosen calibrator. The choice of proper calibrators is not an easy matter (chapter 5), and one measurement is not a sufficiently solid basis for calibration.

Figure 3.1. summarizes the subsequent survey results. It shows calculated average bias of the determined value and the reference method value of all sets of 12 samples. These averages illustrate the bias from the reference point, and reveal occasional periods when the accuracy definitely drifted from the average. For glucose, the overall bias from the reference method values was essentially zero for Rotterdam and - 1.7% for Kaunas. The limited frequency of surveying could not reveal short time trends for analytical and calibration drift. For cholesterol a positive bias was anticipated for the method employed. It must be pointed out that for

Figure 3.1. Laboratory accuracy with human reference serum, KRIS.



■ Rotterdam, mean of 12 samples
● Kaunas, mean of 12 samples

the first half of 1972, the Kaunas cholesterol results were abnormally elevated by a calibration problem with standard solutions. This was revealed by the surveys (3 + 4) when a + 13.6% bias in the Kaunas cholesterol results vs. a + 7.8% bias for Rotterdam occurred. The problem was corrected in autumn 1972 (by the time of survey 5), after which the Kaunas bias averaged +7.9% and that of Rotterdam +6.4%.

3.4.1.2. Interlaboratory comparisons

a. In addition to the surveys, a measure of between-laboratory bias was provided by two exchanges of fairly large numbers of duplicate subject plasma specimens for analysis by the second laboratory (see Table 3.2.). The first such shipment of Kaunas specimens in early 1973 included 133 samples for which chol-

Table 3.2. Comparative results of plasma analyses on the same subject samples in both laboratories.

Date of exchange	number of specimens	Average values (mg%) Kaunas	Rotterdam	Mean difference (%) (K - R)
		Cholesterol		
June 1973	133[1)]	198	184	+ 7.0
	192[2)]	192	188	+ 2.1
October 1974	148[3)]	207	204	+ 1.2
		Glucose		
June 1973	322[1,2)]	139	136	+ 1.8
October 1974	148[3)]	137	138	- 0.7

1) Samples collected and analyzed in Kaunas between March and July 1972 and 2) between September 1972 and June 1973, and reanalyzed in Rotterdam.
3) Samples collected and analyzed in Rotterdam between July and October 1974 and reanalyzed in Kaunas.

esterol values had been determined with the erroneously prepared early 1972 standard solution; the data for this period is shown separately for the cholesterol results, and reveal a deviation from later results of about 5% (this agrees well with the difference shown in figure 3.1. first surveys). After this time the Kaunas results coincided with the 1 - 2% positive difference over Rotterdam results shown in the surveys 5 through 11. Although this type of verification required the collection and storage of separate samples of each specimen and their transportation in the frozen state presented problems, this procedure provided a valuable validation check on the results obtained by the CDC surveys.

b. Comparison of the results for the quality control materials is another and more intensive basis for comparison of interlaboratory bias that could be correlated with time. Two levels of analytes concentration were furnished, the higher of which was effected by a smaller dilution of the 'normal' Metrix material in its reconstitution. To achieve a control concentration approximating the higher levels attained for the 1 hour post-challenge glucose analysis, a 100 mg% glu-

cose standard was used for the reconstitution fluid.

The procedure permitted fairly exact calculation of the concentration ratios for each analyte, which furnished a reliable check on 'recovery' of added analyte. Table 3.3. gives the mean values for the normal and elevated controls averaged over periods of some 20 runs. Only the elevated level regards identical lot numbers of material! The time frame presented for each laboratory extends from March 1973 until December 1974 (Kaunas in this period performed fewer runs). The analyses continued unaltered until early 1976. It is to be noted that with these materials no substantial differences are apparent in the cholesterol values of the two laboratories, but Kaunas is just above Rotterdam on the average. On the basis of the above mentioned exchanges a 2% lower cholesterol and a 2% higher glucose level could be expected, whereas only ½% is found. This confirms the earlier statement (section 3.4.1.1.) that one, or few, lyophilized sera are not of sufficient value to establish bias. This is further discussed in section 6.4. The Rotterdam data have been visualized in figure 3.2. for easy comparison with figure 4.1. (CB Heart Project control sera).

3.4.1.3. Sample Collection as a source of error.

Although great attention was paid to control analytical bias, a much more serious error was introduced into the cholesterol results by the system of blood sampling that was finally evolved.

For many years, blood serum has been the sample of choice for cholesterol measurements, although alterations in plasma lipid concentration during the coagulation process may not be exactly predictable. With the great emphasis placed upon control and documentation of interlaboratory differences by the laboratory planners of the study, it was considered essential that rather large numbers of samples from each population be analyzed in both laboratories. In view of the necessary preservation of glucose both during local processing and in long term storage and shipment, it was decided to use fluorided plasma as the common specimen for both glucose and cholesterol - a decision influenced by the costs and labour of separate collection, handling and shipping of both serum and plasma samples.

Preliminary studies with a vacuum collection tube containing sodium fluoride (2 mg per ml of blood) and EDTA showed that the lowering of cholesterol values in the preserved plasma from the corresponding values obtained in serum (8 - 12%) was of the same magnitude as the positive bias anticipated with the method (This method related bias later proved to be approximately 6.4-7.9%). Such plasma discrepancies are attributable to osmotic shifts of cellular water into the plasma, since

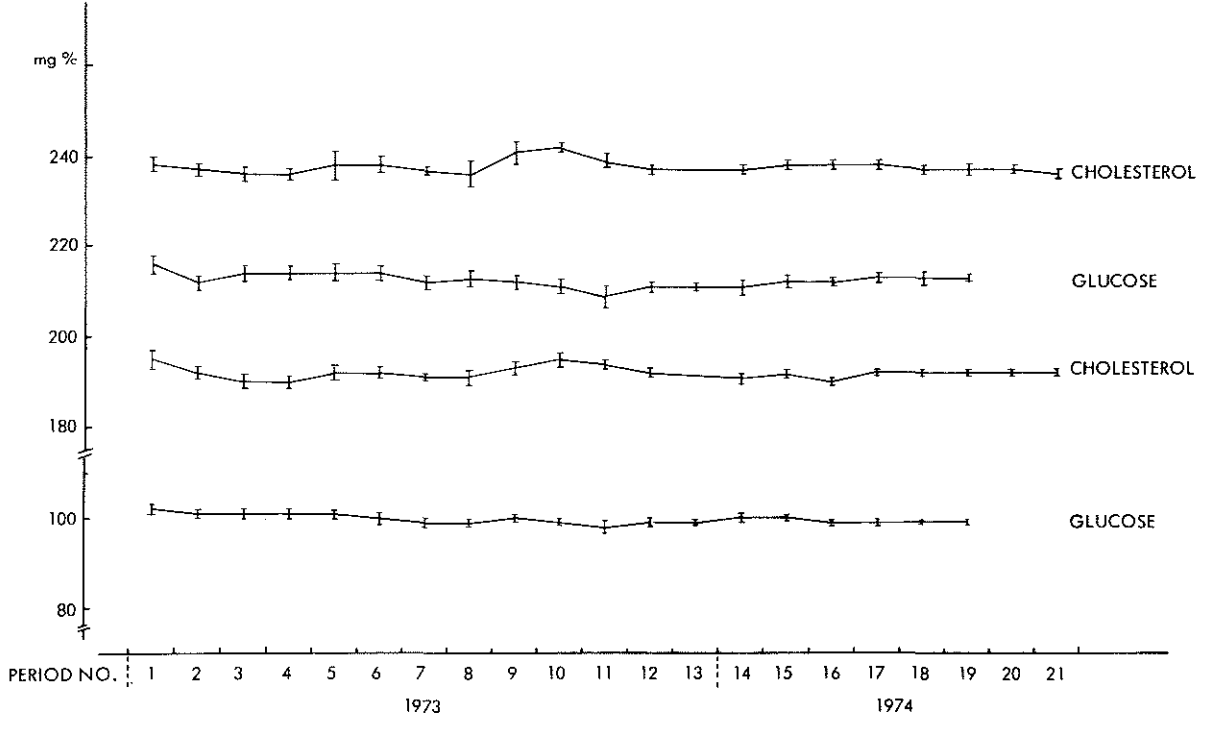
Table 3.3. Interlaboratory bias, common serum control means by twenty-one 20-run periods over the duration of the study.[1)]

Period No.	Glucose mg%				Cholesterol mg%			
	Normal		Elevated		Normal		Elevated	
	K[2)]	R[2)]	K	R	K	R	K	R
1973: 1		102		216	194	195	235	238
2	100	101	212	212		192		237
3		101		214		190		236
4		101		214		190		236
5	102	101	210	214		192		238
6		100		214		192		238
7		99		212	196	191	238	237
8		99		213		191		236
9	101	100	210	212	192	193	235	241
10		99		211	194	195	237	242
11	102	98	210	209	196	194	236	239
12	102	99	211	211	201	192	242	237
13	102	99	208	211				
1974: 14	102	100	211	211	198	191	239	237
15	99	100	210	212	199	192	244	238
16	106	99	217	212		190		238
17	103	99	209	213	198	192	241	238
18	100	99	212	213	195	192	237	237
19	99	100	213	213		192		237
20					195	192	238	237
21					198	192	241	236
$\bar{x}$ (mg%)	101.5	99.8	211.1	212.4	196.3	192.0	238.6	237.6

[1)] Normal control pool was not identical in K and R; Elevated control pool was identical.

[2)] K = Kaunas; R = Rotterdam.

Figure 3.2. Mean values of serum controls in 20-run periods in the KRIS; Rotterdam ($\bar{x} \pm 2$ sem).



oxalate and fluoride per se were found to have no significant effect on the analytical method when they were added to serum (10).

Unfortunately, the selected tubes were not available in Europe except on a special production order with much delayed delivery, and it was hurriedly decided to substitute available oxalate-fluoride tubes. The effects of these concentrations on serum cholesterol were found by subsequent investigation to be greater than expected. A downward adjustment of the cholesterol cut-off level, became necessary. The data for the studies in each laboratory are presented in table 3.4. There is apparently no relation between serum-plasma differences and the cholesterol level, table 3.5.

Owing to the rapid exchange of cellular glucose with plasma, the serum-plasma differences were found to be very small, *on the average* under standardized conditions of sample collection and processing. The epiloque (section 3.5.1.) discusses plasma-serum comparability. The type of specimen and the conditions of collection and processing are extremely important in population and clinical studies and it is essential that steps be taken to standardize these factors. Until they are, strict comparability will be impossible to achieve. For example: separate specimens should be taken for glucose and cholesterol; serum or carefully heparinized plasma for cholesterol, and EDTA and fluoride (1 mg per ml) for glucose, as was initially planned. This second blood sample should be centrifuged as soon as possible. The times of harvesting serum or plasma must be standardized, although this presents a burden on the scheduling of examinees and clinical workload (16, 17, 18). The vacuum tube is suited very well for taking a blood sample without stasis. The subjects should be sitting upright at the occasion.

3.4.2. Precision

In quantitative chemical measurements, precision refers to the random errors of measurement. Measures of actually achieved precision in routine performance are obtained by indirect means that usually underestimate the variability of actual sample results.

3.4.2.1. Repeatability (Within-run precision)

In this study a fixed limit of 3 % of the higher value was placed upon the acceptable deviation between two duplicate values obtained on each sample within each run and samples showing results exceeding this limit were reanalyzed.

Therefore, ordinary estimates of precision cannot be applied to these data. An indirect estimate of the reliability of a single test result can be gained from the rate of rejection of samples outside the 3% limit. These have not been worked up for

Table 3.4. Bias resulting from blood sampling technique.[1)]

Date	n	$(\frac{\text{serum}}{\text{plasma}} - 1)\ 100\% \pm 1\ s$	Analyzed in
December 1972	46	19.2 ± 4.8	Rotterdam
Summer 1974	100	16.4 ± 4.0	Rotterdam
Summer 1974	122	19.9 ± 6.8	Kaunas

1) Serum and plasma obtained from a patient at the same time in two types of Vacutainers, both centrifuged after one hour.

Table 3.5. Ratio cholesterol is not dependent on concentrations.

Rotterdam; 100 subjects

Plasma concentration	numbers	mean ratio serum/plasma
⩽ 150 mg%	4	1.17
151 - 200 mg%	44	1.16
201 - 250 mg%	49	1.17
251 - 300 mg%	2	1.18
> 300 mg%	1	(1.21)

Kaunas; 122 subjects

Plasma concentration	numbers	mean ratio serum/plasma
⩽ 150 mg%	6	1.22
151 - 200 mg%	61	1.19
201 - 250 mg%	45	1.20
251 - 300 mg%	9	1.20
> 300 mg%	1	(1.19)

Table 3.6. Within-run duplicates. Number of sample results rejected and repeated because of duplicate deviation beyond 3% of the higher value (Kaunas results).

year	GLUCOSE			CHOLESTEROL		
	Total n	Rejected n	%	Total n	Rejected n	%
1972	1139	55	4.8	1090	15	1.4
1973	1779	39	2.2	1922	23	1.2
1974	1168	33	2.8	1304	28	2.1
1975	536	12	2.2	815	49	6.0

Rotterdam, but table 3.6. gives the numbers and percentages of all samples rejected as not meeting this criterion in Kaunas. The repeatability in both laboratories was consistently much better than reproducibility, and the low rejection rate for duplicates indicates quite good within-run performance.

3.4.2.2. Reproducibility (Between-run precision)

The use of common lot materials in each analytical run by both laboratories throughout the major part of the study provided a powerful monitor of between-run precision. During the early part of the study until common control materials became available, runs were declared out of control when either control value exceeded 2 s of the predetermined variability; after this time, a fixed limit of 6 % was placed on the allowable deviation of either control. Above this limit the run was rejected. The number of rejected runs is known for Kaunas. In the period 1972 - 1974 of all glucose and cholesterol runs 15% and 4% were discarded. The most reliable estimates of achieved precision are obtained when (a) test samples truly represent the specimen population analyzed and (b) the testing is completely blinded from the technician. A system of this type was provided for in the Laboratory Protocol by a 'Blind Sample Resubmission' programme, in which a sampling of specimens in each run is reserved for reanalysis in the next run. Although the Protocol specified 3 such samples per run, this represented a large portion of the total work-

load, and the number was later reduced to at least one resubmitted sample per run. This procedure was designed to provide the laboratory director with a sensitive estimate of overall reproducibility of performance; the 'blinding' was less rigorous in Rotterdan than in Kaunas until 1976. The complete data for this programme, however, were not collected centrally, as planned.

Summarizations are available in table 3.7. Such estimates for blind tests show greater variability than those obtained from control sample data as is seen when lines 3) and 4) are being compared. The Kaunas data includes results obtained in rejected runs and the difference between Rotterdam and Kaunas coefficients of variation reflects the improvement of precision when our quality control limits are applied.

3.4.2.3. Precision of the survey results

Results of those samples that were inserted as blind duplicates in survey 3 through 11 allow us to calculate standard deviations for repeatability and reproducibility evaluation. Duplicates that were present within the same survey have been analyzed on two consecutive days - occasionally a few days in between may have occurred. This is called day-to-day precision in the table. The s for 2n degrees of freedom were found by applying the 'Snedecor equation' that was also used in table 3.7. (table 3.8.).

To obtain an estimate of long term precision all results of 11 pools, samples whereof appeared in 4 surveys or more, have been selected. For each of these 11 pools reproducibility was estimated by calculating s. A weighed average s is presented in table 3.8.B. Survey 3 and 4 were excluded in the cholesterol calculations.

Especially the long term s values are a clear demonstration of the excellent reproducibility of the two sets of biochemical data obtained in the KRIS study. The overall mean bias between the two participants (1.8% for glucose and 1.4% for cholesterol) confirm that with simple manual methods that are controlled elaborately and performed every day in the same standardized manner a high level of comparability may be achieved.

3.5. Epiloque

3.5.1. Plasma vs. serum cholesterol values

Although the KRIS was officially brought to an end on December 31st, 1974 in both centers the work was continued. All persons 'at risk' were given an opportunity to remain under surveillance and examinations took place half yearly.

In Rotterdam it was then decided to change from plasma to serum measurements in conformity with most other studies, including the CB Heart Project. We

Table 3.7. Between-run reproducibility calculated from the duplicate blind sample resubmission programme for a few time periods.

Analyte	Period	n	$\bar{x}_1$ [1] mg%	$\bar{x}_2$ [1] mg%	s mg%	CV %	Laboratory
CHOLESTEROL							
	[2] I	493	196.8	196.5	12.5	6.3	Kaunas
	II	223	212.2	211.9	11.4	5.4	
	[3] run 1 - 50	47	209.2	209.5	5.3	2.6	Rotterdam
	run 350 - 400	48	233.2	234.2	2.0	0.9	
	[4] August 1975	20	234.6	234.4	8.3	3.5	Rotterdam
	February 1976	30	235.5	232.0	9.9	4.3	
GLUCOSE							
	[2] I	470	141.8	138.9	13.1	9.3	Kaunas
	II	215	99.4	99.0	6.9	6.9	
	[3] run 1 - 50	45	165.0	164.0	4.0	2.4	Rotterdam
	run 350 - 400	51	146.0	146.3	2.1	1.5	

[1] $\bar{x}_1$ and $\bar{x}_2$ are the means of all first and all second analyses. Resubmission in the laboratory was blind [2] or blind as to concentration only [3] or sent blind from outside the laboratory [4].

$s = \sqrt{\frac{\Sigma(\Delta x_i)^2}{2n}}$; (Δx_i) being the difference between $(x_i)_1$ and $(x_i)_2$.

Table 3.8. Precision and bias estimates from external surveillance results.

	A. Day-to-day [1] precision		B. Overall [3] long-term precision		C. Overall [5] bias	
	n	s calculated [2] from duplicate differences (mg%)	n	mean s [4] (mg%)	n	mean bias %
GLUCOSE						
Kaunas	19	4.23	60	4.30	88	- 1.7
Rotterdam	22	2.98	67	3.01	100	+ 0.1
CHOLESTEROL						
Kaunas	25	3.31	50	4.26 [6]	80	+ 7.9 [6]
Rotterdam	25	3.49	69	4.24	107	+ 6.4

[1]) Duplicate values (n pairs) obtained **within** surveys in two consecutive days on identical serum, blinded as to concentration.

[2]) s for 2n degrees of freedom, as in table 3.7.

[3]) Results obtained in **different** surveys in 11 samples that were inserted at least 4 times.

[4]) Average of s values of 11 samples, analysed in 4 - 8 surveys, weighted for n. (n = 4, 5, 6, 7 or 8).

[5]) Average bias of all reported values in 9 surveys.

[6]) Survey 3 and 4 excluded; bias at the time + 13.6% due to calibration problem. Compare figure 3.1.

then evaluated the problem of comparing new serum values with old plasma levels. It was decided also to perform all measurements both ways until serum baseline values would have been obtained for each participant.

This evaluation is based on the large number of duplicate analyses already done on the two materials (table 3.9.) and on our estimate of biological variations in our population.

Table 3.9. (extended table 3.4.).

Laboratory	Period	number of duplicate analyses	ratio $\frac{\text{serum}}{\text{plasma}}$	s of ratios
Rotterdam	Dec. '73	46	1.182	0.048
	Jan. '74	10	1.176	0.026
	Apr. - Sept. '74	100	1.164	0.040
	Aug. - Sept. '75	106	1.169	0.042
Kaunas	Febr. '74	20	1.192	0.077
	Summer '74	122	1.199	0.068

In the equation: $y = ax + \epsilon$,

with y = serum value

x = plasma value in the corresponding sample

ϵ = difference

we calculated the best fitting value for 'a' in such a way that the sum of squared residuals 'ϵ' would be minimal. Other equations were tried out also, but none gave better results than the simple proportional one given here. The value for 'a' is given with the standard deviation and it differs significantly between Rotterdam and Kaunas.

	a	s e m	s_d (mg%)
Rotterdam	1.170	0.004	7.3
Kaunas	1.203	0.007	11.8

When the factor 'a' is applied, small differences remain between measured serum values and calculated ones. The standard deviation of these differences is s_d. To judge whether such differences are of consequence in evaluating plasma choles-

terol changes in our subjects they must be compared with biological variations. Two estimates of that variability can be made with KRIS data:

1. After initial screening (IS) there has been a rescreening (RS) in Rotterdam and Kaunas of 1343 and 821 men respectively. Initially they had one or more borderline risk indicators. The risk pattern at RS was considerably more moderate, see table 3.10. (19).

Table 3.10. Risk pattern at IS and RS in some KRIS subjects.

	n of men with borderline risk (IS)	cholesterol in plasma ⩾ 216 mg% ⩾ 5.58 mmol/l (IS)	⩾ 216 mg% ⩾ 5.58 mmol/l (RS)
Rotterdam	1343	666	500
Kaunas	821	366	288

From these measurements (500 duplicate examinations in Rotterdam and 288 in Kaunas) an estimate of averaged individual variability was made. The calculated standard deviations are 18.9 mg% (0.49 mmol/l) and 22.9 mg% (0.59 mmol/l) for Rotterdam and Kaunas respectively.

2. During 1976, when subjects with various risk patterns were examined half-yearly, we took serum cholesterol values from 23 of them, measured at three consecutive check-ups. The men belonged to the original group of 500 (table 3.10.) and had received dietary and drug treatment since 1973. With all intra-individual changes observed (46 'intervals') a standard deviation of 19.7 mg% (0.51 mmol/l) was calculated. From these two estimates it appears that we must reckon with an average intra-individual biological s of some 19 mg% (0.49 mmol/l). The additional factor of variability - s_d, caused by plasma-serum correction (7.3 mg%) - does not enlarge the variability very much. The average overall s becomes: $(19^2 + 7.3^2)^{1/2} = 20.4$ mg% (0.53 mmol/l). This means that no substantial additional component of variation is introduced when plasma cholesterol values are increased by 17% to compare them with serum levels.

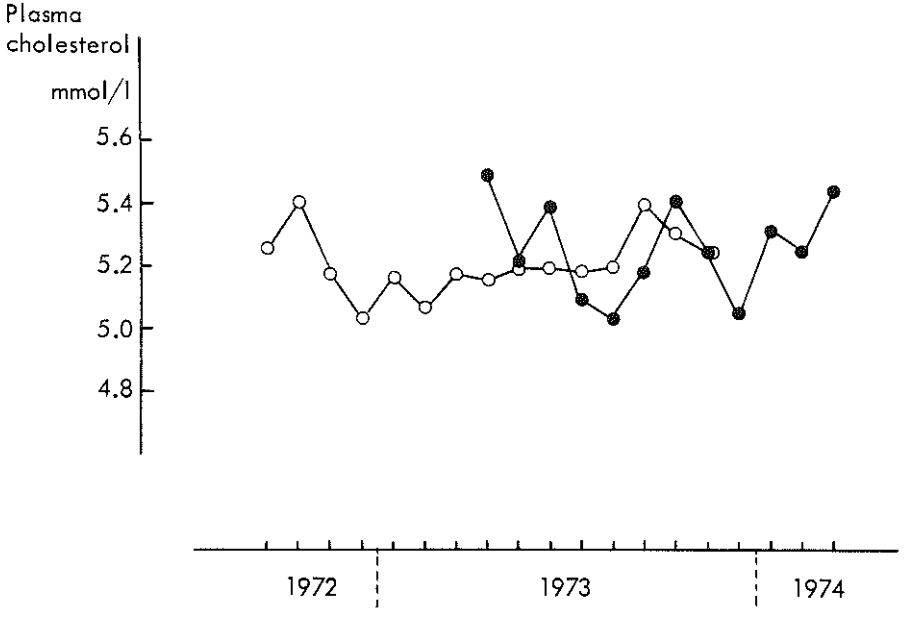
3.5.2. Seasonal variations (figure 3.3.)

The mean cholesterol level was calculated each month in all baseline examinations. The number of men screened each time was large enough to assume that the groups are more or less comparable samples. We observed two relatively high means

in October 1972 and September 1973. For comparison a year from the CB Heart Project is plotted as well. The oldest age group is taken from that study and serum values are turned into plasma values by the correction factor observed in 1973 (1.17).

The selection of men for the KRIS from two particular suburbs of Rotterdam and of certain age groups from other cities in the CB Project make comparisons and general conclusions difficult. Further discussion will follow in section 6.5.

Figure 3.3. Monthly mean plasma cholesterol values at baseline screening in the KRIS, men 45 - 59 yrs (o) compared with CB Heart Project serum values - 'corrected' to plasma level - in men 40 - 49 yrs (●).



REFERENCES

Chapter 3

1. J. P. CALI: Problems of standardization in clinical chemistry. Bull. Wld. Hlth. Org. *43*, 721 - 726 (1973).
2. This chapter is based on the paper:

G. J. M. BOERMA, A. TOLEIKIS, A. MATHER, C. T. BARTELS, L. MARGEVICIENE, I. S. GLASUNOV and B. LEIJNSE: Laboratory standardization in an international study: The Kaunas Rotterdam Intervention Study (KRIS). Clin. Chem. *24*, 1126 - 1130 (1978).

3. The KRIS is referred to in: a, and described in: b :
   a. Official Records of the WHO, *205*, 158. WHO, Geneva (1973); *213*, 99. WHO, Geneva (1974); *229*, 115 - 116. WHO, Geneva (1976).
   b. L. BUREMA, F. STURMANS, H. A. VALKENBURG et al.: De Kaunas Rotterdam Intervention Study (KRIS) I. Doelstellingen. T. soc. Geneesk. *52*, 790 - 796 (1974). II. Recrutering van deelnemers aan de trial. T. soc. Geneesk. *52*, 853 - 859 (1974). III. De trialperiode en de follow-up periode. T. soc. Geneesk. *52*, 924 - 933 (1974).
4. G. L. VAN BOETZELAER and H. A. ZONDAG: A rapid modification of the Pearson reaction for serum cholesterol. Clin. Chim. Acta *5*, 943 - 944 (1960).
5. Technicon AutoAnalyzer Method N-16[b].
6. A. ZLATKIS, B. ZAK and A. J. BOYLE: A new method for the determination of serum cholesterol. J. Lab. Clin. Med. *41*, 486 - 492 (1953).
7. S. NATELSON: Microtechnique of Clinical Chemistry, 2[nd] ed. Ch. C. Thomas, Springfield (1961).
8. T. C. HUANG, C. P. CHEN, Y. WEFLER and A. RAFTERY: A stable reagent for the Liebermann-Burchard reaction. Anal. Chem. *33*, 1405 - 1407 (1961).
9. E. HULTMAN: Rapid specific method for determination of aldosaccharides in body fluids. Nature *183*, 108 - 109 (1959).
10. F. GRANDE, D. S. AMATUZIO and S. WADA: Cholesterol measurement in serum and plasma. Clin. Chem. *10*, 619 - 626 (1964).
11. G. R. COOPER, J. R. AKINS, J. S. BIEGELEISEN, J. HAZLEHURST, G. M. MYERS and D. D. BAYSE: Comparison of selected cholesterol colorimetric and enzymatic reference methods. Clin. Chem. *24*, 1046 (1978).
12. A. MATHER and D. D. BAYSE: Performance requirements of reference methods and materials. Clin. Chem. *24*, 1049 (1978).
13. L. L. ABELL, B. B. LEVY, B. B. BRODIE and F. E. KENDALL: A simplified method for the estimation of total serum cholesterol and demonstration of its specificity. J. Biol. Chem. *195*, 357 - 366 (1952).
14. A. H. KADISH, R. L. LITTLE and J. C. STERNBERG: A new and rapid method for the determination of glucose by measurement of rate of oxygen consumption. Clin. Chem. *14*, 116 - 131 (1968).
15. R. B. PASSEY, R. L. GILLUM, J. B. FULLER, F. M. URRY and M. L. GILES: Evaluation and comparison of 10 glucose methods and the Reference Method recommended in the Proposed Product Class Standard (1974), in: Standard Methods of Clinical Chemistry, Vol. 8, G. R. Cooper ed. AACC,

Washington DC (1978).
16. B. E. STATLAND, P. WINKEL and H. BOKELUND: Factors contributing to intra-individual variation of serum constituents: 1. Within-day variation of serum constituents in healthy subjects. Clin. Chem. *19*, 1374 - 1379 (1973).
17. B. E. STATLAND, P. WINKEL and H. BOKELUND: Factors contributing to intra-individual variation of serum constituents: 2. Effect of exercise and diet on variation of serum constituents in healthy subjects. Clin. Chem. *19*, 1380 - 1383 (1973).
18. B. E. STATLAND, H. BOKELUND and P. WINKEL: Factors contributing to intra-individual variation of serum constituents: 4. Effects of posture and tourniquet application on variation of serum constituents in healthy subjects. Clin. Chem. *20*, 1513 - 1519 (1974).
19. The complete data of the KRIS await publication in the Final Report which is now in the draft stage. Drs. P. Das assisted in preparing this section.

Hence the sages did not treat those who were already ill; they instructed those who were not yet ill... To administer medicines for diseases which have already developed is comparable to the behaviour of those who begin to dig a well after they have become thirsty and of those who begin to cast weapons after they have already engaged in battle.

Huang Ti (The Yellow Emperor) (± 2700 b.C.) (1)

CHAPTER 4

SERUM CHOLESTEROL ANALYSIS IN THE CB HEART PROJECT; INTRA- AND INTERLABORATORY QUALITY CONTROL (2)

4.1. Introduction: the CB Heart Project

In the Netherlands a network of tuberculosis clinics - 'Consultations bureaus' (CB's) - has been active, for more than fifty years, in the field of control and prevention of tuberculosis. The CB's have gained wide confidence among the public and the medical profession. In May 1971 the Ministry of Social Affairs and Public Health (today this is the Ministry of Public Health and Environmental Hygiene) decided to support a study on the possibilities to employ CB's in the prevention of cardiovascular disease. A feasibility study began in October 1972 in three such bureaus (Rotterdam, Tilburg, Doetinchem) differing in size, capacity, geographical area and type of public. Full details of the study have been published elsewhere by members of the Steering Committee (3, 4, 5). In January 1976 three additional CB's were included in the project (Amsterdam, Leiden and Maastricht).

A central laboratory was assigned to perform the measurements of one risk factor: the serum cholesterol concentration.

This chapter will describe how the Central Clinical Chemistry Laboratory of the Rotterdam University Hospital 'Dijkzigt' carried out this task.

4.2. Methods and materials

The selection of methods described in the previous chapter (KRIS) led to the same choice for the cholesterol method. But since in the KRIS and the CB Project the external contacts, the arrival of samples and the reporting of results etc. were entirely different, the two procedures were carried out in different parts of the

CKCL. In the CB Heart Project there was less need to emphasize interlaboratory standardization. The results of the analyses have been continuously compared with the WHO Regional Lipid Reference Center in Prague. In contrast with the KRIS a mechanized pipetting step was adopted and serum was used instead of plasma after the difference between the two (section 3.4.1.3.) became apparent. From April 1973 the following procedure has been in use:

A venous blood sample for serum cholesterol determination is taken from one of the arms with little or no stasis whilst the subject is sitting upright. A Terumo-Venoject tube (type T 200-U; plain; 10 ml) is used; the subject's name and laboratory number are written on a sticker which is then attached to the tube.

The blood is allowed to clot for 45 to 60 minutes at room temperature. The tube is then spun in a centrifuge for 10 minutes at 1000g. The serum is pipetted into 5 ml plastic tubes, one hundred of which fit into a polystyrene box, kept in a refrigerator. Samples are collected at the three CB's throughout one week, before overnight shipment to the laboratory by mail. The same identification as described above is used on the tubes containing the serum when shipped to the laboratory. Boxes are delivered at the post office at 17.00 hours and arrive at the CKCL the next morning at 08.00 hours, at which time the contents are still well cooled.

The total serum cholesterol is measured with the 'direct' method of Huang et al. according to the standardized modification NEN 2415 of the NNI, the NACC and RIV (6, 7).

Without prior extraction, 0.1 ml of serum together with 5.0 ml of Liebermann-Burchard reagent (J.T.Baker Chemicals) which is brought to room temperature, are transferred into a reagent tube with a Hook & Tucker diluter-dispenser (Type III). After mixing (Vortex Mixer) the tube is placed in a water bath kept at 25° C. The optical density of the blue-green colour developed over 25 minutes, is read with a photometer (Vitatron Digital DCP) at a wavelenght of 620 nm. Calibration is done with a solution of pure cholesterol (British Drug Houses) in acetic acid (Merck). Several commercial quality control samples with values known to the technician, are included in each run (process control). Until 1975 about 3% of all patient samples were taken at random and analyzed the next day, thus providing for a 'blind' control. Samples with cholesterol levels of 7.7 mmol/l (296 mg %) or more are re-analysed as well, as an extra check on the elevated level. Since February 1975 a completely 'blind' control scheme has been functioning.

Control materials: Monitrol I and I-X (Dade), Monitrol I-X and II-X made for the Massachusetts Society of Pathologists Interlaboratory Quality Assurance Programme (Dade), R.I.V. Control Serum, Elevated Lipid Control Serum (Lederle

Diagnostics) and Seronorm Control Serum (Nyegaard) have been used in the time periods shown in figure 4.1.

4.3. Terminology

The reliability of experimental results has at least two components: precision and accuracy.

In discussing precision, a distinction has been made in this study between repetition of the experiment under constant experimental conditions or as different as possible. In the former case repeatability, in the latter reproducibility is under discussion, but always the same laboratory with the same technicians and method is performing the measurements. The definitions of repeatability and reproducibility presented allow considerable freedom in interpretation. We prefer to determine the conditions under which precision and accuracy are studied (see section 4.4.), rather than try to define these terms unequivocally.

4.4. Control procedures

The main component of quality control is concerned with checks on stability and reproducibility by monitoring precision. The following four procedures are part of control of precision in the CKCL cholesterol analysis.

4.4.1. Process control

Reconstituted, lyophilized control serum from different sources is included in all the series of measurements (runs). All analyses are performed in side-by-side duplicate because all patient samples are treated in that way to rule out large errors, according to a protocol proposed by the Center for Disease Control, USA. The maximum difference allowed between duplicate patient values is 0.3 mmol/l. Until the summer of 1975, three controls were used with mean cholesterol levels of 2.2; 2.7 and 4.2 mmol/l as determined in our laboratory. Since then the 2.7 mmol/l control was replaced by an 'elevated' control with an 8.2 mmol/l cholesterol level, which in turn has been succeeded by other elevated controls. In each run at least two out of the three controls had to yield values within a range of 2 s from the previously determined mean value. However, in the first period a run was accepted when both 'low' controls were slightly out of range, as the 4.2 mmol/l control was considered more important. From the summer of 1975, runs have been discarded and repeated when the 'normal' and the 'elevated' controls have fallen beyond their respective 2 s limits.

The reproducibility results are summarized in Fig. 4.1 and Table 4.1. The smallest dispersion with regard to repeatability can be expected within one analytical run:

Table 4.1. Precision results in serum cholesterol analysis.

Control serum	Period of observation	No. of months	No. of samples	Monthly mean (mmol/l)	s (mmol/l)
Elevated Lipid Control (2915 - 650 R1)	July 1975 - Feb. 1977	20	1007	8.28	0.157
Elevated Lipid Control (2915 - 413)	Feb. 1977 - Apr. 1978	15	748	7.83	0.173
Monitrol (II) (SPXP - 9570)	Feb. 1978 - Dec. 1978	10	508	7.02	0.196
Monitrol I	Mar. 1973 - Apr. 1975	26	318	4.29	0.152
Monitrol I (SPXL - 345)	June 1975 - Mar. 1978	34	1817	4.16	0.115
Monitrol I (SPXL - 372)	Jan. 1978 - Dec. 1978	12	581	4.40	0.136
R.I.V. Control (Serum B)	Mar. 1977 - Dec. 1978	16	561	2.69	0.111
R.I.V. Control (Serum A)	June 1973 - Dec. 1978	57	1932	2.26	0.113

at 'normal' and 'elevated' levels a coefficient of variation of 1% is observed in our laboratory. However, this sample is not 'blind' and can be identified by the technician.

A check on linearity is included in all runs: the calibration is performed with a 10.0 mmol/l standard solution and a 5.0 mmol/l solution is analysed subsequently. The result must be within 0.2 mmol/l from the expected 5.0 mmol/l level. Graph plots of the control results are maintained in the laboratory.

4.4.2. Blind control with duplicate patient serum; repeatability

In all serum shipments arriving at the CKCL from February 1975 onwards, about 10% of all samples have been divided into duplicates, one of which is labelled with a code known to the CB administration only.

The selection of these sets of samples at the CB is carried out as follows. Each week, the serum of the first person whose case history number ends with 0 is taken as the first serum for double determination of cholesterol. The selection of the person occurs at random, as serum cholesterol is measured not only in newly examined persons but also in those belonging to the 'high risk' group who are re-examined at 4, 8 and 12 months, and also in persons with 7.3 mmol/l or more at entry who are re-examined about 3 weeks later. Serum of each tenth person following the first whose case history number ends with 0 is eligible for the blind control.

Sera chosen for the blind control are divided into two approximately equal parts; the second tube is labelled with a new laboratory number and a fictitious name. A separate register is kept at each CB with information about each person selected for double determinations, or triple as described next.

4.4.3. Reproducibility

In 5% of newly-examined persons, selected at random as described above, the same serum is analysed by the CKCL three times: twice for repeatability and once for reproducibility about three weeks later. For this purpose, two venous blood samples are taken from the chosen persons. After centrifugation, the serum from both tubes is pipetted into one tube, mixed and divided into three parts.

4.4.4. Accuracy (interlaboratory control)

The CKCL takes part in the standardization programme offered by WHO through the WHO Regional Lipid Center for Europe (Director: Dr. D. Grafnetter, Prague, Czechoslovakia) (8). Since early 1973 sets of twenty reference serum samples have been received, the reference method values being unknown to the CKCL. The samples are included in routine analytical runs, usually at a rate of two

samples on two different days per week. The evaluation of the results by the reference laboratory includes precision and accuracy, usually at two different levels.

4.5. RESULTS

Precision

4.5.1. Process control

Precision results for these sera are plotted in fig. 4.1., showing monthly averages. These fluctuate over a range that is within 0.4 mmol/l.

4.5.2. Blind control with duplicate patient serum; repeatability.

The results of serum cholesterol determination in duplicate sera for the period February 1975 to December 1978 are given in table 4.2.1.

The differences between the mean values of the two series of sera fluctuated between + 0.05 mmol/l (August 1975) and - 0.11 mmol/l (February 1975). The standard deviation was relatively large during the period February to July 1975 and became smaller from August 1975 and varied less than before, between 0.10 and 0.19 mmol/l. The last column of table 4.2.1. refers to the coefficient of variation. Again, it was higher during the period February to July 1975 (about 3% to 5%) than afterwards (2% to 3%). In September 1978 two outliers caused an increased CV.

4.5.3. Blind control with duplicate patient serum: reproducibility

The results of serum cholesterol determination for reproducibility in patient's sera for the period August 1975 to December 1978 are given in table 4.2.2. This table contains information about reproducibility given in the same way as repeatability in table 4.2.1.

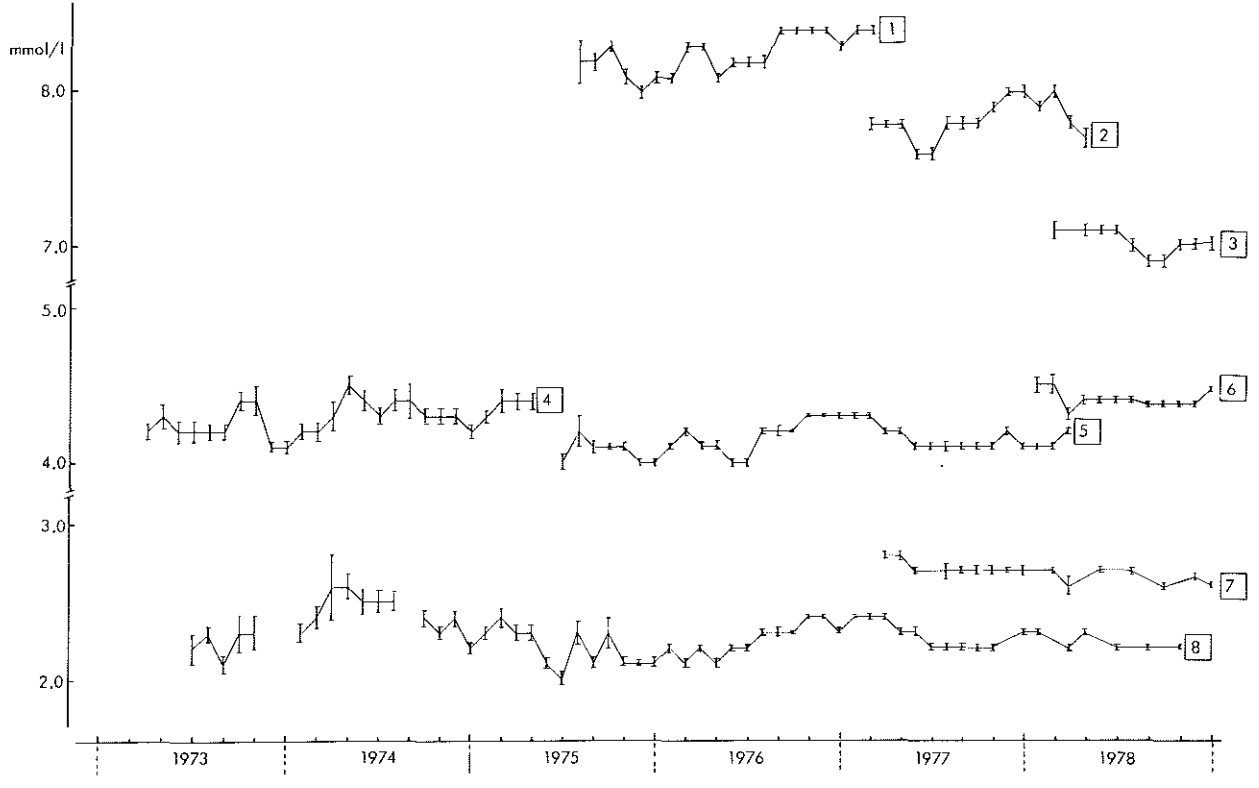
The difference between the mean values of the two series fluctuated between + 0.20 mmol/l (September 1978) and - 0.14 mmol/l (October 1975 and April 1978).

The value for s was, as could only be expected, on average larger than that observed in table 4.2.1. (repeatability). The results of serum cholesterol determination in duplicate sera have also been analysed separately for values <6.2 mmol/l, values between 6.2 and 7.2 mmol/l, and values of 7.3 mmol/l and more. The results for the analytical errors and the coefficients of variation are given in table 4.3.

4.5.4. Accuracy (interlaboratory control)

As already mentioned, several sets of serum samples with reference values unknown to the CKCL have been received from the WHO Regional Lipid Reference

Figure 4.1. Serum controls for cholesterol analysis in CB Heart Project. Monthly $\bar{x} \pm sem$ are shown.



1, 2: Elevated Lipid Control, two lots; 3: Monitrol II-X; 4: Monitrol I; 5, 6: Monitrol I-X, two lots; 7: RIV serum B; 8: RIV serum A.

Table 4.2.1. Blind control with duplicate patient serum.
CB Heart Project, February 1975 to December 1978.

Repeatability

Year and month	No. of pairs	Mean serum cholesterol[1] value (mmol/l)		s[2]	CV (%)
		Chol. I	Chol. II		
1975					
February	38	6.22	6.11	0.20	3.4
March	87	5.94	5.93	0.15	2.6
April	100	5.92	5.86	0.17	2.9
May	81	5.69	5.65	0.28	4.9
June	109	5.53	5.56	0.24	4.3
July	67	5.56	5.49	0.24	4.3
August	100	5.57	5.62	0.14	2.5
September	120	5.95	5.96	0.17	3.1
October	73	5.79	5.76	0.13	2.3
November	81	5.61	5.65	0.12	2.1
December	71	6.19	6.22	0.14	2.2
1976					
January	135	6.00	6.00	0.10	1.6
February	190	5.93	5.93	0.11	2.0
March	187	5.92	5.91	0.17	2.8
April	160	5.65	5.65	0.10	1.9
May	174	5.78	5.80	0.16	2.8
June	171	5.90	5.91	0.15	2.6
July	126	5.98	5.97	0.14	2.3
August	192	5.97	5.98	0.10	1.7
September	205	5.97	5.97	0.12	1.9
October	215	6.19	6.17	0.11	1.9
November	233	6.18	6.19	0.10	1.6
December	187	6.01	6.00	0.13	2.1

[1] Cholesterol I and Cholesterol II are duplicate sera, one of which is labelled with a code known to the CB administration only. The sera are usually examined on the same day.

[2] For calculation of s see equation in table 3.7.

Table 4.2.1. (continued)

Repeatability

Year and month	No. of pairs	Mean serum cholesterol[1] value (mmol/l)		s [2]	CV (%)
		Chol. I	Chol. II		
1977					
January	193	6.08	6.09	0.14	2.5
February	162	6.12	6.08	0.16	2.5
March	223	5.86	5.83	0.12	2.1
April	176	5.79	5.78	0.12	2.1
May	141	5.80	5.80	0.09	1.6
June	181	5.81	5.82	0.19	3.1
July	62	5.69	5.68	0.10	1.8
August	85	5.61	5.55	0.11	1.9
September	109	5.80	5.81	0.15	2.6
October	106	5.86	5.84	0.16	2.7
November	119	6.11	6.09	0.08	1.4
December	106	6.00	6.01	0.10	1.6
1978					
January	108	5.92	5.90	0.07	1.3
February	88	5.98	5.99	0.08	1.4
March	109	5.84	5.85	0.10	1.6
April	91	5.78	5.78	0.11	2.0
May	85	5.80	5.81	0.09	1.6
June	100	5.69	5.69	0.14	2.4
July	75	5.74	5.75	0.11	1.8
August	81	5.90	5.86	0.10	1.7
September	83	5.93	5.93	0.24	4.0
October	124	5.83	5.81	0.11	2.0
November	105	5.90	5.88	0.08	1.3
December	82	6.26	6.26	0.11	1.7

[1] Cholesterol I and Cholesterol II are duplicate sera, one of which is labelled with a code known to the CB administration only. The sera are usually examined on the same day.

[2] For calculation of s see equation in table 3.7.

Table 4.2.2. Blind control with duplicate patient serum.
CB Heart Project, August 1975 to December 1978.

Reproducibility

Year and month	No. of pairs	Mean serum cholesterol[1] value (mmol/l)		s[2]	CV (%)
		Chol. I	Chol. III		
1975					
August	32	5.72	5.72	0.14	2.3
September	38	5.68	5.68	0.28	5.3
October	30	5.74	5.60	0.21	3.8
November	29	5.54	5.54	0.13	2.4
December	28	5.94	5.91	0.21	3.9
1976					
January	65	6.05	6.09	0.19	3.0
February	92	6.05	5.94	0.19	3.1
March	92	5.99	5.92	0.21	3.8
April	79	5.84	5.87	0.14	2.5
May	84	5.71	5.76	0.17	3.0
June	84	5.90	6.08	0.24	4.1
July	84	5.71	5.76	0.17	3.0
August	93	6.04	6.03	0.22	3.7
September	103	6.02	6.02	0.15	2.5
October	103	6.01	6.01	0.15	2.6
November	115	6.32	6.26	0.13	2.1
December	93	5.96	6.04	0.20	3.2

[1] Cholesterol I and Cholesterol III are duplicate sera, one of which is labelled with a code known to the CB administration only. Cholesterol III is measured about three weeks later than Cholesterol I.
[2] For calculation of s see equation in table 3.7.

Table 4.2.2. (continued)

Reproducibility

Year and month	No. of pairs	Mean serum cholesterol[1] value (mmol/l)		s [2]	CV (%)
		Chol. I	Chol. III		
1977					
January	93	6.02	5.98	0.15	2.6
February	75	6.06	5.96	0.20	3.2
March	107	5.87	5.85	0.17	2.8
April	83	5.95	5.90	0.14	2.4
May	66	5.87	5.91	0.15	2.6
June	88	5.81	5.84	0.27	4.4
July	42	5.62	5.60	0.17	3.2
August	82	5.61	5.63	0.14	2.4
September	97	5.72	5.69	0.18	3.2
October	98	5.87	5.89	0.16	2.6
November	108	6.11	6.06	0.14	2.3
December	96	6.03	5.96	0.14	2.4
1978					
January	95	5.98	6.05	0.17	2.9
February	78	5.98	5.86	0.20	3.3
March	98	5.80	5.84	0.29	4.8
April	83	5.81	5.67	0.25	4.4
May	77	5.81	5.85	0.20	3.4
June	91	5.73	5.73	0.17	3.0
July	64	5.73	5.67	0.18	3.6
August	75	5.82	5.86	0.21	3.6
September	73	5.93	6.13	0.36	6.0
October	102	5.83	5.76	0.17	3.0
November	106	5.89	5.96	0.13	2.3
December	81	6.24	6.15	0.23	3.9

[1]Cholesterol I and Cholesterol III are duplicate sera, one of which is labelled with a code known to the CB administration only. Cholesterol III is measured about three weeks later than Cholesterol I.

[2]For calculation of s see equation in table 3.7.

Table 4.3. Blind repeatability with duplicate patient serum. Standard deviation (s) and coefficient variation in samples with 'low' (<6.2 mmol/l), 'borderline' (6.2 - 7.2 mmol/l) and 'elevated' (⩾7.3 mmol/l) serum cholesterol values.
CB Heart Project, June 1975 to December 1976.

Period of observation	s (mmol/l) Serum cholesterol			CV (%) Serum cholesterol		
	Low	Borderline	Elevated	Low	Borderline	Elevated
Apr.-June 1975	0.22 (195)	0.27 (67)	0.23 (28)	4.2	4.1	2.7
July-Sept. 1975	0.17 (189)	0.20 (68)	0.21 (30)	3.3	3.2	2.8
Oct.-Dec. 1975	0.13 (142)	0.13 (51)	0.13 (31)	2.4	1.9	1.7
Jan.-Mar. 1976	0.13 (313)	0.14 (135)	0.12 (64)	2.4	2.3	1.6
Apr.-June 1976	0.15 (331)	0.14 (126)	0.11 (48)	2,7	2.1	1.4
July-Sept. 1976	0.10 (299)	0.15 (158)	0.11 (66)	1.9	2.3	1.4
Oct.-Dec. 1976	0.11 (335)	0.12 (199)	0.12 (100)	2.0	1.8	1.5

Figures given in brackets are the numbers of pairs.

Center for Europe during the study. Each survey set usually contained 10 lyophilized samples with 'low' (4.5 to 5.0 mmol/l) and 10 samples with 'elevated' (7.5 to 10.0 mmol/l) levels of cholesterol. The reproducibility reported periodically did agree quite well with the intralaboratory results (see table 4.4. for s and CV values).

The WHO Reference Center reports also include average differences between the 'reference method values' and those obtained in the CKCL. These biases, expressed as percentages of the target values, are given for the 'low' and 'high' pools for each survey set. For the results see also figure 4.2.

The table shows that many different reference serum pools have been employed. All of them behave slightly different in the Huang determination as will be discussed in more detail in Chapter 6. Sometimes corrections have been introduced by the Reference Center for matrix effects in order to obtain a more consistently arranged graph. Pool B and C are uncorrected. During October 1973 a comparison was made by sending 22 frozen patient samples to the reference laboratory. Our bias was +5.3%, in good agreement with survey set II. Between February and July 1974 a special set of 42 frozen sera was analyzed. Our bias was +7.8%(2), a bit more than the 3,8% appearing in the corresponding survey IV (compare table 4.4.; June-Oct 1974).

Late in 1975 pools L and M were introduced and in September showed a bias of +12.1 en 10.4% respectively (compare figure 6.1.). In December the much higher bias in L and M (as compared with the bias B and C) was confirmed and to counterbalance this matrix effect the target values were increased 8%. It was believed that such corrections would make the bias of our method close to that for human serum.

During 1976 pool N was issued but not continued because our triglycerides method ran into problems with this serum. Pools T and U in April had a bias of 11.4 and 7.2%. Corrections applied were 10% and 6% (increase of target values); discussion: section 6.4.1. In 1978 several new pools were tested. This will be described in more detail in section 6.4.

4.6. Seasonal variations

The table 4.5. and figure 4.3. depict the mean cholesterol values per month in men and women at baseline (at first entry into CB) examinitions. The number of persons examined each month are as a rule many hundreds and we assume these groups to be fairly comparable, although this is not exactly true. For example let us look at the age distribution of the men and women examined in November 1974 and November 1975, table 4.6.

Table 4.4. Interlaboratory quality control results as reported by the WHO Regional Lipid Reference Center for Europe, Prague.
Reproducibility and bias, 1973 - 1978

Period of observation	Pool	$\bar{x}$ (mmol/l)	s	CV (%)	Bias (%)	No. of samples	Reference method value	s of the bias: s_b [1]
Set I								
May-June	B	4.74	0.23	4.8	+0.7	10	4.71	1.60
1973	C	7.55	0.31	4.1	+0.4	10	7.53	1.33
Set II								
July-Sept.	B	4.94	0.23	4.6	+5.3	10	4.71	1.58
1973	C	8.07	0.24	3.0	+7.2	10	7.53	1.03
Set III								
Nov. 1973 -	B	4.74	0.08	1.7	+0.7	10	4.71	0.54
Jan. 1974	C	7.64	0.15	2.0	+1.5	10	7.53	0.67
Set IV								
June-Oct.	B	4.88	0.28	5.7	+3.7	10	4.71	1.91
1974	C	7.80	0.40	5.1	+3.6	10	7.53	1.70
Set V								
Dec. 1974 -	B	4.83	0.16	3.2	+2.7	10	4.71	1.15
Febr. 1975	C	7.72	0.19	2.4	+2.6	10	7.53	0.83
Set VI								
Apr.-June	B	4.45	0.24	5.3	-5.4	10	4.71	1.63
1975	C	7.53	0.20	2.6	+0.2	10	7.53	0.87
Set CT 1								
Aug.-Oct.	L	5.25	0.11	2.1	+2.9	10	5.05	0.77
1975	M	10.38	0.19	1.8	+1.8	10	10.13	0.61

Table 4.4. (continued)

Period of observation	Pool	$\bar{x}$ (mmol/l)	s	CV (%)	Bias (%)	No. of samples	Reference method value	s of the [1] bias: s_b
Set CT 2								
Febr.-March	L	5.19	0.14	2.8	+ 2.7	10	5.05	0.94
1976	M	10.45	0.17	1.6	+ 3.0	10	10.13	0.53
Set CT 3								
Apr.-May	L	5.02	0.16	3.1	- 0.7	10	5.05	1.06
1976	M	10.17	0.16	1.5	+ 0.3	10	10.13	0.52
Set CT 4								
June-July	L	5.09	0.10	2.0	+ 0.8	10	5.05	0.71
1976	M	10.17	0.17	1.6	+ 0.3	10	10.13	0.55
Set CT 5								
Sept.-Oct. 1976	N	5.07	0.12	2.9	+ 6.0	20	4.78	0.91
Set CT 6								
Dec. 1976 - Jan. 1977	N	5.18	0.12	2.2	+ 8.4	20	4.78	0.90
Set CT 7								
Febr. 1977	N	5.18	0.13	2.6	+ 8.4	10	4.78	0.97
	P	5.45	0.14	2.7	+ 4.0	10	5.24	0.93
Set CT 8								
Apr.-May	T	3.65	0.09	2.6	- 1.1	7	3.70	0.94
1977	U	8.94	0.13	1.4	- 2.0	7	9.13	0.50
Set CT 9								
July-August	T	3.55	0.12	3.3	- 1.6	10	3.70	1.16
1977	U	8.82	0.21	2.4	- 3.3	10	9.13	0.76

Table 4.4. (continued)

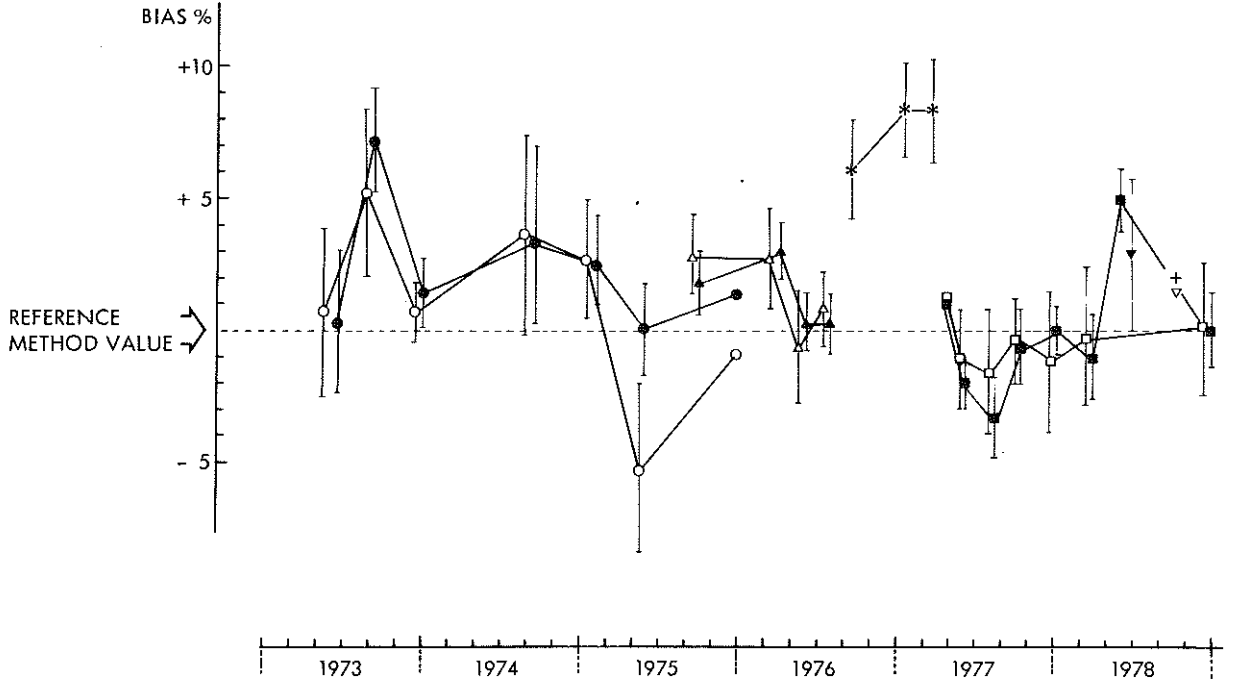
Period of observation	Pool	$\bar{x}$ (mmol/l)	s	CV (%)	Bias (%)	No. of samples	Reference method value	s of the [1] bias: s_b
Set CT 10								
Sept.-Oct.	T	3.59	0.07	2.0	- 0.4	7	3.70	0.80
1977	U	9.07	0.20	2.2	- 0.6	7	9.13	0.73
Set CT 11								
Nov. 1977 -	T	3.65	0.14	3.8	- 1.2	7	3.70	1.32
Jan. 1978	U	9.12	0.11	1.2	- 0.1	8	9.13	0.44
Set CT 12								
March 1978	T	3.60	0.14	3.9	- 0.2	8	3.70	1.31
	U	9.03	0.22	2.4	- 1.0	8	9.13	0.79
Set CT 13								
May-June	AA	4.13	0.16	3.9	+ 2.9	7	3.83	1.43
1978	AB	5.84	0.13	2.2	+ 7.5	7	5.43	0.85
	U	9.32	0.16	1.8	+ 5.1	7	9.13	0.60
Set CT 14								
Sept.-Oct.	AC	4.14	0.13	3.1	+ 1.6	9	-	-
1978	AE	8.00	0.19	2.4	+ 2.0	10	-	-
Set CT 15								
Dec. 1978	T	3.70	0.16	3.5	0	10	3.70	1.30
	U	9.12	0.23	2.3	- 0.1	10	9.13	0.70

[1] The standard deviation of the bias was estimated with the following equation

$$s_b^2 \approx \frac{100}{\bar{y}} \cdot \frac{1}{n} s^2 + (1 + \frac{b}{100})^2 \, S_{\bar{y}}^2 \ ,$$

in which: n = number of samples; $\bar{y}$ = reference method value; s = our standard deviation; $S_{\bar{y}}$ = s of reference method; b = bias

Figure 4.2. Laboratory accuracy with reference serum, IKEM, Prague.



pools: B (o), C (●), L (Δ), M (▲), N (*), T (□), U (■), AA (▽), AE (+). Each time $\bar{x} \pm 2$ sem are given and 7 - 10 analyses are represented.

Table 4.5. Monthly mean serum cholesterol values.

		Men			Women		
		n	$\bar{x}$	s	n	$\bar{x}$	s
1973	April	261	5.95	1.357	247	5.86	0.943
	May	257	5.77	1.090	269	5.70	0.820
	June	228	6.08	1.253	247	5.66	0.786
	July	130	5.55	1.209	124	5.38	0.779
	August	160	5.51	1.075	180	5.33	0.805
	September	174	5.84	1.068	183	5.70	0.812
	October	238	6.06	1.203	274	5.67	0.828
	November	187	5.84	1.121	237	5.55	0.770
	December	97	5.51	0.965	118	5.43	0.652
1974	January	260	5.98	1.113	329	5.66	0.907
	February	204	5.81	1.285	259	5.37	0.805
	March*	421	5.91	1.231	478	5.69	0.875
	April	240	6.01	1.224	186	5.66	1.146
	May	198	5.74	1.224	189	5.57	1.059
	June	161	5.87	1.040	198	5.54	0.999
	July	130	6.13	1.482	110	5.74	1.133
	August	293	6.03	1.147	300	5.67	1.178
	September	283	5.81	1.043	291	5.59	1.024
	October	170	6.03	1.108	206	5.52	0.919
	November	252	6.02	1.397	293	5.56	1.027
	December	194	5.68	1.365	187	5.37	1.080

*including 15-3 - 31-3-1973.

Table 4.5. (continued)

		Men			Women		
		n	$\bar{x}$	s	n	$\bar{x}$	s
1975	January	206	6.02	1.177	222	5.47	0.983
	February	209	5.89	1.113	260	5.48	0.951
	March	200	5.73	1.131	206	5.44	0.990
	April	259	5.90	1.239	267	5.48	0.931
	May	211	5.39	1.046	201	5.09	0.964
	June	365	5.53	1.165	356	5.22	1.132
	July	192	5.53	1.053	184	5.27	0.868
	August	248	5.47	1.071	279	5.27	1.002
	September	349	5.68	1.214	350	5.51	0.992
	October	182	5.45	1.025	178	5.38	1.134
	November	238	5.50	1.034	251	5.12	0.808
	December	212	5.79	1.092	212	5.44	1.107
1976	January	378	5.82	1.108	376	5.57	1.125
	February	410	5.82	1.114	398	5.48	1.057
	March	504	5.85	1.122	534	5.42	1.017
	April	461	5.72	1.031	465	5.36	1.014
	May	552	5.76	1.104	544	5.38	0.980
	June	576	5.91	1.104	531	5.39	0.922
	July	429	5.92	1.181	414	5.59	1.017
	August	597	6.00	1.075	631	5.61	0.955
	September	388	6.15	1.221	472	5.67	1.065
	October	459	6.15	1.071	520	5.87	1.026
	November	630	6.11	1.079	637	5.68	0.959
	December	517	5.91	1.114	568	5.55	0.953

Figure 4.3. Monthly mean serum cholesterol values in men (●) and women (o) at baseline examination in the CB Heart Project.

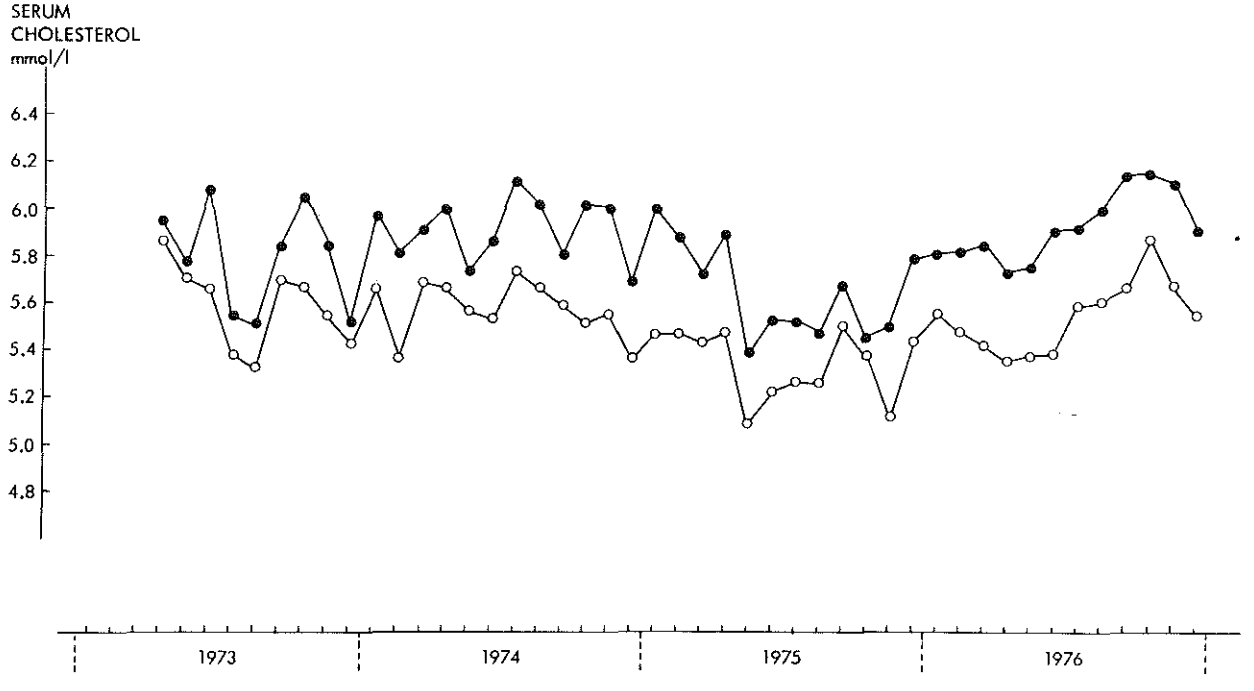


Table 4.6. Examples of cholesterol values in groups of men and women.

	November 1974		November 1975		
	n	$\bar{x}$	n	$\bar{x}$	expected from 1974
Men					
20 - 29 yrs	33	4.94	34	5.24	
30 - 39 yrs	81	5.90	145	5.48	
40 - 49 yrs	138	6.34	59	5.71	
Total	252	6.02	238	5.50	5.87
Women					
20 - 29 yrs	32	4.87	36	5.09	
30 - 39 yrs	108	5.53	146	5.07	
40 - 49 yrs	153	5.73	69	5.24	
Total	293	5.65	251	5.12	5.49

Number examined: n Mean serum cholesterol: $\bar{x}$

This example shows how the age group 40 - 49 dominates in November 1974 and in the next year the group aged 30 - 39 years is larger. At the same time it can be noted how the cholesterol level in the youngest group increased while at the same time it has become lower in the 30 - 39 years old subjects. The oldest group again has a lower level.

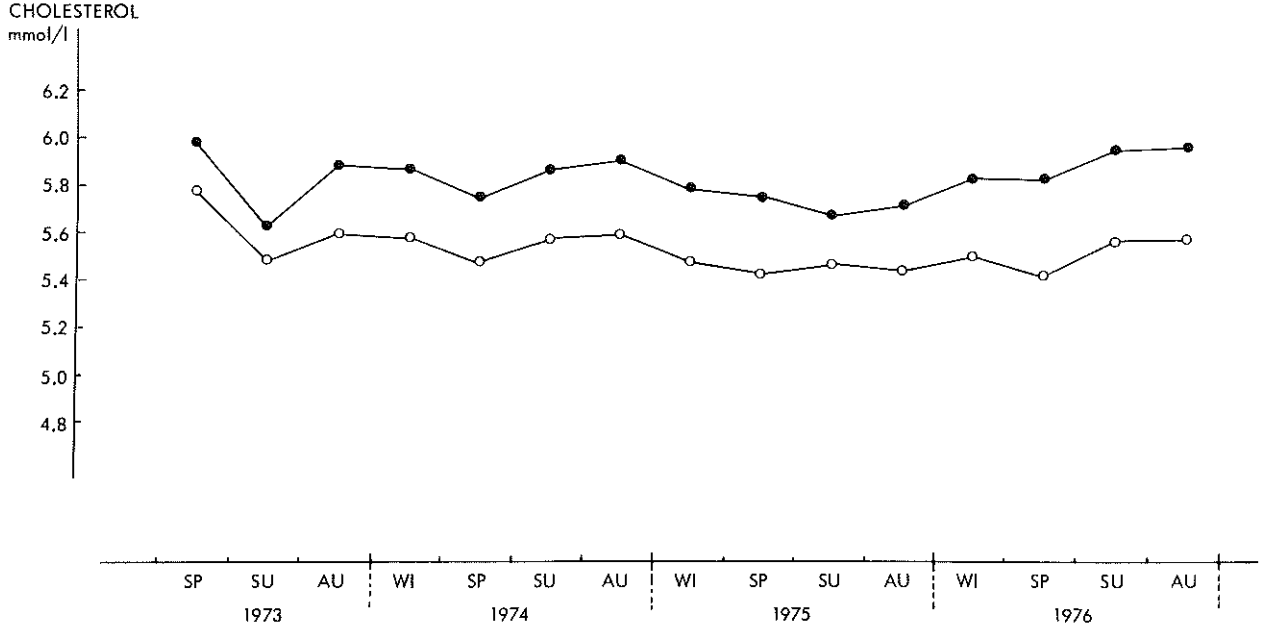
So we must realize that the selection of subjects from certain age groups and from various cities introduces variables which make an exact explanation for all fluctuations seen in figure 4.3. a rather tricky matter.

According to the data of figure 4.1. the laboratory measured about 0.2 mmol/l lower in 1975 when compared with 1974. The lower mean age in November 1975 further decreases the average cholesterol level. If for example the levels from November 1974 and the numbers of subjects examined a year later are combined, the expected cholesterol values in men become 5.87 mmol/l and in women 5.49 mmol/l.

In the literature the term 'seasonal variations' is often used and in accordance with this we took the values of three months periods together, and also those of the quality control sera. The latter value enabled us to correct the average levels in the studied population for laboratory drift.

The seasonal variations - figure 4.4. - now appear to be very small. These results are further discussed in section 6.5.

Figure 4.4. Seasonal mean serum cholesterol values in men (●) and women (o) at baseline examination, corrected for laboratory variations. CB Heart Project.



4.7. DISCUSSION

The results of our process control summarized in fig. 4.1. and table 4.1. demonstrate a fairly constant level of performance and little long-term drift, although some variations do occur. It is interesting to note the rather constant s-values over the whole range of concentrations. It has to be stressed that when the introduction of 'known' samples is not performed by a 'blind' procedure the laboratory precision is easily overestimated (9). The same holds true for automated methods; accuracy is not influenced as much (10). Our built-in control with duplicate patients specimen is firstly an absolute 'blind' procedure and, secondly, uses human material only. The table 4.2.1. shows that the analytical error (s) for repeatability was between 0.07 and 0.28 mmol/l. For reproducibility (table 4.2.2.) these values were between 0.13 and 0.36 mmol/l. Most high values reflect the presence of one or a few outliers. For example: two samples analyzed and having 5.1 and 6.5 mmol/l, in the next analysis show the following cholesterol values: 6.6 and 5.0. A mix-up of samples in such a case is hard to trace - it may have happened in the clerical or in the laboratory area - and the large difference is not deleted from calculations.

The accuracy of our method was assessed by analyzing lyophilized reference sera provided by the WHO Regional Lipid Reference Center in Prague. The expected positive bias of our 'direct' Liebermann-Burchard technique does appear and not always does it remain within the WHO limits of ± 5% (table 4.4.). The low level in survey VI (May-June 1975) was caused by unsatisfactory reagents. Comparisons of two lots of this reagent with our own preparation confirmed a low level of results for this particular reagent lot, although all calibrations were done with identical standards. The high cost and several operational problems connected with the shipment of frozen human serum make its usage difficult. It would, however, be highly desirable to do it. Our results indicate a material-related bias instead of an absolute bias. This will be discussed in Chapter 6. In our opinion a 'blind' continuous quality control is feasible in a large study and measures accurately the extent of laboratory errors while it remains free from several short-comings of lyophilized control preparations (matrix, vial-to-vial variation, changed solubility, need to reconstitute). We do not find seasonal variations in the study population but there are fluctuations (after correction) in the monthly averages within a range of about 0.3 mmol/l. These variations cannot be attributed to analytical instability as they are corrected by means of the results in figure 4.1. Indeed the low values in May and June 1975 are, at least in part, a result of a downward change in analytical bias, which also is revealed by figure 4.2. Laboratory bias cannot fully explain the rising curve through 1976.

Further discussion of the findings appears in chapter 6.

REFERENCES

Chapter 4

1. quoted in: A. G. OLSSON: Studies in asymptomatic primary hyperlipidemia. Clinical, biochemical and physiological investigations. Acta Med. Scand. Suppl. 581, 1 (1975).
2. This chapter is based in part on a. and b.:
   a. G. J. M. BOERMA, K. STYBLO and B. LEIJNSE: Serum cholesterol analysis in the CB Heart Project. Intra- and interlaboratory continuous quality control. Clin. Chim. Acta *83*, 211 - 222 (1978).
   b. K. STYBLO, G. J. M. BOERMA and A. C. ARNTZENIUS: CB Heart Project in the Netherlands. Cholesterol. Hart Bull. *7*, 47 - 54 (1976).
3. MEMBERS OF THE STEERING COMMITTEE OF THE CB HEART PROJECT: CB Heart Project. Hart Bull. *7*, (2), 38 - 81 (1976).
4. MEMBERS OF THE STEERING COMMITTEE OF THE CB HEART PROJECT and I. SUTHERLAND: CB Heart Project - intervention results. Hart Bull. *8*, (2), 33 - 71 (1977).
5. A. C. ARNTZENIUS: The Consultation Bureau Project (CB Project) in the Netherlands. A feasibility study on evaluation and management of coronary risk factors. J. de Coeur et des Vaiss. 143 - 147 (1976).
6. T. C. HUANG, C. P. CHEN, V. WEFLER and A. RAFTERY: A stable reagent for the Liebermann-Burchard reaction. Anal. Chem. *33*, 1405 - 1407 (1961).
7. NEDERLANDS NORMALISATIE-INSTITUUT: Procedure NEN 2415. Klinisch-chemisch onderzoek. Cholesterol (totaal) in serum, fotometrisch zonder extractie (1975).
8. D. GRAFNETTER: World Health Organization (WHO) coordinated quality control in the lipid laboratory. Giorn. Arterioscl. *2*, 113 - 128 (1977).
9. J. R. ALLEN, R. EARP, E. C. FARRELL and H.-D. GRÜMER: Analytical bias in a quality control scheme. Clin. Chem. *15*, 1039 - 1044 (1969).
10. M. S. WEINBERG and R. N. BARNETT: Absence of analytical bias in a quality control program. Am. J. Clin. Pathol. *38*, 468 - 472 (1968).

The unsolved problem of establishing long-term stability in accuracy and precision stands in the way of extended longitudinal studies, and prevents proper utilization of the mass of information accumulating in the records of CC laboratories all over the world.
IFCC Committee on Standards; Expert Panel on Nomenclature and Principles of Quality Control in Clinical Chemistry. (1977) (1)

CHAPTER 5

STANDARDIZATION OF CHOLESTEROL ANALYSES IN THE NETHERLANDS

5.1. Introduction

The history of interlaboratory comparisons in clinical chemistry begins in 1947 with the survey of Belk and Sunderman (2) in the USA and on an international scale in 1956 with the trial organized in England by Wootton (3). The organization of many comparative studies has been undertaken in the past 30 years by professional associations, federal institutions and commercial organizations. Between 1957 and 1965 the Dutch RIV has, yearly, performed trials by distribution of equine serum samples or bovine albumin solutions spiked with several biochemical compounds and electrolytes (4, 5). Invariably the results of the trials demonstrated a wide variability in results. Deviations of 100% from the mean values were no exception! Besides we must reckon with the fact that several of these results were obtained in analyses that were done with extra care and not in daily routine runs (4). It has always been hoped that such information would stimulate laboratory directors to abandon obsolete methods and to do something about intralaboratory differences. An unmistakable trend towards better comparability is present and an ever increasing number of laboratories take part in one or several trials (6).

Since 1972 in the Netherlands a new scheme with 6 surveys per year is functioning under the auspices of the Foundation for Quality Control in Clinical Chemical Hospital Laboratories (7). The intention remains the same: the trials are educational and the clinical chemist himself must improve his techniques in such a way as to increase comparability and if possible accuracy. Such improvements are brought about slowly. Much more than average attention and aiming at a few analytical techniques at a time are requirements for a faster process. As a stimulus,

to speed up the standardization, a purposive set-up is necessary.

In 1976 the Foundation for Quality Control of Chemical Analysis in Epidemiological Investigations (Stichting Kwaliteitsbewaking Chemische Analyse ten behoeve van epidemiologisch onderzoek: KCA) began to implement the first national standardization programme in clinical chemistry.

An initiative was taken by Prof. Dr. C. den Hartog, Wageningen, who assembled a first meeting in December 1975 to discuss the non-comparable cholesterol findings. Subsequent conferences led to the birth of the KCA in the office of the Netherlands Heart Foundation (8). Several associations take part in it: The Netherlands Heart Foundation, The Netherlands Association for Clinical Chemistry, The Netherlands Association for Occupational Health and the Foundation for Quality Control in Clinical Chemical Hospital Laboratories furnish board members. The Ministry of Health and Environmental Protection, The Royal Netherlands Medical Association, The Royal Netherlands Tuberculosis Association appoint advisers in the board. The scientific advisory board of the foundation decided to adopt the WHO Cooperative Lipid Standardization Programme as a model for this enterprise (9).

After having consulted several experts from CDC (Dr. A. Mather, Dr. J.H. Boutwell and Dr. G.R. Cooper visited Rotterdam) the following plans were developed with the aim of increasing interlaboratory precision and overall accuracy of the methods used in our screening projects. In the first year 31 laboratories enlisted in the project.

5.2. KCA Cholesterol Standardization Programme

Part 1 Participants will, in routine runs, analyze 30 samples from each of three serum pools. The precision and the accuracy are then evaluated according to the protocol as used in CDC's own Lipid Standardization Programme. The standard deviation will reveal intralaboratory proficiency; the WHO narrow limits (table 2.7.) will be applied.
The pools have been analyzed in the Lipid Reference Laboratory so that comparisons with the reference method values will enable us to determine accuracy. Matrix effects may pose a problem here.

Part 2 Participants with sufficient precision will use the above mentioned serum pools as calibrators. The positive bias of the direct LB-reaction (often 10 - 15% higher than target values) may be reduced to less than ±5% (WHO narrow limits).

Part 3 Continuous surveillance by the KCA with a number of serum controls, to be mailed, or delivered in person for on the spot analysis.

5.2.1. Methods of the participants

Since all laboratories had at the start submitted a questionnaire with details on their methods, it was known that considerable variation among results could be expected. According to the questionnaires filled in by all KCA laboratories quality control was sketchy in a number of cases. Important shortcomings are listed below in table 5.1.

Table 5.1. Deficiencies in quality control procedures.

Problem	Prevalence
no serum controls in every run	5
low level serum control only	3
high level serum control only	1
no linearity check with more than 1 standard	11
no blind controls	(16)
(it is not always clear what is meant with 'blind', but some people obviously regard 'unknown' controls as blind, even when their identity as a control is known).	
no external comparison	9
commercial serum calibrators or aqueous standards	7
(this may work very well, but not without extensive checking upon the matrix effects in comparison with human serum).	

Any improvement of comparability among laboratories must begin within the laboratories. The factors just listed must be taken into consideration to ensure internal repeatability and stability of the method.

Quality control, preferably with blind samples, the identity of which is entirely unknown to the laboratory, over a sufficient period of time will reveal how the situation is. Only with such information do we have a starting-point to try and bring about a favourable change in the bias and to check whether such a change has had the desired result. After ample discussion the KCA has recommended the KRIS laboratory protocol as an excellent example of a controlled analytical method (Appendix I). This protocol was then distributed to the participants. During the first year (1977) 31 laboratories entered part I of the standardization programme. Three laboratories used two different methods as they were in the process of changing. Among these laboratories there were 10 hospitals, 11 occupational health services and 10 research or private institutes, table 5.2.

Table 5.2. Methods used by KCA participants.

Methods	labs.	group
1. Netherlands recommended Huang method NEN 2415	18	A
2. Huang on AutoAnalyzer (Technicon Corp.)	3	
3. Huang on SMAC (Technicon Corp.)	2	
4. Huang on AKES (Vitatron Co.)	1	
5. Huang calibrated with reference serum	1	B
6. Direct Liebermann-Burchard on Hycel 1600	1	
7. Manual Abell-Kendall	1	
Total Liebermann-Burchard reaction	27	
8. Manual enzymatic procedure	3	
9. AutoAnalyzer II enzymatic procedure	2	
10. SMAC enzymatic procedure	1	C
11. Centrifichem enzymatic procedure (Roche)	1	
Total ChOD reaction	7	

The advisory board has encouraged the continuing use of the direct LB-reaction, at least for a few years. The recommended Huang modification is by far the most frequently adopted procedure in the Netherlands. Recently, however, the enzymatic reagents have been gaining increasing popularity. The advocated advantages are: specificity, absence of unpleasant odour and corrosiveness and rather simple analytical manipulations. However, there are now several products commercially available with different reagent composition. In addition reaction conditions exist sometimes where specificity of the method may be lost. Surely the standardization among laboratories will not be fostered when part of them changes over from LB to a number of different enzymatic methods. This does not alter the possibility that eventually a reliable enzymatic procedure may become the method of choice for everybody.

5.2.2. Materials

Pooled serum, from samples sent to the RIV (Dr.A.H. Holtz, Drs. J.B.A. Terlingen) for Hepatitis B-antigen testing and found negative, was used to harvest human lipoproteins employing an alcoholic precipitation technique (10). This procedure was carried out by Drs. C.M. van Gent, Gaubius Institute-TNO, Leiden.

Concentrated lipid fractions were added to equine serum which itself had a cholesterol level of approximately 2.0 mmol/l (Pool A_1). After mixing, a filtration was done to sterilize the pools. Several filters clogged and the entire filtration took many hours. When the first pools appeared not to be sterile the entire procedure was repeated. Finally pools A_1, A_2 and A_3 were ready and dispensed in 7000 glass ampoules containing 0.6 ml of serum. Later three more pools (A_4, A_5, A_6) were prepared at the RIV, employing the Ca^{++}- heparine separation technique of Burnstein (11) for isolation of a human lipid fraction to be added to the equine serum. Recently human serum pools, numbered '7, 8, 9 and 10' were added (Dr. J.C. Koedam).

5.3. KCA Programme part I

To each laboratory 90 ampoules of pools A_1, A_2 and A_3 were mailed in a polystyrene container with dry ice. All material had to be kept frozen untill the time of analysis. Then, three ampoules were to be taken from the freezer, thawed and homogenized well, and analyzed in a routine run. These measurements had to be spread over a minimum of 15 days. All results were then sent back and evaluated with a simple calculation of means and standard deviations (overall, within-run and among runs) like the CDC Standardization Programme requires in its part I.

5.3.1. Precision:

In accordance with the WHO criteria the overall s for these pools should be below 0.18 mmol/l (7 mg %) with regard to pools A_1 and A_2 whereas s for pool A_3 had to be smaller than 0.21 mmol/l (8 mg %). These requirements had to be met for entering part II. Bias in this first part was not marked because in part II of the programme this would change anyhow for the majority of the participants. Table 5.3 presents a summary of the results.

It turned out that for one or more pools in 13 cases $s_{overall}$ was unsatisfactory and 21 methods qualified for part II. A second serum shipment for part I was sent to the first 13 to repeat the series after reviewing their analytical procedure. All of them qualified in the second series.

5.3.2. First results of part I:

It appears that the comparability within the manual Huang group (A) and within the enzymatic group (C) is not bad in pool A_1 and A_2. The overall spread of results among all methods is of course very wide in all pools.

As could be expected the LB-reaction yields elevated values. This bias is relatively high at lower cholesterol levels and highest in group A. Group B includes several methods employing serum calibrators which may explain a lower bias as

Table 5.3. Results of KCA programme, part I.

			group (compare table 5.2.)			
		CDC reference method	A n = 18	B n = 9	C n = 7	All n = 34
pool A_1	mean (mmol/l)	2.14	2.45	2.35	2.00	2.33
	s (mmol/l)	0.054	0.092	0.397	0.160	0.279
	bias (%)	0	+ 14.5	+ 9.8	- 6.3	+ 9.0
pool A_2	mean	4.16	4.54	4.44	4.03	4.41
	s	0.088	0.194	0.407	0.182	0.325
	bias	0	+ 9.3	+ 6.7	- 3.0	+ 6.1
pool A_3	mean	6.85	7.40	7.09	6.64	7.16
	s	0.101	0.314	0.535	0.280	0.475
	bias	0	+ 8.1	+ 3.5	- 3.1	+ 4.6

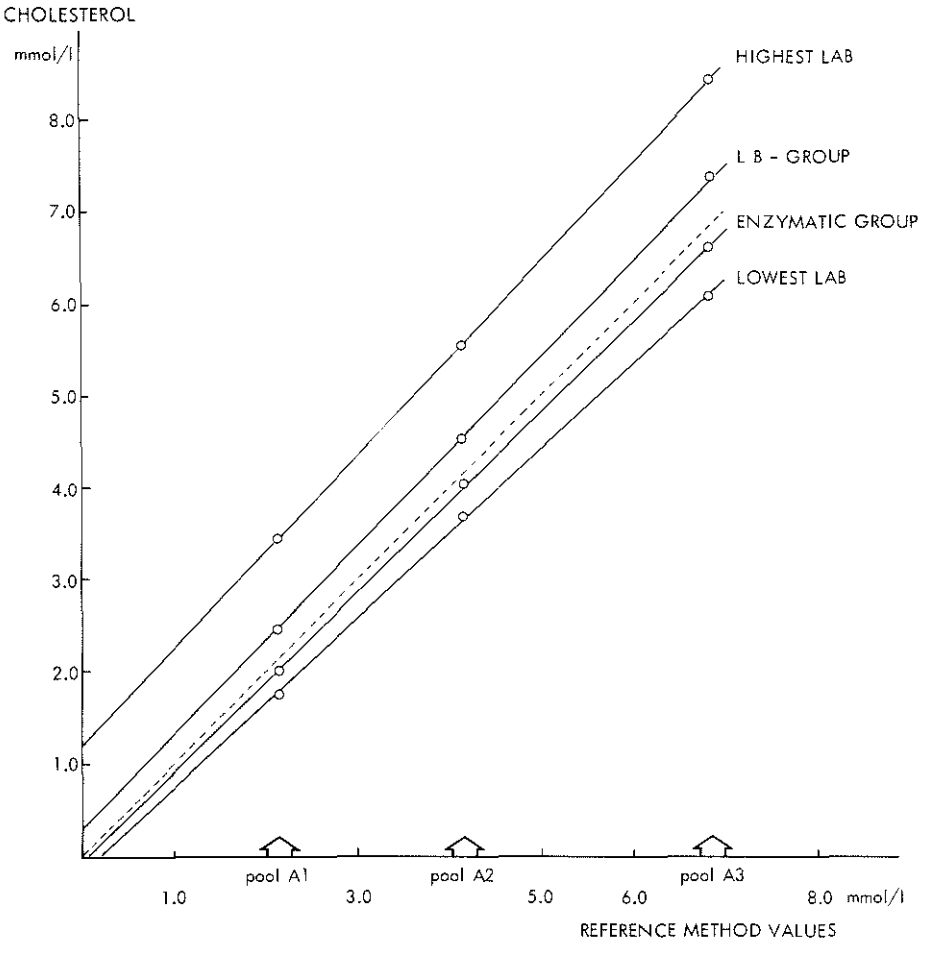
compared with group A. The enzymatic method gives low results: an experience often seen when measuring serum that has been processed to change the usual physiological levels of constituents or to increase stability (lyophilization) (12, 13, 14).

5.3.3. Bias

Figure 5.1. shows the relationship between several methods and the reference method. There appears to be an absolute and a relative component that together create the difference. Each individual laboratory of course will need a special correction suited for their own need. As shown in figure 5.2. some have produced a calibration curve through the origin (no. 10) and others (no. 1^a) have parallel upward shift with a constant bias in each of the three samples. No. 13 has a negative bias for the high serum and positive bias for the lower values.

When measurements are repeated several months later the calibration curve may have shifted significantly over a range equal for all samples ($1^a \longrightarrow 1^b$) or the slope may change ($11 \longrightarrow 11^a$); Figure 5.3. An other shift is seen in the highest laboratory results shown below. These intralaboratory results illustrate the necessity to standardize a method with internal quality control before interlaboratory comparability is within reach. A search for the highest and lowest values among the initial and the repeated analyses in part I resulted in a surprise: the lowest values were produced in a direct LB-reaction: table 5.4.

Figure 5.1. Mean results of KCA-laboratories; LB-group, enzymatic group and highest and lowest results.



Every laboratory performed 30 analyses for each pool.

Figure 5.2. Examples of KCA-laboratory results. Each dot represents the mean of 30 samples. Bias varies from proportional to constant values.

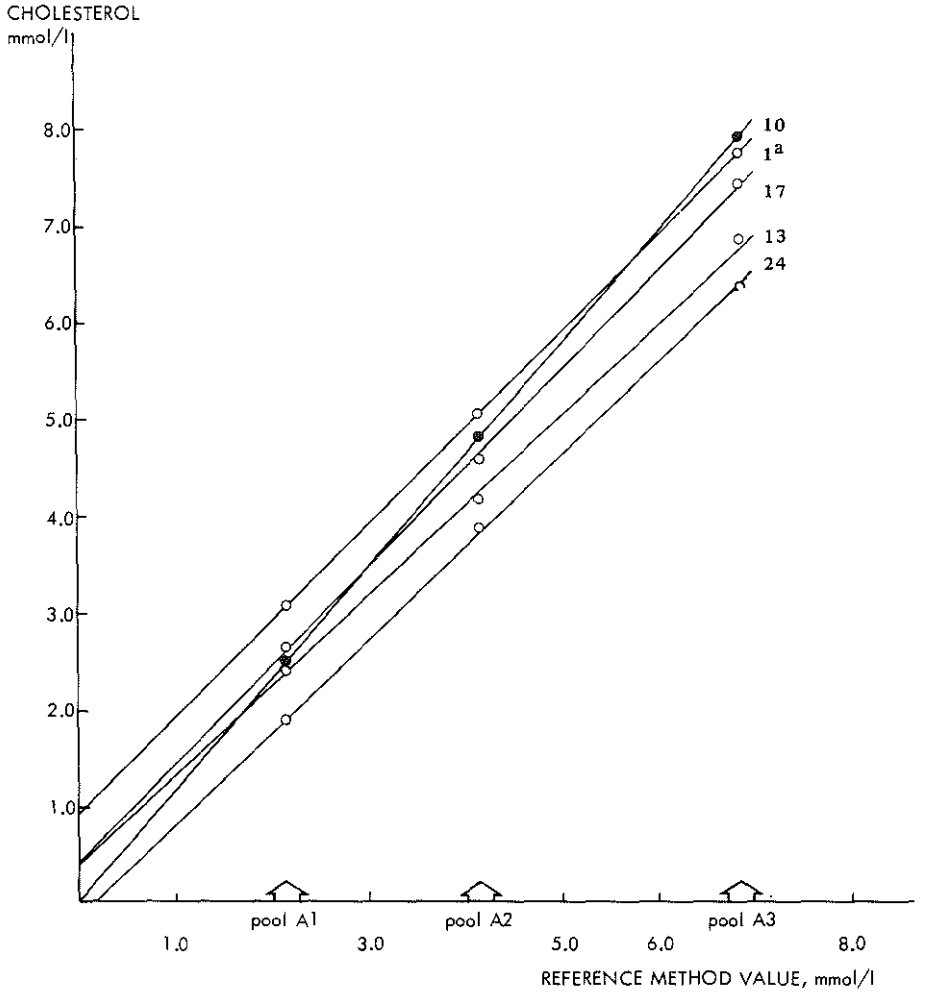
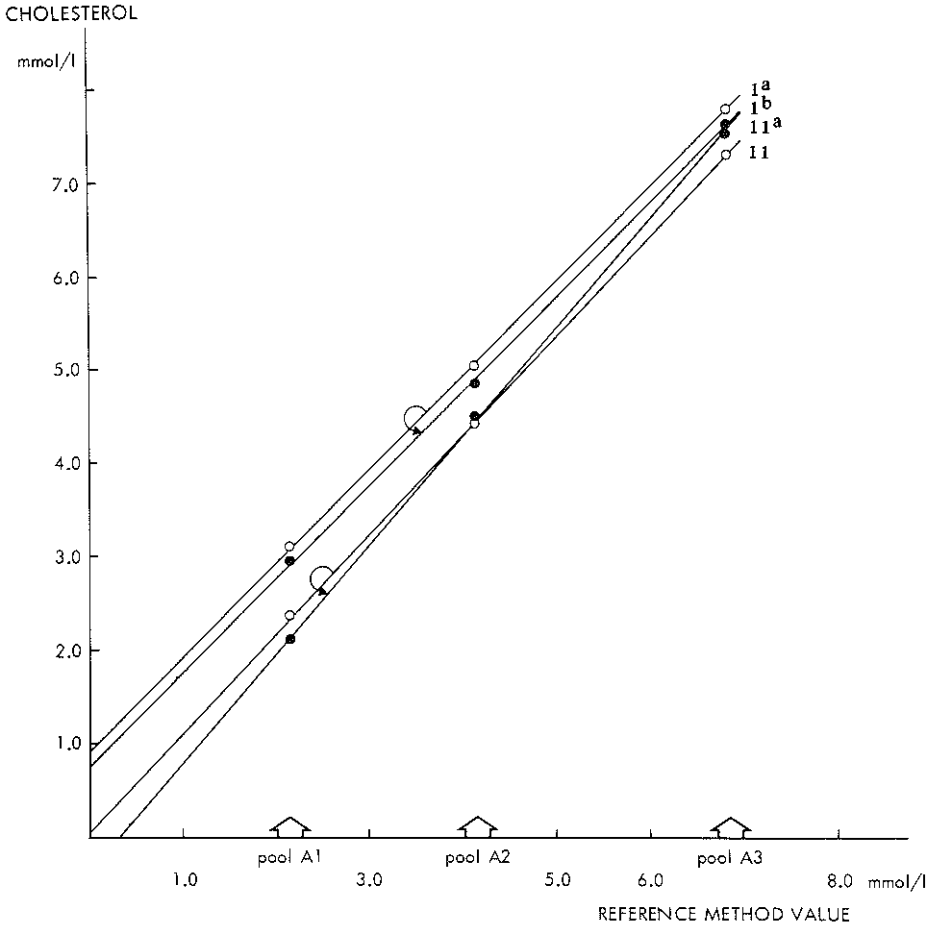


Figure 5.3. Intralaboratory shift in two KCA-laboratories (parallel and angular change). o and • show mean results with a few months time interval.



Every laboratory performed 30 analyses for each pool.

Table 5.4. Low and high values in KCA, part I (both LB-reaction) including an example of intralaboratory drift.

Pool	lowest value (mmol/l)	highest value (mmol/l)		drift	
		first series	second series [1)]	%	mmol/l
A_1	1.74	2.81	3.45	23	0.64
A_2	3.69	4.92	5.56	13	0.64
A_3	6.11	7.66	8.46	9	0.80

[1)]The high laboratory in the first series showed insufficient precision and repeated the measurements 4 months later. Significant upward drift had occurred.

5.4. Earlier experiments with common calibration procedures and methods

The main effort of the KCA was focussing on reduction of interlaboratory differences. A plan was developed to use serum calibrators to aim at reference method values at the same time. It was only reasonable of course to expect lower values in a direct LB-reaction when serum is used to calibrate. This would not, in itself, decrease the interlaboratory differences, unless the present calibration procedures would be among the main causes of them.

The equalization of calibration procedures had been attempted earlier in the Foundation for Quality Control in Clinical Chemical Hospital Laboratories. In three of the surveys a common standard solution was distributed along with two serum samples. Interlaboratory comparability did improve, but not very much, at the time (Table 5.5.).

An aqeous cholesterol solution, made by addition of a detergent (15), was centrally prepared and distributed along with the survey samples.

A second approach was tested among colleagues in the south-west region, near Rotterdam, during a number of national surveys. The RIV control serum was analyzed along with the survey samples. Its label values were used in calibrating 12 constituents.

Unfortunately the low cholesterol content (2.3 mmol/l) made the material unfit for calibrating this particular analyte. Typical results are shown in table 5.6.

Table 5.5. National QC survey with common cholesterol standard and/or common cholesterol method.

		Mean and s overall.[1] (mmol/l)	Same, but common standards.[2]	Common standards and common LB-method.[2]
Survey 9				
sample 9A.	$\bar{x}$.	4.1 (n = 97)	4.17 (n = 89)	4.25 (n = 77)
(Wellcome)	s	0.32	0.27	0.24
sample 9B	$\bar{x}$.	5.2	5.30	5.43
	s	0.35	0.33	0.26
Survey 10				
sample 10A	$\bar{x}$.	2.7 (n = 97)	2.86 (n = 80)	2.88 (n = 57)
(Nyegaard)	s	0.34	0.36	0.26
sample 10B	$\bar{x}$.	2.8	2.88	2.97
	s	0.32	0.26	0.27
			Common LB-method.[3]	Enzymatic methods only.[3]
Survey 11				
sample 11A	$\bar{x}$.	2.4 (n = 101)	2.33 (n = 71)	1.89 (n = 15)
(RIV)	s	0.38	0.27	0.21
sample 11B	$\bar{x}$.	2.4	2.40	2.04
	s	0.31	0.26	0.16

[1] Results from the National Surveys as sent out by the QC Foundation, Dr. A. P. Jansen, Nijmegen.

[2] Calculated from Youden-plots sent out with survey results.

[3] Two extreme values excluded (values not plotted).

Table 5.6. Cholesterol results with and without common serum calibrator (RIV-pool).

Survey	Sample	Mean ; corrected		s ; corrected	
13 (n = 11)	13A	3.99	3.72	0.32	0.44
(Dade serum)	13B	3.38	3.15	0.21	0.31
17 (n = 15)	17A	6.27	5.77	0.69	1.01
(Nyegaard)	17B	3.70	3.42	0.35	0.52
19 (n = 11)	19A	6.09	6.18	0.45	0.93
(Nyegaard)	19B	3.58	3.60	0.35	0.33

5.4.1. Conclusions

The experiments illustrate that taking just one common quality control serum or one common standard solution does not automatically change interlaboratory comparability very much. However, it must be noted that the surveys with common aqueous standards suffered from a few deficiencies:

1. The number of laboratories performing the LB-reaction in addition to their routine procedures was considerably smaller than the total.
2. Several laboratories not routinely employing the common LB-method only did so upon request for the purpose of the survey. In such cases the unexperienced, single attempt does not really reveal the merits or disadvantages of the method.
3. The aqueous standard itself was a novelty for the large majority of the participants.

The results obtained in the south-west region in survey 13 and 17 confirm that serum calibrators tend to lower the positive bias of direct LB-reactions. Survey 19 is an exception. Interlaboratory precision was not improved at all in this local recalibration enterprise. And it was confirmed at the same time that the different control pools behave differently in our methods.

5.5. KCA Programme part II:

It was expected that we produced a suitable serum calibrator and the KCA-material was used in a new trial.

The three pools used in part I of the standardization programme were designated to serve as standards. Each laboratory would employ its usual routine procedures and recalculate the same results later on the basis of measurements of serum A_1, A_2 and A_3 and the CDC-reference method values.

Five pooled, human sera, numbered I - V, were sent to a number of laboratories for analysis with the direct LB-method. A few of them also did enzymatic and Abell measurements. Six KCA serum calibrators were also shipped, numbered A_1 - A_6 (enriched horse serum). All samples were analyzed in duplicate in 4 different runs. Each laboratory prepared calibration curves by plotting absorbance values of serum A_1, A_2 and A_3 against the reference method values. Later, when assigned values for serum A_4, A_5 and A_6 had been obtained from CDC, these were also used in a retrospective, calculated calibration.

Tabel 5.7. demonstrates in what way the results are influenced by this common calibration method. The left section shows results, as obtained in the routine methods of the participants. The central part contains mean results and s values based on serum calibrators A_1, A_2 and A_3. On the right a recalculated result is obtained in such a way that each participant's mean result in serum A_4, A_5 and A_6 coincides with the reference method values mean.

As a first result serum calibrators clearly improve standard deviations of the Huang group.

To check whether the improvement of all standard deviations was significant a Morgan-Pitman-test was performed; table 5.8. The common F-test cannot be used here, because the two series of values each time are strongly correlated: the two series of results were taken from the same runs. Therefore a 'correction' is needed in order to find the proper value for t (16).

$$t = \frac{F - 1}{2\sqrt{F(1-r^2)}} \cdot \sqrt{n-2} \qquad \text{(t was tested one-sided)}$$

It must be noted that this test is applied to only 9 values which is a rather low number; therefore the enzymatic results were evaluated after a second trial.

The values for p are given for serum calibrators A_1, A_2 and A_3 and for A_4, A_5 and A_6. They are discussed in chapter 6 (table 6.4.). Two striking effects on the standard deviations are the improved homogeneity in the Huang group and the excellent result of calibrators A_1, A_2 and A_3 in the measurement of the other three preparations. This must be ascribed to the matrix similarity.

The second quantity, the mean result x, may be compared with the mean of Abell-measurements because these closely approach true values as can be seen in the case of A_4, A_5 and A_6. In the middle section of this table Huang values obviously have become much too low. In the right hand part they are only slightly lower (table 5.10).

We also observe how the average results of enzymatic methods closely approach reference method values in human serum. The prepared pools A_4, A_5 and

Table 5.7. Huang method (average of 4 analyses).

	Routine results								Results with serum calibrators A1, A2 and A3								Results with calibrators A4, A5 and A6				
Sera:	I	II	III	IV	V	A4	A5	A6	I	II	III	IV	V	A4	A5	A6	I	II	III	IV	V
LAB 1:	4.81	6.54	7.11	7.04	8.56	4.86	7.15	10.09	4.12	5.58	6.06	6.01	7.30	4.13	6.09	8.60	4.20	5.72	6.21	6.15	7.48
2:	4.77	6.22	6.99	6.89	8.53	4.79	7.02	9.77	4.12	5.45	6.11	6.13	7.58	4.19	6.17	8.62	4.26	5.56	6.24	6.15	7.62
3:	4.67	6.20	7.06	6.96	8.66	4.67	6.94	9.67	4.14	5.48	6.22	6.12	7.61	4.12	6.12	8.51	4.24	5.63	6.41	6.31	7.86
4:	4.85	6.39	7.17	7.11	8.84	4.94	7.05	9.88	4.01	5.37	6.06	6.01	7.54	4.09	5.95	8.46	4.27	5.62	6.31	6.25	7.78
5:	4.96	6.62	7.24	7.28	8.86	4.95	7.33	10.09	4.08	5.45	5.96	5.99	7.30	4.07	6.04	8.32	4.27	5.70	6.24	6.27	7.63
6:	4.42	5.77	6.38	6.28	7.83	4.67	6.79	9.33	3.88	5.12	5.64	5.54	6.93	4.10	6.00	8.26	4.09	5.34	5.90	5.81	7.24
7:	4.75	6.31	7.06	7.00	8.62	4.91	7.02	9.88	4.02	5.33	5.97	5.92	7.30	4.16	5.94	8.36	4.19	5.57	6.23	6.18	7.61
8:	4.80	6.23	6.90	6.89	8.48	4.70	6.80	9.47	4.29	5.63	6.25	6.25	7.73	4.19	6.16	8.67	4.41	5.72	6.33	6.32	7.78
9:	4.68	6.09	6.66	6.67	8.26	4.74	6.80	9.38	4.12	5.41	5.93	5.92	7.37	4.19	6.07	8.40	4.30	5.59	6.12	6.13	7.59
$\bar{x}$:	4.75	6.26	6.95	6.90	8.52	4.80	6.99	9.73	4.09	5.42	6.02	5.99	7.41	4.14	6.06	8.47	4.25	5.61	6.22	6.17	7.62
s:	0.150	0.250	0.273	0.287	0.315	0.115	0.181	0.287	0.112	0.148	0.180	0.199	0.239	0.047	0.085	0.143	0.087	0.117	0.146	0.154	0.185
Ref. lab. CDC:						4.24	6.28	8.79													
Enzymatic methods (average of 4 analyses).																					
10:	4.05	5.59	5.99	6.09	7.42	3.96	5.94	8.30	4.35	5.90	6.30	6.40	7.74	4.25	6.24	8.63	4.30	5.94	6.36	6.47	7.88
11:	4.36	5.63	6.33	6.15	8.09	3.76	5.90	8.23	4.74	6.06	6.80	6.62	8.16	4.10	6.34	8.79	4.74	6.12	6.88	6.68	8.79
12:	4.22	5.53	6.09	6.08	7.50	3.94	5.84	8.23	4.57	5.99	6.61	6.60	8.14	4.27	6.34	8.93	4.53	5.93	6.54	6.52	8.05
13:	4.33	5.76	6.40	6.38	7.91	3.93	5.94	8.40	4.65	6.17	6.84	6.82	8.44	4.24	6.35	8.96	4.59	6.11	6.79	6.77	8.39
$\tilde{x}$:	4.24	5.63	6.20	6.18	7.73	3.90	5.90	8.29	4.58	6.03	6.64	6.61	8.12	4.22	6.32	8.83	4.54	6.03	6.64	6.61	8.28
Abell-Kendall method:																					
14:	4.17	5.53	6.13	6.12	7.51	4.06	6.03	8.36	4.18	5.58	6.19	6.11	7.61	4.07	6.09	8.48	4.36	5.78	6.41	6.40	7.85
15:	4.32	5.98	6.66	6.60	8.14	4.42	6.56	9.14									4.14	5.74	6.39	6.33	7.81
16:	4.29	5.79	6.38	6.36	7.91	4.26	6.31	8.87	4.23	5.71	6.31	6.29	7.84	4.19	6.24	8.78	4.26	5.75	6.34	6.32	7.86
17:	4.40	5.75	6.31	6.40	8.05	4.28	6.30	8.81	4.32	5.59	6.18	6.14	7.76	4.19	6.02	8.47	4.38	5.72	6.28	6.37	8.01
$\bar{x}$:	4.30	5.76	6.37	6.37	7.90	4.26	6.30	8.80	4.24	5.63	6.23	6.18	7.74	4.15	6.13	8.58	4.29	5.75	6.36	6.36	7.88

Table 5.8. Calculated p values indicating significance of improvement of interlaboratory comparability (9 laboratories with direct LB-method).

Sample	calibration with A_1, A_2, A_3		recalculated with A_4, A_5, A_6	
	t	p	t	p
I	0.8998	0.200	1.8684	0.052
II	1.8408	0.054	3.9899	0.003
III	1.4324	0.097	3.1739	0.008
IV	1.3073	0.116	3.0998	0.009
V	0.8981	0.200	2.1953	0.032
A_4	2.9482	0.011		
A_5	2.2138	0.031		
A_6	1.8904	0.050		

A_6 demonstrate the often observed negative bias (14) in control materials. This negative shift in calibrators A_1, A_2 and A_3 leads to elevated results in the 'corrected' parts of the table, proving these sera to be unfit for calibration of enzymatic methods. Calibrators A_4, A_5 and A_6 do not change the mean Abell results, whereas in the central section of the table these are lowered.

A larger study was then carried out to see whether our results could be confirmed. Also the enzymatic group had to be enlarged. Five fresh human serum pools, numbered VI - X, and the KCA-sera were distributed to 25 laboratories. Some of them used more than one method.

The results, collected in tables 5.9.1. and 5.9.2. do indeed correspond with those from table 5.6.

Again standard deviations in the Huang-laboratories improved and the mean result is lower - in fact there is an overshoot, like the first time (table 5.10) (This 'Huang-group' is **not** group 'A' from table 5.3., but is mixed with 'B').

The Morgan-Pitman test reveals the largest improvement again in the standard deviations for pools A_4, A_5 and A_6 (matrix effect), whereas in the human pools significance is observed only when the second series of calibrators is used (table 5.11.). One-sided t was tested.

In the enzymatic methods a decrease of interlaboratory comparability occurred in some samples (negative t), but in the KCA pools the improvement is great also. The comparability among the original enzymatic results is better than it is in the LB-group; after recalibration this is reversed.

Table 5.9.1. Huang method (average of 4 analyses).

	Routine results								Results with serum calibrators A1, A2 and A3								Results with calibrators A4, A5 and A6				
Sera:	VI	VII	VIII	IX	X	A4	A5	A6	VI	VII	VIII	IX	X	A4	A5	A6	VI	VII	VIII	IX	X
LAB 1:	10.44	7.55	6.91	6.24	3.57	4.80	7.11	9.71	8.87	6.52	5.87	5.30	3.05	4.08	6.07	8.25	9.30	6.73	6.16	5.56	3.18
2:	10.34	7.81	6.99	6.39	3.88	4.95	7.08	9.68	8.85	6.64	5.90	5.36	3.19	4.13	6.00	8.26	9.14	6.90	6.18	5.65	3.43
3:	10.04	7.39	6.63	6.00	3.60	4.79	6.93	9.58	8.76	6.38	5.71	5.16	3.07	4.13	5.97	8.30	9.07	6.67	5.99	5.42	3.25
4:	10.05	7.50	6.80	6.03	3.73	4.77	6.90	9.43	9.14	6.77	6.11	5.40	3.26	4.22	6.15	8.56	9.15	6.83	6.19	5.49	3.40
5:	9.91	7.34	6.57	5.98	3.51	4.66	6.78	9.36	9.13	6.73	6.00	5.46	3.15	4.20	6.17	8.63	9.17	6.79	6.08	5.53	3.25
6:	9.52	6.89	6.25	5.47	3.21	4.66	6.67	9.20	8.39	6.12	5.52	4.85	2.86	4.13	5.90	8.12	8.91	6.45	5.85	5.12	3.00
7:	9.03	6.97	6.25	5.64	3.43	4.41	6.39	8.85	8.65	6.67	5.96	5.48	3.28	4.20	6.09	8.47	8.84	6.82	6.12	5.52	3.36
8:	9.84	7.25	6.50	5.86	3.43	4.46	6.58	9.14	8.96	6.66	5.94	5.38	3.20	4.11	6.03	8.03	9.40	6.93	6.21	5.60	3.28
9:	10.18	7.36	6.59	5.88	3.49	4.64	6.91	9.55	8.95	6.48	5.80	5.18	3.10	4.07	6.09	8.30	9.31	6.73	6.03	5.38	3.19
10:	9.33	6.95	6.25	5.75	3.42	4.60	6.65	9.27	8.44	6.29	5.66	5.20	3.09	4.17	6.01	8.40	8.75	6.52	5.86	5.39	3.21
11:	9.16	6.96	6.17	5.62	3.39	4.41	6.29	8.69	8.81	6.69	5.95	5.40	3.25	4.25	6.04	8.33	9.07	6.89	6.11	5.57	3.36
12:	9.77	7.30	6.51	5.84	3.47	4.59	6.76	9.28	8.75	6.50	5.80	5.20	3.10	4.09	6.02	8.29	9.12	6.81	6.08	5.45	3.24
13:	9.96	7.27	6.53	5.78	3.50	4.66	6.77	9.50	8.76	6.41	5.76	5.09	3.12	4.11	5.98	8.37	9.17	6.70	6.01	5.32	3.22
14:	10.91	7.94	7.27	6.47	3.92	5.12	7.45	10.23	8.97	6.46	5.89	5.22	3.06	4.08	6.04	8.38	9.20	6.70	6.13	5.46	3.31
15:	10.13	7.42	–	5.96	3.54	4.63	6.79	9.36	9.04	6.62	–	5.32	3.16	4.14	6.04	8.35	9.39	6.88	–	5.52	3.28
16:	9.31	6.80	6.11	5.44	3.21	4.32	6.35	8.85	8.88	6.50	5.84	5.20	3.08	4.14	6.07	8.45	9.20	6.72	6.04	5.37	3.17
17:	8.48	6.50	5.85	5.29	3.34	4.84	6.38	8.23	7.64	5.70	5.03	4.50	2.54	4.00	5.53	7.35	8.28	6.34	5.71	5.16	3.26
Ref. lab. CDC:						4.24	6.28	8.79													
$\bar{x}$:	9.80	7.25	6.51	5.86	3.51	4.66	6.75	9.83	8.76	6.48	5.80	5.21	3.09	4.13	6.01	8.30	9.09	6.73	6.05	5.44	3.26
s:	0.593	0.367	0.364	0.321	0.196	0.204	0.306	0.457	0.355	0.262	0.249	0.241	0.173	0.062	0.140	0.281	0.273	0.163	0.139	0.144	0.101

Table 5.9.2.

Enzymatic methods (average of 4 analyses)

	Routine results								Results with serum calibrators A1, A2 and A3								Results with calibrators A4, A5 and A6				
Sera:	VI	VII	VIII	IX	X	A4	A5	A6	VI	VII	VIII	IX	X	A4	A5	A6	VI	VII	VIII	IX	X
LAB 18:	9.06	6.94	6.54	5.79	3.67	3.63	5.48	7.57	10.18	7.83	7.39	6.55	4.20	4.22	6.31	8.66	10.50	8.04	7.58	6.71	4.25
19:	9.08	6.78	6.05	5.34	3.29	3.90	5.76	8.02	9.80	7.45	6.69	5.90	3.72	4.36	6.35	8.78	9.91	7.40	6.60	5.83	3.59
20:	9.40	6.77	6.01	5.43	3.34	3.88	5.86	8.27	10.20	7.35	6.53	5.88	3.63	4.21	6.36	8.99	10.11	7.28	6.47	5.84	3.59
21:	10.10	7.45	6.42	5.55	3.44	4.25	6.36	9.00	9.44	7.06	6.14	5.36	3.47	4.20	6.09	8.46	9.97	7.35	6.34	5.48	3.40
22:	8.94	6.66	5.93	5.32	3.25	3.85	5.75	8.00	9.73	7.29	6.49	5.83	3.61	4.25	6.30	8.73	9.81	7.31	6.51	5.84	3.57
23:	8.92	6.57	5.92	5.20	3.20	3.84	5.75	8.06	9.65	7.15	6.45	5.67	3.54	4.21	6.27	8.74	9.77	7.20	6.49	5.70	3.51
24:	9.51	6.93	6.28	5.53	3.34	4.01	5.86	8.28	10.06	7.39	6.68	5.88	3.55	4.26	6.20	8.81	10.11	7.37	6.68	5.88	3.55
$\bar{x}$:	9.29	6.87	6.16	5.45	3.36	3.91	5.83	8.17	9.87	7.36	6.62	5.88	3.67	4.24	6.27	8.74	10.03	7.42	6.67	5.90	3.64
s:	0.423	0.288	0.249	0.193	0.156	0.187	0.265	0.435	0.288	0.248	0.384	0.357	0.245	0.056	0.095	0.160	0.248	0.281	0.416	0.384	0.278
Abell-Kendall methods																					
LAB 25:	9.54	6.99	6.27	5.54	3.32	4.30	6.43	8.92	9.17	6.72	6.03	5.34	3.21	4.14	6.21	8.56	9.38	7.43	6.67	5.89	3.53
26:	9.33	6.94	6.24	5.41	3.33	–	6.15	8.63	9.33	6.87	6.17	5.33	3.21	–	6.07	8.60	9.15	6.81	6.12	5.31	3.27
27:	9.48	6.89	6.14	5.39	3.28	4.22	6.31	8.64	9.43	6.84	6.09	5.34	3.24	4.17	6.26	8.58	9.53	6.93	6.18	5.42	3.30
$\bar{x}$:	9.45	6.94	6.22	5.45	3.31	4.26	6.30	8.73	9.31	6.81	6.10	5.34	3.22	4.15	6.18	8.58	9.35	7.06	6.32	5.54	3.37
Ref. lab. CDC:						4.24	6.28	8.79													

Table 5.10. Mean results compared with Abell values and presented as percentage bias.

	Routine procedure		Calibration with A_1, A_2, A_3		Calibration with A_4, A_5, A_6
Huang method	Human serum	KCA serum	Human serum[1]	KCA serum	Human serum
1st survey (n = 9)	+ 8.9%	+ 11.4%	- 5.7%	- 5.6%	- 2.7%
2nd survey (n = 17)	+ 5.3%	+ 9.7%	- 6.4%	- 4.2%	- 2.3%
Enzymatic methods			[2]	[3]	
1st survey (n = 4)	- 2.3%	- 6.9%	+ 4.4%	- 0.1%	+ 4.6%
2nd survey (n = 7)	- 0.4%	- 7.4%	+ 7.1%	- 0.3%	+ 7.7%

[1] Calibrating the Huang method with A_1, A_2 and A_3 creates considerable negative bias; A_4, A_5 and A_6 perform much better.
[2] Calibrating the enzymatic methods with either set of serum calibrators creates a positive bias, whereas the original averages were quite accurate.
[3] Note that the similarity of matrices yields excellent results in KCA serum in the enzymatic procedures.

Table 5.11. Calculated p values indicating significance of change in interlaboratory comparability.

	Huang (n = 17)			Enzymatic (n = 7)	
Sample	t	p	sample	t[1]	p
VI	2.9293	0.005	VI	0.9110	0.202
VII	1.5991	0.065	VII	0.3387	0.374
VIII (n = 16)	1.7439	0.052	VIII	- 1.1153	0.157
IX	1.3134	0.104	IX	- 1.7680	0.068
X	0.4985	0.312	X	- 1.6220	0.082
A_4	6.4180	< 0.001	A_4	3.4708	0.008
A_5	3.4762	0.002	A_5	4.2692	0.004
A_6	2.2603	0.020	A_6	2.8283	0.018
VI	5.0172	< 0.001	VI	1.2632	0.131
VII	4.1561	< 0.001	VII	0.0586	0.478
VIII (n = 16)	5.5129	< 0.001	VIII	- 1.5177	0.095
IX	4.4857	< 0.001	IX	- 2.2385	0.038
X	3.5407	0.001	X	- 2.2156	0.039

[1] with negative t a decline of comparability is indicated.

Considering our initial (routine) Abell values for A_4, A_5 and A_6 the serum calibrators A_1, A_2 and A_3 decrease those values unfavourably which might suggest that the assigned calibrator values are on the low side. The downward shift was also observed in the first experiment (table 5.7.) and at present the serum calibrators A_1, A_2 and A_3 are again investigated in the CDC Lipid Reference Laboratory.

5.6. Conclusions

1. The KCA-samples were analyzed in the Reference Laboratory at CDC (17). The reference method values are:

   A_1 : 2.14 mmol/l s = 0.054 n = 32
   A_2 : 4.16 mmol/l s = 0.088 n = 31
   A_3 : 6.85 mmol/l s = 0.101 n = 32
   A_4 : 4.24 mmol/l s = 0.051 n = 16
   A_5 : 6.28 mmol/l s = 0.083 n = 16
   A_6 : 8.79 mmol/l s = 0.081 n = 16

   In our Abell-Kendall procedures we find that the results of pool A_4, A_5 and A_6 are lowered when we calibrate with A_1, A_2 and A_3. This has led to a re-investigation of the target values.
2. The LB-reaction in these surveys on the average overestimates cholesterol by 8.9% to 5.3% respectively. After the first correction the mean results become 5.7% to 6.4% too low. The KCA-pools (A_4, A_5 and A_6) at the same time have a larger bias and show a better correction. The serum calibrators lead to an overshoot effect, unless their matrix is similar to that of the analyzed material. Calibration with the three new serum pools A_4 - A_6 brings Huang means within WHO limits. Bias in human sera must still be thoroughly checked.
3. The serum calibrators in enzymatic analysis tend to turn accurate results into high results. Only the KCA-pools (A_4, A_5 and A_6) do improve considerably. As in the LB-procedure, it is quite obvious that enriched pools behave in a way that differs from that of human serum. In the enzymatic analysis no general change of calibration is necessary. The remaining problem is: how to prepare in sufficient quantity a serum control, fit to check the enzymatic accuracy in the KCA-programme.
4. Although all laboratories in part I produced acceptable precision, variations among laboratories employing the same methodology are large. In addition the intralaboratory reproducibility is not stable proving that sustaining good precision is a problem. It is in fact the problem that must be solved before all other ones.
5. It appears that serum calibrators of a type like A_1, A_2 and A_3 are able to pro-

duce more comparable Huang values in pool A_4, A_5 and A_6, but are less adequate for normal human serum. Calibration with A_4, A_5 and A_6 yields better results. Their application will mean a step forward on the road to interlaboratory comparability. Before this step is taken consultation with the laboratories and the several projects must take place.

In chapter 6 these findings and some very recent developments are discussed.

REFERENCES

Chapter 5

1. IFCC. Provisional recommendation on quality control in clinical chemistry. Part 6. Quality requirements from the point of view of health care. Clin. Chem. *23*, 1066 - 1071 (1977).
2. W. P. BELK and F. W. SUNDERMAN: A survey of the accuracy of chemical analyses in clinical laboratories. Am. J. Clin. Pathol. *17*, 835 - 861 (1947).
3. I. D. P. WOOTTON: International biochemical trial 1954. Clin. Chem. *2*, 296 - 301 (1956).
4. A. H. HOLTZ: Klinisch-chemisch enquête-onderzoek in Nederland. Ned. T. Geneesk. *103*, 2366 - 2376 (1959).
5. A. H. HOLTZ: Methoden voor klinisch-chemische analyses 1957 - 1964. Chem. Weekblad, *61*, 424 - 427 (1965).
6. N. S. LAWSON and G. T. HAVEN: The role of regional control programmes in the practice of laboratory medicine in the United States. Am. J. Clin. Pathol. *66*, 268 - 275 (1976).
7. A. P. JANSEN, E. J. VAN KAMPEN, B. LEIJNSE, C. A. M. MEIJERS and P. J. J. VAN MUNSTER: Experience in the Netherlands with an external quality control and scoring system for clinical chemistry laboratories. Clin. Chim. Acta, *74*, 191 - 201 (1977).
8. G. J. M. BOERMA and B. LEIJNSE: Standaardisatie van de analyse van serum cholesterol in het kader van de preventie van hart- en vaatziekten in Nederland: een mededeling van de Stichting KCA. Mededelingen NVKC, *2*, 130 - 132 (1977).
9. G. R. COOPER: The WHO-CDC Lipid Standardization Program, in: Quality Control in Clinical Chemistry. Transactions of the VIth International Symposium. G. Anido, E. J. van Kampen, S. B. Rosalki, M. Rubin eds. W. de Gruyter, New York (1975).
10. J. H. WILLIAMS, L. TAYLOR, M. KUCHMAK and R. F. WITTER: Preparation of hypercholesteremic and/or hypertriglyceridemic sera for lipid

determination. Clin. Chim. Acta *28*, 247 - 253 (1970).
11. P. N. M. DEMACKER, H. E. VOS-JANSEN, A. P. JANSEN and A. VAN 't LAAR: Evaluation of the dual precipitation method by comparison with the ultracentrifugation method for measurement of lipoproteins in serum. Clin. Chem. *23*, 1238 - 1244 (1977).
12. J. C. KOEDAM: Comments on the conversion of cholesterol values obtained using different methods. Clin. Chim. Acta *79*, 607 - 608 (1977).
13. A. M. DE BRUYN, F. KLEIN, C. M. VAN GENT and H. A. VAN DER VOORT: On the calibration of automated enzymatic cholesterol determinations. Clin. Chim. Acta *85*, 215 - 218 (1978).
14. This phenomenon can also be noted when studying stated values, on the package inserts of serum control manufacturers, that have been obtained with various methods.
15. L. VASS: Wässerige primäre Cholesterinstandardlösung mit Cholesteringehalt im oberen pathologischen Bereich. Zubereitung und Bedeutung. Ärztl. Lab. *5*, 148 - 151 (1974).
16. a. E. J. G. PITMAN: A note on normal correlation. Biometrika *31*, 9 -12 (1939).
    b. W. A. MORGAN: A test for a significance of the difference between the two variations in a sample for a normal bivariant population. Biometrika *31*, 13 - 19 (1939).

'Nos vetera instauramus, nova non prodimus'
(Our business is to restore the old, not further the new)

Erasmus

That may have been Erasmus' business, but ours is the reverse. In this fast-moving field the old is soon replaced.

Eds. of Clin. Chem. (1977) (1)

CHAPTER 6

DISCUSSION OF QUALITY CONTROL, STANDARDIZATION AND THE USEFULNESS OF SERUM CHOLESTEROL DETERMINATION

6.1. Introduction

As an illustration of what this thesis is all about let us look into the manner in which a new analytical method, or an old method with new diagnostic potential, becomes a part of general daily practice. HDL-cholesterol introduced recently as a risk indicator will be our example. It is not at all a newly discovered serum component but it was never analyzed as frequently as other lipoproteins since it is not used in the Fredrickson typing of lipidemias (2). However, recent publications have boosted the interest of many scientists and HDL has become known as an important inverse risk indicator (3, 4, 5). Although the chemical composition is complex and several distinct classes exist among the HDL-lipoproteins (HDL-2 and HDL-3 being most important) in the present situation HDL-cholesterol is often measured without further differentiation. Electrophoresis is not precise enough for good, quantitative analysis (6, 7). The ultracentrifuge is expensive and time consuming, although a miniature instrument is now available (8). A number of procedures have been published in which the other lipoproteins are being precipitated (LDL and VLDL together or in separated steps) and HDL remains in the supernatant after centrifugation (9). The HDL is then quantitated by means of its cholesterol content. The apoproteins A1 or phospholipids may also be analyzed (10).

High levels of HDL are protective against atherosclerotic complications because HDL facilitates cholesterol transport away from blood vessel walls towards the liver where excretion must take place - thus reads the hypothesis (11, 12).

A clinical chemist considering the possible introduction of this analysis - his

enthousiasm springing from the literature data or from being hurried by his fellow clinicians - will begin to screen published analytical methods, finds one or another promising technique and begins to evaluate HDL measurements. The selection of the technique is based upon workload, available instrumentation, previous experience, personal preferences and such. This results in a variety of methods. There is no standard material or calibration serum readily available (although it does exist) (13). A means for a clinical chemist to check his analyses somehow, is to see whether values in healthy people as obtained by his laboratory, agree more or less with published reference values. Soon the first results begin to flow towards the requesting physician. This is a realistic account, as we proved in the following way: In order to obtain insight into the methodology and precision of performance of HDL-cholesterol determinations two trials were organized. Serum samples were sent by mail to sixteen laboratories, actively engaged in HDL-research of some kind. The majority of them recently became interested in HDL-measurements. A questionnaire about methods and reference values was included. To study the effect upon comparability of the HDL-isolation separately from that of the cholesterol measurement a few HDL-fractions, prepared by tungstate precipitation of LDL and VLDL, were distributed at the same time for cholesterol analysis only. All results are collected in table 6.1 and 6.2.

Conclusions from the questionnaire are:

a. At least 8 different modifications of HDL-cholesterol procedures are in use.
b. Precipitation techniques include Ca-heparine, Mg-heparine, Ca-dextrane sulfate, Mn-dextrane sulfate, Na-tungstate in various concentrations.
   The cholesterol method includes Abell-Kendall, Huang and enzymatic procedures in several modifications.

Conclusions from the tables are:

c. The total CV values range from 16% to 48% and in human serum from 16% to 30%. And s lies between 0.2 and 0.3 mmol/l.
d. The cholesterol analysis alone yields s values of 0.1 - 0.3 mmol/l - CV: 13% to 20% - and is the main cause of overall variability.
e. An inserted duplicate sample reveals that intralaboratory repeatability as a rule is satisfactory (table 6.1. serum 2 and 5).
f. After a history of 30 years of widely published interlaboratory trial results with poor comparability, clinical chemists continue to introduce analytical procedures without coordination. Even when the measurement relates to an important risk factor in cardiovascular epidemiology, comparability does not have the required priority.

Table 6.1. First HDL-survey, August 1978.

LAB	Precipitation method	Cholesterol method	Serum pools[1] 1	2	3	4	5	6	7	HDL-fractions[2] I	II	III
1	Na-tungstate	Abell	1.47	1.28	4.64	0.99	1.19	–[3]	0.74	1.79	1.23	0.79
2	Mn-heparine	ChOD-PAP	1.88	–	2.54	1.27	1.31	–	1.11	1.71	1.21	–
3	Mg-dextr.sulfate	ChOD-PAP	1.52	–	4.95	1.01	1.12	–	0.98	1.71	1.21	–
4	Mn-heparine	ChOD-PAP	2.43	1.26	1.95	1.29	1.32	1.20	1.18	1.59	1.21	0.68
5	Na-tungstate	ChOD-PAP	0.76	–	2.31	0.90	1.08	0.68	0.85	1.51	–	0.70
6	Ca-dextr.sulfate	ChOD-Katalase	–	0.66	3.05	0.44	0.73	0.62	0.44	1.67	1.24	0.73
7	Na-tungstate	Abell	1.29	1.11	1.09	0.85	1.11	0.72	0.78	1.78	1.40	0.75
8	Na-tungstate	ChOD Pierce-Kit	1.37	1.11	1.11	0.88	1.11	0.78	0.75	1.78	1.45	0.83
9	Mn-heparine	ChOD-Katalase + EDTA	1.87	1.18	2.84	1.04	1.19	–	0.94	1.63	1.30	–
10	Mn-heparine	Abell	1.7	1.0	3.3	1.1	1.1	–	0.9	1.6	1.2	0.7
11	Na-tungstate	ChOD Pierce-Kit	0.99	0.58	–	0.67	0.63	–	–	0.77	0.61	0.37
12	Mg-tungstate	ChOD-PAP	1.0	0.8	2.3	0.6	0.7	0.5	0.6	1.5	1.1	0.6
13	Na-tungstate	ChOD-Katalase	1.35	1.09	–	1.27	1.11	–	1.40	1.89	1.63	0.88
14	Mn-heparine	Huang with calibration serum	1.52	1.06	1.21	0.94	1.09	0.64	1.59	1.73	1.35	0.73
n			13	11	12	14	14	7	13	14	13	11
$\bar{x}$			1.47	1.01	2.61	0.95	1.06	0.73	0.94	1.62	1.24	0.71
s			0.44	0.23	1.26	0.25	0.21	0.22	0.32	0.27	0.23	0.13
VC %			29.9	23.1	48.3	26.8	20.4	30.4	33.5	16.6	18.9	19.1

With lab. 11 deleted from HDL results (fractions I, II and III) the following results are obtained:

	I	II	III
n	13	12	10
$\bar{x}$	1.68	1.29	0.74
s	0.11	0.14	0.08
VC %	6.8	11.1	10.8

Table 6.1. (continued).

[1]The pools consisted of the following materials: 1, Monitrol II (Dade); 2 = 5, human pool; 3, Elevated Lipid Control (Lederle); 4, Monitrol I (Dade); 6, human pool and 7, human pool with bilirubin 110 μmol/l.
[2]Prepared by Na-tungstate precipitation of LDL and VLDL.
[3]Several vials were damaged in the mail, results not reported.

Table 6.2. Second HDL-survey, November 1978.

LAB[1]	Precipitation method[2]	Cholesterol method	Serum pools[3] 8	9	10	11	12	HDL-fractions IV	V	VI
1			1.25	1.30	1.21	1.28	1.41	1.15	0.87	0.99
2			1.58	1.40	1.34	–	1.30	0.96	0.87	0.99
4			1.27	1.35	1.23	1.22	1.39	1.10	0.88	1.00
6			0.58	0.80	0.62	1.06	1.16	–	0.87	1.04
7			1.49	1.33	1.22	1.30	1.35	1.32	1.00	1.11
9			1.24	1.29	1.12	1.16	1.25	1.24	0.97	1.06
10			1.39	1.48	1.32	–	1.77	1.15	0.94	1.00
11	Na-tungstate	ChOD	1.49	1.53	1.46	1.65	–	1.03	0.71	0.87
12			1.2	1.2	1.1	1.3	1.2	1.20	0.94	1.01
13			1.56	1.53	1.45	1.53	1.76	1.53	1.22	1.32
14			1.24	1.33	1.26	1.14	1.52	1.25	0.97	1.09
15	Mg-dextr.sulfate	ChOD	1.07	1.13	0.99	1.05	1.14	1.04	0.73	0.82
16	Mn-heparine	Huang	1.62	1.67	–	1.60	1.68	–	1.20	1.30
n			13	13	12	11	12	11	13	13
$\bar{x}$			1.31	1.33	1.19	1.30	1.41	1.18	0.91	1.03
s			0.27	0.21	0.22	0.20	0.22	0.15	0.18	0.16
VC %			21.1	16.3	19.0	16.1	16.0	13.4	20.7	16.2

Notes
[1]Same as in table 6.1.; three did not report values; two new participants added.
[2]Same as in table 6.1., unless stated here.
[3]Human pools

g. Reference materials and standardization are urgently needed in lipoprotein analysis.

In the following sections (6.2. - 6.4.) a number of steps will be discussed that have to be taken to achieve a quality of cholesterol analysis wanted in epidemiological investigations. Section 6.7.7. will discuss the different needs for different purposes.

6.2. Terminology

Communication of meaningful data is possible only when members of a scientific community practice some linguistic normalization. Clinical chemistry is a field in which misunderstanding as a result of the use of personal interpretation of terms occurs easily. Fortunately in many countries professional or legal authorities have compiled norms for terminology. However, these do not always coincide in a convenient way. *Accuracy* for instance in the IFCC definition relates to systematic error in the best estimate (14), whereas the ASTM relates it to single measurements and thus includes the precision concept in it (15).

The NNI defines both: firstly: accuracy in the IFCC-sense (Dutch: juistheid, zuiverheid) and secondly: exactitude (nauwkeurigheid) which encompasses both systematic and random errors (16). It is a pity that 'nauwkeurigheid' is often used in this country for precision. Two methods with the same 'exactitude' may well differ in both accuracy and precision and the expression does not generally provide useful information.

As for *precision* it is always necessary to state clearly whether repeatability is under discussion or reproducibility, and ideally, whether one or more instruments or technicans are involved and which time intervals have occurred.

The concept of *error* itself is not a simple one either. Sometimes the distinction between random or systematic is hard to establish - causes for systematic error may be variable entities themselves and random errors may be caused by unrecognized phenomena. Sometimes 'bias' or 'analytical bias' have been used for precision (17) or prejudice (18) which of course must lead to confusion.

Materials, like primary standards or secundary standards or reference sera have been defined by a number of committees in different ways and most of this is in the stage of draft proposals. The way to uniformity seems to be long. Even the simple information about how *serum or plasma* was obtained, or which of the two was used, is deleted in several papers (19). The consensus acceptance of definitions of *methods* is probably not more at hand and among the standard methods, selected methods, reference methods, optimized methods and definitive methods a

lot of clearing up has to occur. Particularly the latter term: definitive, is an example where the linguistic and philosophical arguments frequently obscure the real issue of methodology. It sometimes seems as if some are afraid of the process of approaching true values, which inevitably teaches us the limitations of long used and trusted procedures (compare section 2.4.).

Quality control in Dutch means 'kwaliteitsbeheersing' or 'kwaliteitsbewaking', but until recently the term 'kwaliteitscontrole' (quality checking) was in use predominantly. This term indeed is less compulsive and may seem to require less engagement. Control systems imply preventive measures whereas checking systems are more aimed at inspections after the event. In accordance with the modern approach, involved with earliest identification of sources of error, the term 'quality assurance' is also used (20). The latest version of various, officially proposed terminology was very recently published in the proceedings of a conference (21). In the field of epidemiology confusion may also arise regarding terminology. When several kinds of investigations are described, a clear distinction ought to be made between 'epidemiological surveys', 'screening', 'case-finding' and 'surveillance' (22). There are different requirements for test properties in these cases (23); see section 6.7.7.

6.3. Organization of quality control

This is very much a matter of delegation and sharing of responsibilities. Of course general principles of care, paying attention to neatness and maintenance of equipment, sufficient know-how and a good working atmosphere are prerequisites for quality. In case an analytical procedure lacks precision and accuracy a large variety of factors may be causative and each should be checked.
Summarising such factors one will think of:

Glassware	: pipettes, volumetric flasks and such must be of sufficient quality and well cleaned.
Weighing devices	: must be well calibrated.
Reagents	: all chemicals used must be of analytical grade, (p.a.) purity, including the water;
Standards	: several standards are available from the NBS, Washington DC (standard reference materials = SRM) including cholesterol being more than 99.4% pure. But cholesterol of comparable purity is now commercially available also. All chemicals and standards must be properly stored - some cold, some at room-temperature, some in the dark, and shelf life must be carefully observed.

Photometers	: Wavelength and absorbance must be calibrated from time to time.
Specimens for analysis	: the serum or plasma samples (make sure which is which always) should be as fresh as possible; preferably without hemolysis, turbidity, lipaemia or drugs. Storage in the cold (4°C) in well closed tubes is possible over several days. At room temperature the ratio free: esterified cholesterol may change.
Quality control materials	: proper quality control must include at least two levels of control serum and clearly outlined rules for accepting and discarding an analytical run.
Analytical procedure	: all steps must be performed in as much the same manner each time as possible (pipetting, mixing, timing, temperature). For this a written and argued account of every step is needed, because oral transmissions from technician to technician will not do justice to all necessary details.

(In reference (24) an exhaustive checklist is given, as employed by CDC Laboratory Inspectors in the USA).

6.3.1. Our recommended organization scheme in a simple form is the following:

I *Process control* must be executed entirely by the technicians. The analytical procedure be performed exactly according to the written protocol. It is controlled at the bench and produces analytical results within predetermined control limits. It is the responsibility of the technician to reject results always when a run is out of control. For this purpose known serum control samples are used. Details were described in chapter 3 and 4. The system ensures a sufficiently constant precision and no seizable drift will be accepted.

II *Retrospective quality control* must evaluate the proficiency attained. Unknown serum controls treated as routine patient samples are introduced into all runs. The results must become available in time to confirm that process control mechanisms are functioning to a satisfactory degree. This must lead to changes in the method or the materials used whenever results shift out of control limits. The modern tendency to board out this retrospective control to serum manufacturers can threaten the readiness of action of the clinical chemist. The director may delegate this task and considerable responsibility to a 'control officer' who is not involved in routine measurements a great deal and who will be able to spot unexpected errors timely.

As a result of I and II the precision of the method is known and absence of

drift is confirmed. Accuracy is pursued by making specificity a principle point in method selection.

III Interlaboratory comparison; the laboratory director must set up an *external quality control* system. By inserting a proper reference serum into several runs the bias of a method will become established.

It is also desirable to know the position of the own method among many others. Participation in one or more interlaboratory trials makes this position surveyable. It is important to decide what is to be done with survey results of this kind before a selection out of the many programmes is made. In general, comparability within a limited area is more important than comparability among the continents, at least so in the clinical situation.

6.4. Materials for quality control

A great number of quality control serum preparations are available commercially in lyophilized form. Mostly bovine or equine serum constitutes the main fraction of such pools. A limited number of them has an added lipid content that can provide the high cholesterol level of man and the matching cholesterol ester variety. Human serum is used sometimes in which case hepatitis-B antigen tests must be negative. With adequate dispensing and lyophilizing procedures these sera show, after reconstitution, a sufficiently homogeneous vial-to-vial composition (VC $< ½\%$) (25). Inhomogeneity is difficult to ascertain but is sometimes revealed in large interlaboratory trials (26). Usually lyophilized pools are fit to be used in reproducibility control as stability can be guaranteed for a period of at least two years. Most laboratories can make human serum pools with or without added lipoproteins; several methods are available in the literature for this purpose (27).

When different kinds of serum are in use a serious problem arises regarding the estimation of analytical bias. The biological matrix in which the analyte is present influences the results by a number of factors. An analytical result is the sum of them all. For instance we will get:

Result = total cholesterol concentration + a + b + c + d + e.

where a = error introduced by cholesterol esters
b = error introduced by bilirubin
c = error introduced by other components (drugs)
d = error introduced by physio-chemical properties, like high viscosity
e = analytical errors (from calibration or instrumental problems etc.).

Evidently most of these contributions will be different from serum to serum, in the case of patient sera as well as the various lyophilized preparations on the market.

We have seen in the KCA trials that even serum with most of the lipid content

being from human sources gives rise to unexpected behaviour in several analytical methods.

6.4.1. Materials in the KRIS and CB studies

In the KRIS-study accuracy was checked by CDC with a number of human serum pools (fig. 3.1.) whereas in the CB Heart Project serum pools of varying origin were employed by the Prague Reference Center (fig.4.2.). In 1973 already, this material dependent behaviour became apparent. It was attempted to make corrections each time new pools showed a bias distinctly different from previous ones. As a result the graph in fig.4.2. is suggestive of a fairly stable laboratory performance; on the average a small positive bias is present. The survey means are within or just above WHO narrow limits. However, when uncorrected values are displayed, often a much higher bias appears to be present (fig. 6.1.).

Pools B and C were used from 1973 till 1975 in five comparisons (surveys). In late 1975 new pools -L and M- were introduced, when the new combined cholesterol-triglycerides programme started. A higher bias than before was noticed at that instant. In December 1975, at a visit to Rotterdam, Dr. Grafnetter provided us with several samples of the pools B, C, L and M. Comparisons were made during several days and the shift of the bias could thus be attributed to the new materials with reliability. Our target values for pools L and M were increased by the Reference Laboratory accordingly (+8%), table 6.3. The graphic representation in figure 4.2. continues to demonstrate a low bias during 1976.

The following pools, N and P, gave us some problems in triglycerides analysis and were soon replaced by T and U.

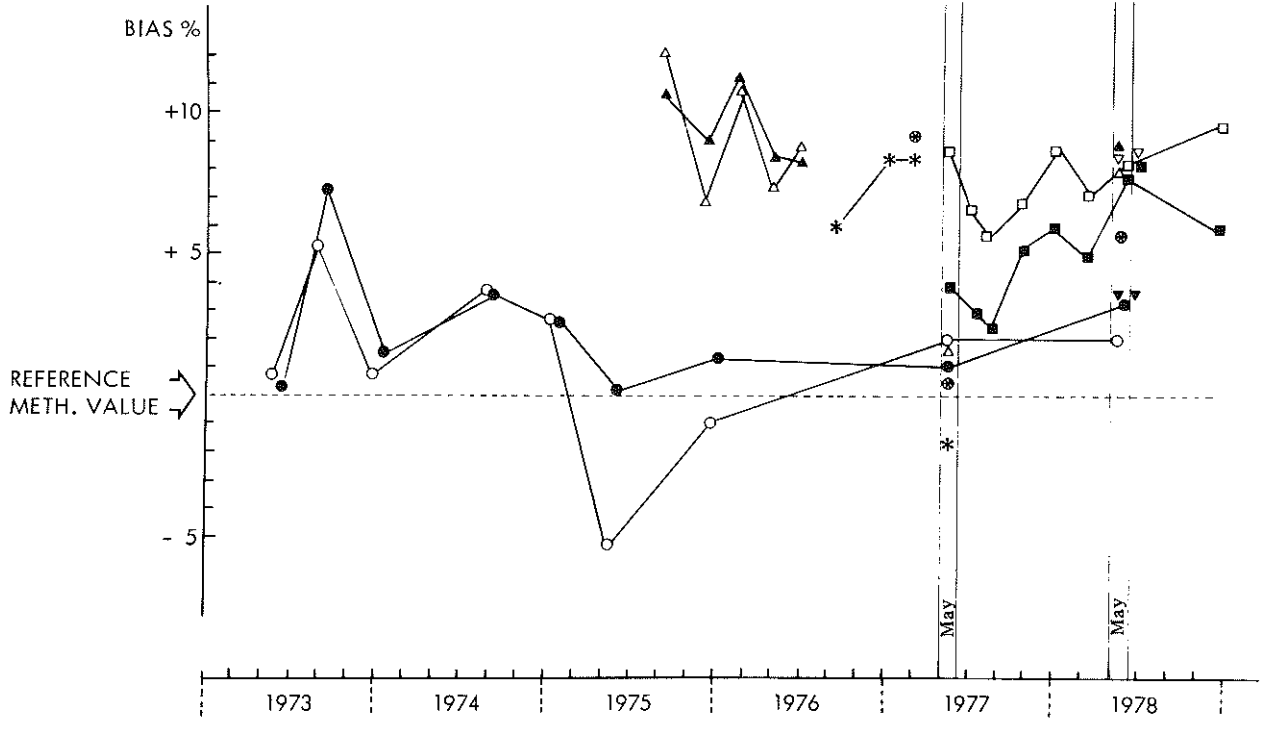
Again their matrix effect was studied, however in a limited series of measurements while at the same time no reference method values had yet been established. These were provisionally estimated and a correction of +10% and +6% was applied to neutralize our shifted bias. It later turned out that the target values were a little lower than the previously estimated ones and our bias correction was overdone. In the next period our results often seemed to have a negative error.

Nevertheless an exchange of human serum (Autumn 1977) and comparisons during 1978 and 1979 with the Abell-Kendall method of our own Reference Laboratory indicate that our bias indeed is small (positive and sometimes zero) and quite comparable with the results appearing in figure 4.2.

During 1978 some new pools again were put into use but comparisons are not yet completed.

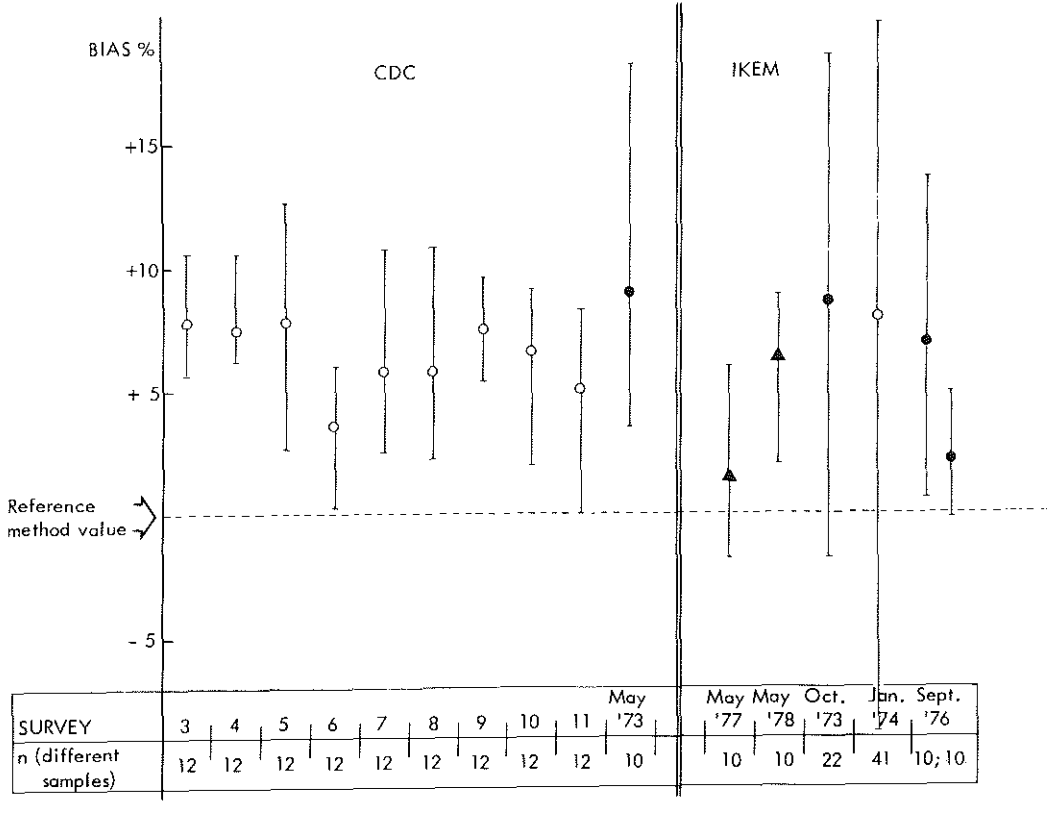
The importance of making a clear distinction between bias and random error is seen when survey results of single analyses are studied. The KRIS-surveys - 12 hu-

Figure 6.1. Material related bias in reference serum, IKEM, Prague. Matrix effects not corrected.



pools: B (o), C (●), L (Δ), M (▲), N (*), P (⊛), T (□), AA (▽), AB (▼).
In May 1977 and 1978 various pools were analyzed, but with few measurements per pool which makes precision of great importance.

Figure 6.2. Mean and range of serum analyzed once in mixed survey sets. Human (CDC) and lyophilized (IKEM) serum.



Patient serum from Rotterdam (●), lyophilized serum (▲) and liquid material (o).

Table 6.3. Comparison: December 1975. Bias in WHO Prague Lipid Reference Center Surveys.

Pool	Reference method	n	$\bar{x}$	bias	corrected bias
B	4.71	14	4.66	- 1.1%	- 1.1%
C	7.53	13	7.63	+ 1.4%	+ 1.4%
L	4.68	26	5.03	+ 6.9%	- 0.5%
M	9.39	26	10.23	+ 9.0%	+ 0.8%

The assigned values for pools L and M were increased by 8% to compensate for anomalous behaviour due to incompatible matrix. After this correction, bias found in analysis of control pools was believed to continue representing 'true' bias.

man samples from CDC - were each time analyzed in two runs only. A few CB-surveys with frozen serum, or with Rotterdam sera sent to Prague, were likewise performed in single determinations.

Figure 6.2. takes together the range of differences of our results and reference method results. These differences are determined much more by the precision of our method and of the reference method than by our methods bias. It is clear that there are no apparent signs now to indicate which set of results stems from human samples and which from lyophilized animal materials. But the mean bias of a set is a useful figure when the Huang method is employed.

Conclusions:

The use of simultaneous sets of several lyophilized samples in the external surveys gives results in which matrix effects are averaged out and approach the results obtained from sets of human sera. But to ascertain bias with reliability, human material must be considered superior.

6.4.2. KCA programme

When the KCA programme began, its intention was to standardize the cholesterol analysis in a number of laboratories of which practically all used a direct LB-reaction. The plan to develop a common calibration system with serum seemed ideal to reduce bias and improve comparability at the same time. Human serum was considered preferable to other possibilities. To get hold of freshly frozen human serum was not quite easy. Arguments against a plan based solely on employing it were for example:

1. Law in the Netherlands does not permit the use and commerce of human blood other than for the purpose of transfusions and through official blood banks.
2. The cost of human material purchased abroad would be high.
3. It was presumed that a WHO appeal to developing countries to oppose commercial blood donations or plasmapheresis would in the future make it considerable less obtainable (29).
4. CDC programmes make use of bovine serum mixed with human lipids in addition to human serum.
5. It was considered important to make a start within a limited time span because a number of epidemiological projects were being conducted without standardization.

As a consequence enriched equine pools became our compromise. These were to be frozen and not freeze-dried, to avoid errors in the reconstitution and inhomogeneity caused by dispensing and lyophilization.

When part I of the KCA plans was completed and part II began with experimental trials to test several serum calibrators a new situation emerged: the direct enzymatic methods were gradually beginning to replace the direct LB-reactions. New participants using enzymatic methods entered the programme.

Clearly this did not make standardization any easier, but there can be little doubt that the future will see enzymatic methods to be more and more applied in clinical chemistry. But since a considerable number of procedures and modifications exists already, each with their own problems, standardization still remains an unsolved problem.

In order to calibrate and control enzymatic methods as well, the need for human serum again becomes urgent, although improvements of the enzymatic reagents may eventually release this strain. For direct LB-methods our enriched serum served its purpose fairly well. When in both trials the final results obtained with calibrator sets A_1, A_2, A_3 and A_4, A_5, A_6 are compared, looking firstly at interlaboratory comparability, the second set is clearly superior. The standard deviations, already improved by calibrating with A_1, A_2 and A_3 are again much better when A_4, A_5 and A_6 are used.

Table 6.4. shows which improvements were statistically significant in the Huang method; one-sided t is used.

Table 6.5. in the same manner presents the enzymatic values; these do no justify the use of these calibration sera.

For enzymatic analysis human pools were very recently tried out and we are confident that calibration with this serum produces cholesterol levels that are quite

Table 6.4. (summarizing table 5.7. - 5.11).
Interlaboratory comparability. Huang method with serum calibrators (mmol/l): a=calibrators A_1, A_2 and A_3; b=calibrators A_4, A_5 and A_6; n = number of laboratories

Sample	Abell values n = 4	Huang routine n = 9	Huang a n = 9	Huang b n = 9	Standard deviations routine	Standard deviations a	Standard deviations b
trial 1							
I	4.30	4.75	4.09	4.25	0.150	0.112	0.087
II	5.76	6.26	5.42	5.61	0.250	0.148	0.117[2)]
III	6.37	6.95	6.02	6.22	0.273	0.180	0.149[2)]
IV	6.37	6.90	5.99	6.17	0.287	0.199	0.154[2)]
V	7.90	8.52	7.41	7.62	0.315	0.239	0.185[1)]
A_4 (4.24)	4.26	4.80	4.14	-	0.115	0.047[1)]	-
A_5 (6.28)	6.30	6.99	6.06	-	0.181	0.085[1)]	-
A_6 (8.79)	8.80	9.73	8.47	-	0.287	0.143[1)]	-
trial 2	n = 3	n = 17	n = 17	n = 17			
VI	9.45	9.79	8.76	9.09	0.593	0.355[1)]	0.273[2)]
VII	6.94	7.25	6.48	6.73	0.366	0.262	0.163[2)]
VIII	6.22	6.51	5.80	6.05	0.364	0.249	0.139[2)]
IX	5.45	5.86	5.21	5.44	0.321	0.241	0.144[2)]
X	3.31	3.51	3.09	3.26	0.196	0.173	0.101[2)]
A_4 (4.24)	4.26	4.66	4.13	-	0.204	0.062[1)]	-
A_5 (6.28)	6.30	6.75	6.01	-	0.306	0.140[1)]	-
A_6 (8.79)	8.73	9.83	8.30	-	0.457	0.281[1)]	-

[1)]significant improvement: $p < 0.05$

[2)]significant improvement: $p < 0.01$

Table 6.5. (summarizing table 5.7. - 5.11).
Interlaboratory comparability. Cholesterol oxidase methods with serum calibrators: a = calibrators A_1, A_2 and A_3; b = calibrators A_4, A_5 and A_6; n = number of laboratories.

Sample trial 2	Abell values n = 3	Enzymatic values (n = 7) routine	a	b	Standard deviations routine	a	b
VI	9.45	9.29	9.87	10.03	0.423	0.288	0.248
VII	6.94	6.87	7.36	7.42	0.288	0.248	0.281
VIII	6.22	6.16	6.62	6.67	0.249	0.384	0.416
IX	5.45	5.45	5.88	5.90	0.193	0.357	0.384
X	3.31	3.36	3.67	3.64	0.156	0.245	0.278
A_4 (4.24)	4.26	3.91	4.24	-	0.187	0.055[1)]	-
A_5 (6.28)	6.30	5.83	6.27	-	0.265	0.095[1)]	-
A_6 (8.79)	8.73	8.17	8.74	-	0.435	0.160[1)]	-

[1)]significant improvement with $p < 0.05$

close to Abell-results. Whether this also leads to better comparability is being tried out in a new comparative study.

A rough calculation of what is necessary in terms of volume of donor blood can be made as follows. If 40 laboratories calibrate every day with three pools - two ml of each daily - and an equal amount of a fourth pool is needed for enrichment, we need per year: 40 x 0.008 x 250 (working days) = 80 liter of serum, or approximately 480 donor units of blood. This should not be any burden at all upon regular blood supply and serve an excellent purpose in a most convenient way.

6.5. Classification errors

The question how many persons could be classified in a wrong category as a result of random analytical errors and biological variability may be approached in the following manner. The serum cholesterol frequency distribution in the population examined from October 1974 till the end of 1975 in the CB Heart Project is used as an example (table 6.6.). How many individuals put into the normal and borderline group really had a cholesterol level above 7.25 mmol/l at that moment and, conversely, what percentage of our 'elevated cholesterol' group in fact did have a borderline or normal serum cholesterol value?

Table 6.6. Frequency distribution of serum cholesterol. All baseline examinations, October 1974 - December 1975. CB Heart Project.

Cholesterol mmol/l	number of persons
< 3.2	16
3.2 - 3.6	88
3.7 - 4.1	347
4.2 - 4.6	717
4.7 - 5.1	1104
5.2 - 5.6	1155
5.7 - 6.1	971
< 6.2	**4398**
6.2 - 6.6	678
6.7 - 7.2	499
> 6.2 - ⩽ 7.2	**1177**
7.3 - 7.7	221
7.8 - 8.2	113
8.3 - 8.7	45
8.8 - 9.2	33
> 9.2	19
> 7.2	**431**

The following assumptions are made for this particular calculation:
Symbols are listed as well:

s_l = standard deviation in laboratory: 0.25 mmol/l
s_b = standard deviation of intra-individual variation: 0.40 mmol/l
$s_t = \sqrt{s_l^2 + s_b^2} = 0.47$ mmol/l
y = serum cholesterol level expected at examination (best estimate)
x = actual laboratory result at examination
k_j = mean cholesterol result of section j in the frequency distribution
e = y - x (laboratory error, having a normal distribution)
G = 7.25 mmol/l, being the cut-off level between elevated and borderline cholesterol

e_j = G - k_j; smallest laboratory error that causes wrong classification

P(x=k) = fraction of population (with cholesterol level between (k ± 0.2) mmol/l)

P(e > (or <) G - k) = chance that laboratory error (e) is larger (or smaller) than the difference between someones cholesterol level (k) and cut-off level G

P_1 = P(y > G|x < G); fraction of the population that **has** an elevated cholesterol level, but was analyzed as borderline or normal

P_2 = P(y < G|x > G); fraction of the population that was put into elevated group, but does **not have** high cholesterol levels.

The two equations for P_1 and P_2 are as follows:

$$P_1 = \frac{\sum_{k<G} P(x = k) \,.\, P(e > G - k)}{\sum_{k<G} P(x = k)}$$

$$P_2 = \frac{\sum_{k>G} P(x = k) \,.\, P(e < G - k)}{\sum_{k>G} P(x = k)}$$

For this simplified approach it is assumed that all individuals within a certain interval of 0.4 mmol/l had serum cholesterol levels equal to the mean of their interval (k_j). The numbers of people belonging to each interval are given in table 6.6.

The chance P ($e \gtrless (G - k)$) was obtained by transforming the normally distributed variable e into a standard normally distributed variable Z in the following way:

$e \cong N(0, \sigma_m^2)$; becomes: $Z = \frac{e}{\sigma_m} \cong N(0, 1)$

For example the probability of 'e' exceeding 'a' now becomes equal to the probability that Z is larger than $\frac{a}{\sigma_m}$ or, conversely, the probability that Z is smaller than $-\frac{a}{\sigma_m}$:

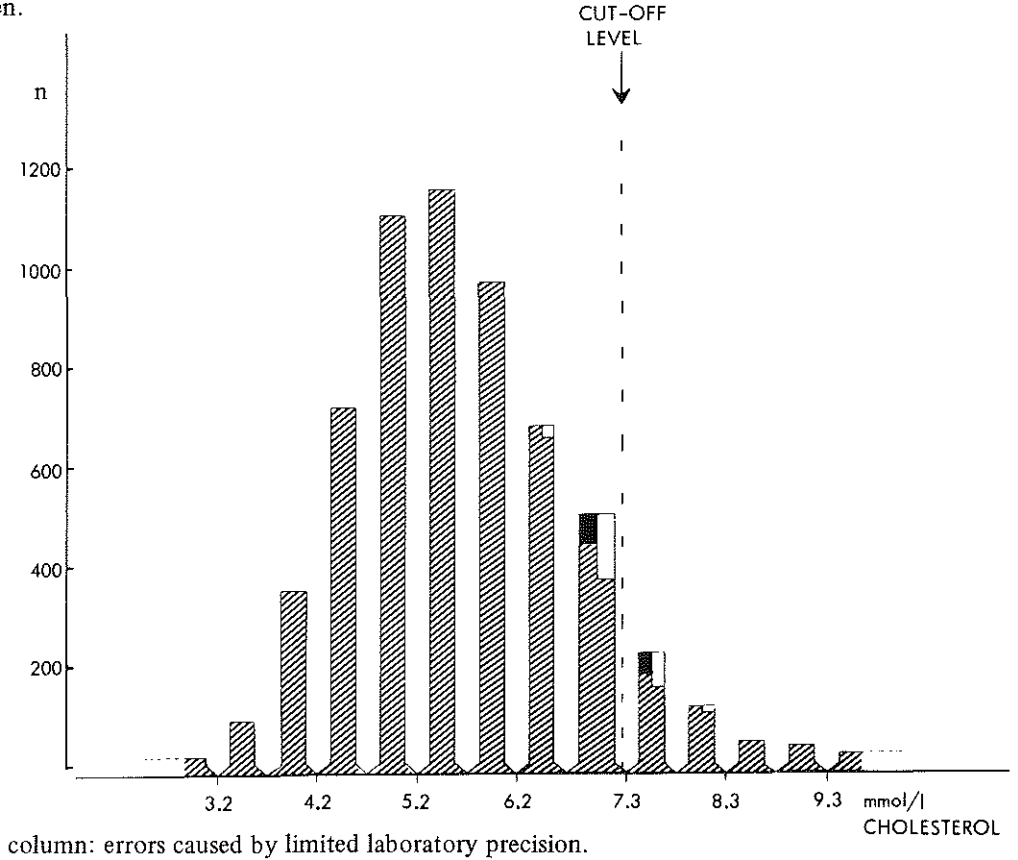
$P(Z < -\frac{a}{\sigma_m})$

The tail-probability can now be found in a table of the standard normal distribution after substituting (G - k) for 'a' and s_l and s_t for σ_m.

This tail-probability is practically zero for cholesterol sections that are one mmol/l or more away from G. The calculated values for P ($e \gtrless (G - k)$) were also included in table 6.7. together with all other necessary quantities.

The results are at the same time shown in figure 6.3.

Figure 6.3. Wrong classifications in frequency distribution of serum cholesterol, CB Heart Project 1975. Total of 6006 men and women.



height of black column: errors caused by limited laboratory precision.
height of white column: 'errors' if biological 's' is also considered.

Table 6.7. Data used in calculating erroneous classifications and results.

cholesterol category	n	k_j mmol/l	e_j mmol/l	P(x=k)	P(e < G - k) s: 0.25 mmol/l	0.47 mmol/l	P(e > G - k) 0.25 mmol/l	0.47 mmol/l
Total	6006	-	-	1				
5.2 - 5.6 mmol/l	1155	5.4	- 1.85	0.192	0	0		
5.7 - 6.1	971	5.9	- 1.35	0.162	0	0.002		
6.2 - 6.6	678	6.4	- 0.85	0.113	0.0003	0.0359		
6.7 - 7.2	499	6.9	- 0.35	0.083	0.1151	0.2611		
7.3 - 7.7	221	7.5	0.25	0.037			0.1587	0.2980
7.8 - 8.2	113	8.0	0.75	0.019			0.001	0.0559
8.3 - 8.7	45	8.5	1.25	0.007			0	0.004
8.8 - 9.2	33	9.0	1.75	0.005			0	0.0001
below G	5575			0.928				
above G	431			0.072				

Results [1)]	s = 0.25 mmol/l	s = 0.47 mmol/l
P_1	57 (1%)	154 (2.5%)
P_2	35 (8%)	72 (17%)
Total misclassified	92 (1.5%)	226 (4%)

[1)] WHO narrow limits at this concentration level require s_I to be 0.19 mmol/l. Using that value the number of errors will be slightly lower than those in the left column.

We do not know very much about biological variation in serum cholesterol but we can use literature data suggesting a fluctuation with s = 0.40 mmol/l is a reasonable estimate. When combined with laboratory precision (reproducibility is known, table 4.2.2.) a total standard deviation of 0.47 mmol/l is obtained:

Laboratory CV = ± 3% = 0.25 mmol/l

Biological CV = ± 5-6% = 0.40 mmol/l

s overall = $\sqrt{0{,}25^2 + 0.40^2}$ = 0.47 mmol/l

We calculated that among 431 persons in the elevated group, 35 are probably misclassified, whereas 57 persons were missed. Suppose we do not consider the cholesterol level which the examined subjects happen to have on the day of baseline screening important. And that instead we would like to let the mean level, around which the biological fluctuations occur, decide about classification. Then the persons accidentally having a 'relatively high' cholesterol concentration during the examination day will be classified 'wrongly' as elevated. Then the overall s must be employed and we find 72 persons in the elevated group and 154 in the normal group that in fact should have been on the 'other side'. However, because our decision level of 7.25 mmol/l is an arbitrary one, and does not change for different age groups, the term 'wrong classification' is a very relative one. Besides, the calculations are performed with a frequency distribution having 0.4 mmol/l intervals. This introduces an element of guess work but we believe a realistic impression has been given.

Moreover, it is not possible to make measurements for checking these calculations.

After completion of the classifying process there are examinations at a later time. Elevated cholesterol values are checked for instance. We now run into the 'regression-towards-the-mean' phenomenon, which will not be discussed here (30). The regression is such that 'elevated' persons upon rescreening show a lower mean value. This effect must be carefully distinguished from lowering as a result of intervention (diet, drugs).

Table 6.8.1. and 6.8.2. show what happens when individuals with increased risk are rescreened. The cut-off point in the CB Heart Project was higher than that in the KRIS and therefore the regression towards the mean there is more pronounced. In addition to regression we have the biological variation which moves the cholesterol level across the cut-off level in both directions for several of these persons.

Table 6.8.1. Serum cholesterol changes in 'elevated group' in a 3 week period. CB Heart Project, 1976.

	first analysis ⩾ 7.3 mmol/l	repeated analysis in the same group ⩽ 7.2 mmol/l	percentage lower
Men	623	211	33.9%
Women	295	118	40.0%
Total	918	329	35.8%

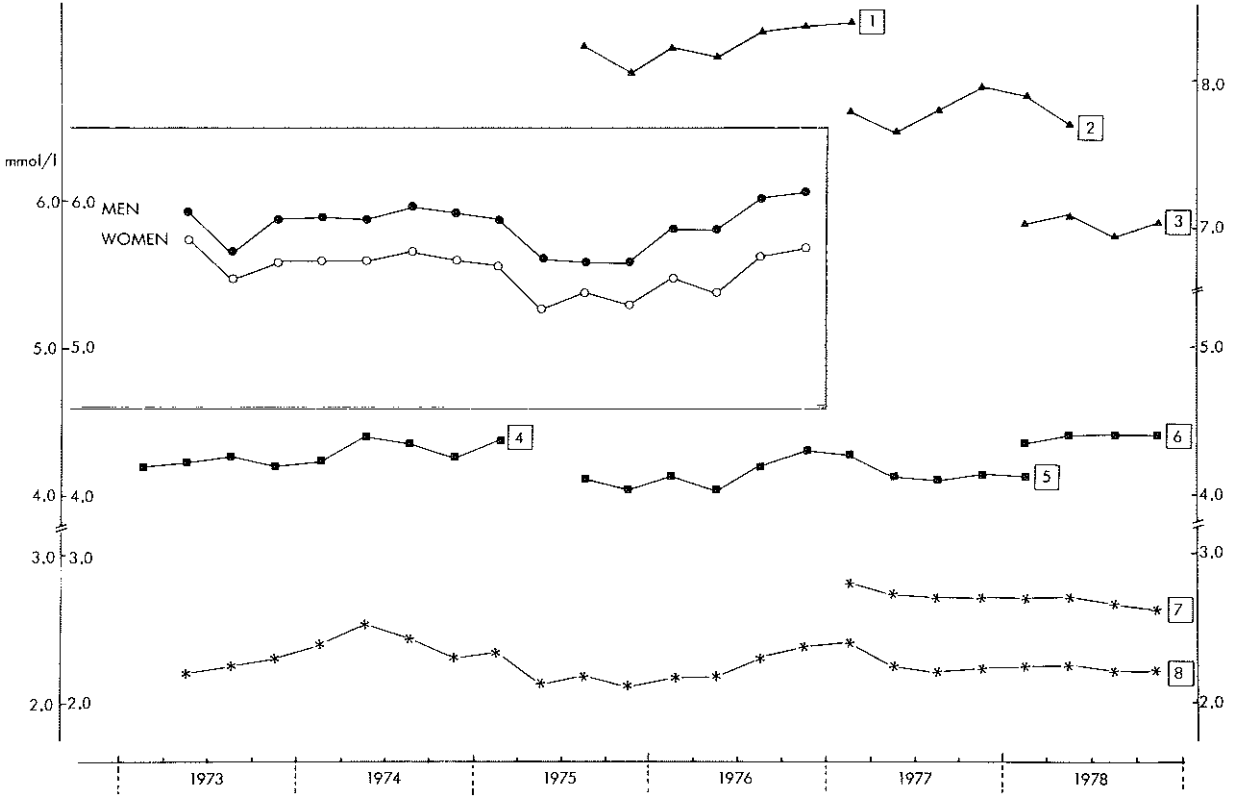
Table 6.8.2. (from table 3.10.). Plasma cholesterol changes in 'borderline group' in a 6 week period (KRIS, 1973).

	serum ⩾ 6.5 mmol/l plasma ⩾ 216 mg%	second analysis serum ⩽ 6.4 mmol/l	percentage lower
Rotterdam men	666	166	25%
Kaunas men	366	78	21%

6.6. Seasonal variations

The eventuality of serum cholesterol levels being dependent of the time of the year is the subject of many a study. This was reviewed in chapter 2. Some studies have followed a group of volunteers, others have compared the average cholesterol levels in different samples of the population. In this case comparability of these samples may be limited. Observations over periods longer than a year are exceptional and the observed variations do not have a consistent pattern. Sometimes winter levels seem to be higher than those in summer and the absence of trends has been reported as well. Men and women sometimes have shown differing trends. But pertinent data on quality control procedures are not communicated. In our CB Heart Project seasonal changes in the screened population may be visualized by compiling the averages of three months and plotting those graphically. In figure 6.4. they are put together with results that were likewise obtained from process control data. In several parallel segments of the curve the similarity is striking. Sometimes, however, the lines diverge. An estimate of laboratory drift was used to

Figure 6.4. Seasonal mean cholesterol values in men, women and control sera.



1, 2: Elevated Lipid Control, two lots; 3: Monitrol II-X, 4: Monitrol I; 5, 6: Monitrol I-X; 7: RIV serum B; 8: RIV serum A.

draw a corrected version of seasonal average serum cholesterol levels. Clearly there is no predictable, regular pattern present and only slight, irregular fluctuations occur. This corrected seasonal trend is shown in figure 4.4.

In the KRIS study seasonal trends were not observed either (figure 5.3.) although in this case the number of examined subjects per month was relatively small and did not represent a random sample and therefore we should be careful making general conclusions.

In the case of the CB Project the seasonal mean values are calculated for very large groups. These are every time composed of people aged 20 - 50 years. Due to the fact that varying age groups have been examined within the CB's and their relative contribution to the total number was subject to changes also, it proved to be quite difficult to single out a number of rather narrow age categories and follow them through all 3½ years. Only some groups could be followed over single year periods and a Friedman test did not indicate that one month or another had high or low cholesterol results.

We did not find a seasonal effect that might be used to 'correct' intervention results or to modify other data obtained in the screening procedures. We can also conclude that a seasonal effect - should it exist - is most probably equal for men and women and rather a small one.

6.7. Requirements for cholesterol analysis in epidemiological investigations, some final observations

6.7.1. Precision

To establish the risk of getting an infarction for an individual by measuring his blood cholesterol a high precision is needed. Incorrect classifications on both sides of cut-off points, however, always occur as the precision is never perfect. Their number is not equally distributed on either side (section 6.5.). Precision is preeminently a challenge for each laboratory itself. The most reliable way to measure analytical standard deviations is by blind analysis of human serum samples (section 4.4.). The possibility of clerical errors remains in blind set-ups but this adds to a realistic estimate of overall precision.

A 'sufficient' level of precision must be predetermined by the experts setting up a particular study. When thinking in terms of biological variation (section 2.2.) one should be aware of interindividual differences. The intra-individual standard deviations of the human serum cholesterol variations - considering time intervals of more than a day - may be as low as 10 mg% (0.26 mmol/l) - table 2.4. If we now accept the total value for s to be 10% higher than the biological one, this total s

may become 0.28 mmol/l (11 mg%) and the laboratory s must be lower than 0.12 mmol/l (4.6 mg%), in the normal cholesterol range. The WHO recommended (narrow) limits (table 2.7) are shown to be attainable in most KCA laboratories. They may be considered as minimum requirements, although we must note that the measurements were not done by a blind procedure. With this norm ($s < 0.19$ mmol/l at $x = 7.25$ mmol/l) incorrect classifications occur at a rate comparable with those in figure 6.3. Very recently in a conference on reference methods and materials it was stated that analytical CV must be $< \frac{1}{2}$ intra-individual CV. In our example $\frac{1}{2}$ intra-individual s equals $\frac{1}{2} \times 10$ mg% = 5 mg%, which agrees well with the above-mentioned 4.6 mg%. The proceedings of this conference contain a wealth of useful discussions, that were held in the USA on the development of reference methods and materials (21).

6.7.2. Accuracy

Accuracy largely depends on analytical specificity. In laboratories where the Abell-Kendall procedure can be properly carried out - or a partially mechanized version - this is not an insurmountable problem. When larger work loads preclude the use of extraction steps and direct methods have to be selected, one has to decide whether bias is acceptable. Identical bias in laboratories of multi-centre studies still permits the comparison of data (section 2.7. and 3.4.). For prolonged comparability this bias must be established in cooperation with an official reference laboratory. Bias reduction by means of properly selected calibration sera is a useful improvement (section 5.5.).

The WHO criteria for bias are not automatically sufficient for pooling of data (section 2.7.). Besides, the proportional limit of 5% does seem too rigid for low cholesterol levels and rather large for elevated levels. In any case bias must be estimated at several concentration levels by every single laboratory (section 5.4. and figure 5.1.).

This bias of the particular laboratory must be taken into account when its results are studied.

6.7.3. Methods

Methods for cholesterol analysis, and evaluations of them, have been published by the hundreds. A recent, and very informative bibliography lists 972 references! (31).

Normalization of methodologies to be used in epidemiology reduces the number of problems one may expect to run into during standardization procedures (section 3.3. and 5.2.). When several methods are adopted at first and then standardization is attempted we are obviously closing the stable-door after the steed has

been stolen! (section 5.3. and 6.1.) Even when improved techniques come into being they cannot at once be introduced into running studies. Method normalization was carried through in the USA for this reason in the Lipid Research Clinics and the Hypertension Detection and Follow-Up Programme (HDFP) (32, 33) and in the KRIS. All of the laboratories involved are using a LB-reaction.

Today the enzymatic procedures are available as well (section 2.1.6.). The production and delivery of stable reagents with guaranteed enzyme activities is not easy. Interferences are by no means ruled out and differ among the many modifications. If kits are to be used it is also necessary that reagent manufacturers refrain from introducing product modifications without announcing them in advance. The enzymatic methods will come into more general use and it must be realized that standardization is not less strenuous than with other methods (section 5.3.).

6.7.4. Quality control

Quality control procedures must include process control (section 6.3.) based on a careful description of the analytical procedure and on limits for rejecting or accepting results (section 3.3. and Appendix I). Blind control is not difficult to include but is does require special care on the side of the sender of the extra samples (section 4.4.). The laboratory must also be prepared to make all quality control results available to the evaluators of the study results. Evidently the best way to obtain such objective information is the external part of the control organization (table 3.7.).

6.7.5. Materials

Control materials must resemble human serum closely enough to avoid erroneous conclusions caused by matrix effects (section 6.4.). Although precision estimates may be valid, the bias seen in measurements of control sera does not reflect the 'true' bias. The approach of the WHO Regional Lipid Reference Center in Prague of making corrections for these matrix effects probably does not create substantial errors but it is not an optimal alternative. Additional comparative measurements and communication were necessary to unravel all problems connected with such corrections which meant an additional burden on both parties involved. One may expect that the pretentious WHO Standardization Programme employs materials that clear up accuracy problems in stead of creating bias. For that purpose adequate funding of the Reference Laboratory is elementary. Purchase or production costs and shipments in dry-ice containers require more financial support and man-power. Such support must be provided by the WHO - it is not primarily the task of the laboratory director. This does not imply that every clinical chemistry laboratory must have human serum to verify every measurement of every analyte - a limited volume of serum suffices for this sort of programme.

6.7.6. Reference laboratories

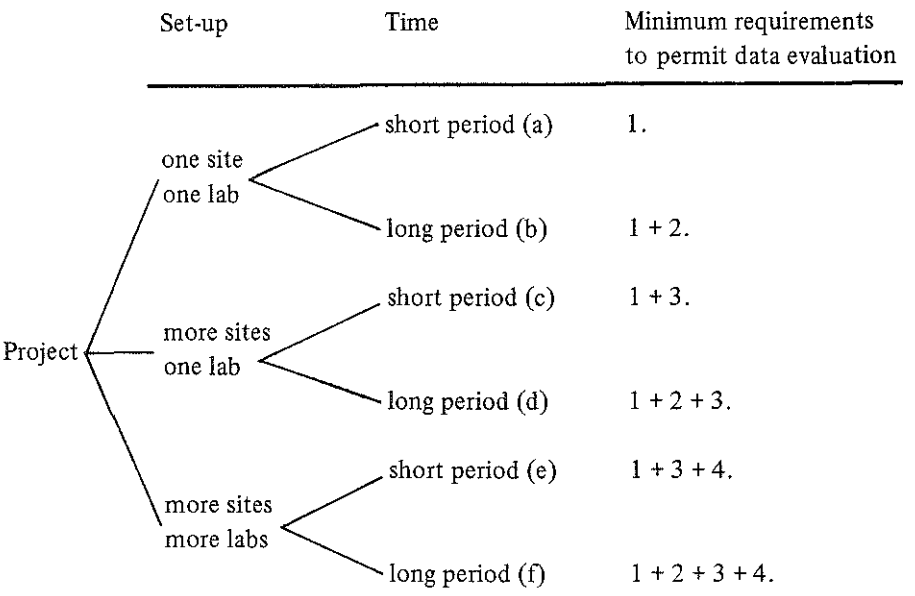
When many laboratories get involved in screening a great number of people, care must be taken that enough time is spent in thorough standardization. More reference centers are needed than the present two (CDC and IKEM). CDC already stimulates the development of more, regional or national reference centers. Such coordination already is taking place in South Australia (Dr. T. D. Geary). These centers are ideal for concentrating communication and for distribution of reference serum in a particular area. Technical support and advice can be given easier. A reference method must be established and maintained in these centers. When quality control procedures and the contacts with CDC and/or IKEM are extensive, reliable reference method values can be produced locally and issued as assigned values. There is no need for large networks of reference laboratories for serum cholesterol analysis. One or two might do this job in the European Community. A cost-benefit analysis should be made to study the financial consequences of such enterprises. Research into definitive methods may be done (but does not have to be) in the same institute; one such centre - somewhere in the EC - would be a wellcome partner of America's NBS to further explore definitive methodology.

6.7.7. Epidemiological investigations

This thesis has emphasized the necessity of comparability of results obtained in large scale studies. But there is a countless number of ways to set up such investigations. It is conceivable that epidemiological studies are planned as 'closed' entities in which final conclusions are drawn in such a way that these are universely useable. For instance one can always single out high risk groups, using arbitrary cut-off points, and measure intervention results by means of methods that are precise. If this is the scope of a project no comparability is required.

A simple scheme is drawn below in which minimum requirements regarding laboratory proficiency in various studies are outlined. These requirements have been listed in four headings:

1. precision: keep s within predetermined limits
2. same as 1; with special attention for long-term reproducibility
3. extra attention for uniformity of sample collection and handling
4. accuracy: if the methods are not accurate, at least their bias must be established.



An example of the first category (a) might be a plan to determine whether persons with illness X have higher or lower serum cholesterol levels than persons without the disease. A target population is selected and the collection of blood and the analysis take place within a short period of time. The conclusion 'illness X is attended with elevated serum cholesterol levels' is useful information. However a publication of the serum cholesterol values is not. Without proper standardization they have no clear meaning outside the scope of the study.

An example of the second category (b) could be an evaluation of the effectiveness of drug treatment. Serum cholesterol levels are determined in patients before and after the drug has been taken for some time. The results are compared with changes found in an untreated group. After a proper evaluation of the data a conclusion like 'drug Y reduced serum cholesterol levels on average more than 1 mmol/l' is useful information. (Unless the presence of the drug in the serum decreased the analytical result by interference). Publishing the cholesterol results separately is useless by lack of a frame of reference.

In set-ups like the last category (f) meaningful information on serum cholesterol concentrations alone can be gathered, because these are comparable with other standardized results. Precision is always necessary and particularly when extended periods of time are involved, the assessment of s by means of a 'blind' control mechanism is essential. Once the reproducibility is under control the job of standard-

ization is practically and in theory an easy one. It involves the measurement of bias with reference serum having a compatible matrix, or with series of human serum samples from the study, in cooperation with a reference centre. This particular measurement does not need frequent repetition as long as the checks on precision confirm an extended method stability. In terms of effort - including financial cost - the extra work, after s or CV has been brought within specified limits (and that is always vital) is relatively small. The importance of proper blood sampling and handling procedures is described in section 3.4.1.3 and Appendix I.

All measurements regarding risk indicators - whether they are used as a basis for planning prevention of atherosclerotic diseases or to inform the individual about 'risk' - must lead to uniform information. When people in various parts of the country have their serum cholesterol level checked and learn in one place that 7.0 mmol/l is perfectly normal and elsewhere it is considered elevated, no successful preventive strategy can be initiated. Normalized information must provide a sound substructure.

Requirements for properties of tests that are employed in various mass health examinations have been proposed in the literature (22, 23). The concept of precision in these papers is treated along the same lines as is done above. Precision is crucial in the diagnosis of individual cases and less important for overall estimates that concern groups.

'Accuracy' is defined next as 'providing a true measurement of the attribute being sought' and it is therefore closely related to clinical sensitivity and specificity. In the case where the test is serum cholesterol analysis and the characteristic sought is susceptibility to atherosclerotic disease, these test properties can hardly be applied. The exact numbers of those with the disease (risk) and of those free from it cannot be provided. However it is possible to make approximations and it is fairly safe to state that both sensitivity and specificity in this particular situation are rather low.

Therefore in screening and case-finding manoeuvres, the measurement of only this single risk indicator will not be worth-while and several others have to be taken into account. In epidemiological surveys sensitivity and specificity are of minor importance.

Additional properties that are important for all of these studies are simplicity and low cost.

In these requirements the comparability of results obtained in the numerous surveys, screening and other programmes is not mentioned although all of these will eventually be published - and have been published. The main issue in the discussion given above, is the need to reduce the confusion that is created when analytical data without a basis for comparison appear in the literature.

6.8. Final remarks

Strict comparability is not always needed (section 6.7.7.) in every clinical or biochemical research setting. Although these areas will eventually benefit from the described efforts to improve the quality of chemical analysis they may still, in a satisfactory way, employ the results furnished by the procedures currently offered by the clinical chemist.

Though not all, certainly many epidemiological investigations in which laboratory analyses are carried out require that these be performed with more than average proficiency. In selecting the laboratories, proof of such capability will be demanded ever more by the planners of the study. The laboratory must then produce a certificate issued by an official standardization programme and it should continuously take part in external surveillance. The laboratories must be large enough, and sufficiently expert, to carry out the necessary quality control procedures. One may also expect future financial support from government or other institutions to depend on the degree of standardization of measurement; certainly when risk indicators are involved. As a result, eventually all hospital laboratories will pursue comparability of results and thereby reduce the need for test duplications and increase the value of literature data.

The methods used must be of proven reliability and control materials must be available. The ways and means to achieve this have been outlined and discussed in this dissertation.

In the Netherlands the KCA and the Netherlands Lipid Reference Laboratory will pursue proper standardization of all determinations done for large scale medical examinations. International cooperation will be needed to ensure general comparability.

A cost-benefit analysis of such international enterprise is outside the scope of this thesis.

REFERENCES

Chapter 6

1. Clin. Chem. *23*, 1517 (1977).
2. B. LEWIS: The Hyperlipidemias. Clinical and Laboratory Practice. Blackwell Scientific Publications, Oxford (1976).
3. G. J. MILLER and N. E. MILLER: Plasma-high-density-lipoprotein concentration and development of ischaemic heart disease. Lancet *1*, 16 - 19 (1975).
4. T. GORDON, W. P. CASTELLI, M. C. HJORTLAND, W. B. KANNEL and T.

R. DAWBER: High density lipoprotein as a protective factor against coronary heart disease. Amer. J. Med. *62*, 707 - 714 (1977).

5. G. J. MILLER: Editorial. High-density lipoprotein, low-density lipoprotein, and coronary heart disease. Thorax *33*, 137 - 139 (1977).
6. S. A. COBB and J. L. SANDERS: Enzymatic determination of cholesterol in serum lipoproteins separated by electrophoresis. Clin. Chem. *24*, 1116 - 1120 (1978).
7. J. M. GOLDBERG: Test precision and high-density-lipoprotein cholesterol. Clin. Chem. *24*, 2061 (1978).
8. T. J. BRONZERT and H. B. BREWER Jr: New micromethod for measuring cholesterol in plasma lipoprotein fractions. Clin. Chem. *23*, 2089 - 2098 (1977).
9. J. J. ALBERS, G. R. WARNICK, D. WIEBE, P. KING, L. SMITH, C. BRECKENRIDGE, A. CHOW, K. KUBA, S. WEIDMAN, H. ARNETT, P. WOOD and A. SCHLAGENHAFT: Multi-laboratory comparison of three heparin-Mn^{2+} precipitation procedures for estimating cholesterol in high-density lipoprotein. Clin. Chem. *24*, 853 - 856 (1978).
10. W. A. BRADLEY and A. M. GOTTO: Structure of intact human plasma lipoproteins, in: Disturbances in lipid and lipoprotein metabolism. J. M. Dietschy, A. M. Gotto, J. A. Ontko eds. Am. Physiol. Soc., Bethesda Md. (1978).
11. J. A. GLOMSET: Plasma-lecithin:cholesterol acyltransferase, in: Blood lipids and lipoproteins: Quantitation, composition and metabolism. G. J. Nelson ed. Wiley-Interscience, New York (1972).
12. G. BONDJERS, G. OLSSON, L. L. NYMAN and S. BJOKERUND: High density lipoprotein (HDL) dependent elimination of cholesterol from normal arterial tissue in man, in: Proceedings of the Fourth International Symposium on Atherosclerosis. G. Schettler, Y. Goto, Y. Hata, G. Klose eds. Springer Verlag, Berlin (1977).
13. The 'alpha Q2' control from CDC. Its target value is the mean of results from 12 Lipid Research Clinics. See a, b and c:
    a. K. O. ASH and W. M. HENTSCHEL: High-density lipoprotein estimated by an enzymatic cholesterol procedure, with a centrifugal analyzer. Clin. Chem. *24*, 2180 - 2184 (1978).
    b. R. J. LIEDTKE, B. BUSBY and J. D. BATJER: Use of the DuPont'aca' to measure cholesterol in high-density lipoprotein fractions prepared by the heparin-Mn^{2+} precipitation method. Clin. Chem. *24*, 161 - 165 (1978).
    c. F. VAN DER HAAR, C. M. VAN GENT, F. M. SCHOUTEN and H. A. VAN DER VOORT: Methods for the estimation of high density cholesterol, comparison between two laboratories. Clin. Chim. Acta, *88*, 469 - 481 (1978).

14. IFCC. Provisional recommendation on quality control in clinical chemistry. Part 1. General Principles and Terminology. Clin. Chim. Acta *63*, F25 - F38, following p. 108 (1975).
15. ASTM. Recommended Practice for Use of the Terms Precision and Accuracy as applied to Measurement of a Property of a Material. Publication E 177-71 (1971).
16. NNI. Terminologie voor het beschrijven van meetnauwkeurigheid. Norm NEN 3114, The Hague (1973).
17. J. R. ALLEN, R. EARP, E. C. FARRELL and H.-D. GRUMER: Analytical bias in a quality control scheme. Clin. Chem. *15*, 1039 - 1044 (1969).
18. M. S. WEINBERG and R. N. BARNETT: Absence of analytical bias in a quality control program. Am. J. Clin. Pathol. *38*, 468 - 472 (1968).
19. C. E. DENT: Blood, plasma or serum? Lancet *2*, 866 - 867 (1975).
20. R. J. EILERS: Quality assurance in health care: missions, goals, activities. Clin. Chem. *21*, 1357 - 1367 (1975).
21. J. H. BOUTWELL ed.: A National Understanding for the Development of Reference Materials and Methods for Clinical Chemistry. Proceedings of a Conference. AACC, Washington DC (1978).
22. D. L. SACKETT and W. W. HOLLAND: Controversy in the detection of disease. Lancet *2*, 357 - 359 (1975).
23. H. A. VALKENBURG: Vroege opsporing en bevolkingsonderzoek; in: Oncologie, 2e druk. Stafleu, Leiden (1978).
24. M. DHARAN: Total Quality Control in the Clinical Laboratory. The C.V.Mosby Co. Saint Louis (1977).
25. R. LAESSIG, D. HASSEMER and G. HOFFMAN: Determination of optimal (absolute minimum) vial-to-vial variations in four commercial and one special control serum: Implications for proficiency testing and intralaboratory quality control. Clin. Chem. *24*, 1049 (1978).
26. A. P. JANSEN, E. J. VAN KAMPEN, C. A. M. MEIJERS, P. J. J. VAN MUNSTER and G. J. M. BOERMA: Quality control and the quality of commercial test sera. Clin. Chim. Acta *84*, 255 - 258 (1978).
27. G. N. BOWERS, R. W. BURNETT and R. B. McCOMB: Preparation and use of human serum control materials for monitoring precision in clinical chemistry, in: Selected Methods of Clinical Chemistry. Vol. 8, G. R. Cooper ed. AACC, Washington DC (1977).
28. D. GRAFNETTER: External quality control carried out by the WHO European Lipid Reference Center in Prague, in: International Conference on Atherosclerosis. add. L. A. Carlson, ed. Raven Press, New York (1978).
29. DIRECTOR-GENERAL OF THE WHO: Utilization and Supply of Human Blood and Blood Products. Official Records of WHO *226*, Part I. Resolutions and decisions; Annexes, 110 - 114. WHO, Geneva (1975).

30. C. E. DAVIS: The effect of regression to the mean in epidemiological and clinical studies. Amer. J. Epid. *104*, 493 - 498 (1976).
31. R. W. A. OLIVER and R. WARBURTON: An indexed, comprehensive bibliography on the methodology for the chemical analysis of serum cholesterol. Publication no. 22, Biological Materials Unit. Dept. of Chemistry and Applied Chemistry. University of Salford. Peel Park, Salford M4 4WT, UK (1977).
32. Manual of Laboratory Operations. Lipid Research Program, Vol. I. Lipid and Lipoprotein Analysis 1974. National Heart and Lung Institute, NIH, Bethesda Md. (1974).
33. HDFP COOPERATIVE GROUP: The Hypertension Detection and Follow-up Program. Prev. Med. *5*, 207 - 215 (1976).

SUMMARY

The rate of clinical chemical measurements of serum cholesterol has been influenced particularly by its role in predicting the risk of developing cardiovascular disease. In this thesis some important aspects - mainly quality control - of this laboratory test are studied, that are particularly related to its function in epidemiological investigations.

Chapter one describes the appearance of large scale studies - mainly in the industrialized countries - on the spread of diseases resulting from progressive atherosclerosis. Serum cholesterol analysis provides a tool for the prediction of susceptibility to such disease. The measurement of this risk indicator - serum cholesterol - therefore must produce reliable results. In view of the common interest and the international communication among epidemiologists and other scientists, comparability of the analytical methodology and results is required. The multiplicity of techniques for cholesterol estimation combined with a lack of common bases for accuracy and precision have stood in the way of universal comparability of such results.

Chapter two reviews the literature on several subjects that are connected with the measurement of serum cholesterol concentrations and their interpretation. The principal problems regarding the most frequently used Liebermann-Burchard reaction and the enzymatic method are outlined.

Interpretation of cholesterol values requires knowledge about biological (physiological) variations. Requirements for analytical precision and accuracy are discussed and these are closely connected with biological variability, within and between human beings. The way to attain proper laboratory proficiency, through quality control and the use of reference methods and materials, is then described. It is also demonstrated that the WHO recommended limits for accuracy do not permit pooling of data from several laboratories. Recently a Netherlands lipid reference laboratory has been established.

Chapter three deals with international standardization between two clinical chemistry laboratories and a reference laboratory. In the KRIS, during three years, plasma cholesterol and glucose measurements were carried out with results that on the average did not differ more than 2% between Rotterdam, the Netherlands and

Kaunas, Lithuania, USSR. Furthermore, the glucose values were accurate; the cholesterol measurements had a positive bias but it was maintained constant. The precision was rigidly controlled and the CV kept within predetermined limits. The overall laboratory proficiency was very much a result of strict adherence to the protocol (which is entirely reproduced in Appendix 1), and of intensive communication between the reference laboratory (CDC) and between the two participating centres.

Chapter four especially focusses on the intralaboratory quality control system that grew into its present form as a part of the CB Heart Project in which the CKCL is the single, central laboratory. The most important parts of its quality control procedures are: process control, blind external control and the survey programme of the WHO European Regional Lipid Reference Center (IKEM, Prague).

The process control enabled us to monitor long-term, small fluctuations in the laboratory and to make retrospective corrections in seasonal variation data. Blind control furnishes the most reliable method to determine repeatabiliy and reproducibility. Analysis of blind duplicates with three week intervals results in a CV of 2 - 4% and some values exceeding this level are a result of relatively rare outliers.

The materials employed in quality control greatly influence the bias one measures, which may complicate standardization procedures.

Chapter five is devoted to the standardization programme of the KCA foundation. This programme is being developed in the Netherlands in order to make laboratory results comparable, which can at present only be interpreted with difficulty. Several epidemiological investigations and health check-up schemes are still carried out with non-standardized laboratories. The evaluation phase (part I) and the preliminary standardization phase (part II) are described. A number of practical difficulties, mostly concerned with procurement or preparation of proper reference sera, had to be conquered. It also seems that a lack of rigid intralaboratory quality control is not uncommon. We have found that calibration based on a curve obtained with three serum cailbrators with known assigned values (Abell-Kendall) greatly improves interlaboratory comparability when a direct LB-reaction is used. For enzymatic methods human serum is prepared and studied for the same purpose. The bias is this case does not need much improvement.

In chapter six follows a discussion about the results of cholesterol analysis. Clinical chemists in general are not sufficiently concerned with standardization and still introduce a multitude of methods for the same analyte. Thus HDL-cholesterol results are being produced today that can hardly be interpreted satisfactorily by the average reader of publications.

For better understanding firstly some uniformity of terminology is necessary. Then an organization of quality control in the laboratory is proposed that doesn't need to be complicated at all.

Usually, commercially available materials for quality control may very well be employed for checking precision. For establishing analytical bias they are often less suitable. Even the serum used by the WHO Reference Laboratory in Prague needs improvement in this respect. The serum calibrators prepared at the RIV for the KCA programme may well serve the purpose in the near future of bringing the results obtained in Dutch population studies in closer agreement. Comparability of data depends greatly on the proper establishment of analytical bias.

The biological variability of cholesterol in the blood is such that laboratory precision does not play a decisive role when changes of the serum cholesterol level are studied; as long as laboratory 'drift' is avoided the mean cholesterol change in population samples can be evaluated properly. Great precision is needed when at the first examination (e.g. baseline screening) individuals are classified into risk categories. Several subjects may for instance be suspected wrongly of having elevated cholesterol levels when their values are close to cut-off points. Properly recorded quality control data can retrospectively provide a means to correct e.g. the calculated mean cholesterol level in seasons. No seasonal variation appears to be present in this country although the conclusion is somewhat restricted by the selection of the population samples.

It is finally concluded that good precision is always vital, particularly in large studies even when the investigators should not require comparability with others. However, as soon as this precision has been brought within specified limits, the measurement of bias is not too much of an extra effort. At present the proper establishment of bias by very many laboratories is impossible by lack of sufficient reference serum; but those relatively few involved in mass health examinations certainly should perform the additional work at little extra costs. Undoubtedly the selection of laboratories for participation in epidemiological investigations will, in the near future, occur mainly on the basis of proficiency, proven beforehand by authorized certification.

The Appendix, last but not least, presents the full details of the KRIS laboratory procedures, that have led to a unique international, interlaboratory standardization. It was largely based on the experience of Dr. A. Mather, CDC, who advised the KRIS laboratories throughout the official life time of the study.

SAMENVATTING

Het aantal serum-cholesterol bepalingen verricht in klinisch-chemische laboratoria is bijzonder beïnvloed door de mogelijkheid de uitkomsten te hanteren als voorspeller van het risico op het ontstaan van hart- en vaatziekten. In dit proefschrift worden enkele belangrijke aspecten van de kwaliteitsbewaking van deze laboratorium bepaling bestudeerd welke verband houden met de plaats ervan in epidemiologische onderzoekingen.

Hoofdstuk 1 signaleert de onderzoekingen op grote schaal, vooral in de geïndustrialiseerde landen, naar de verspreiding van ziekten die een gevolg zijn van een voortschrijdende atherosclerose. De bepaling van serum-cholesterol verschaft een mogelijkheid de vatbaarheid voor dergelijke ziekten te voorspellen. Daartoe dient de meting van die risicofactor of risico-indicator - het serum-cholesterol gehalte - betrouwbare uitkomsten op te leveren. Gezien de wijdverbreide belangstelling en de internationale communicatie tussen epidemiologen en andere onderzoekers is vergelijkbaarheid van methodieken en resultaten een vereiste. De veelheid aan analysetechnieken voor cholesterol en het ontbreken van een gemeenschappelijke basis voor precisie en juistheid hebben een algemene vergelijkbaarheid van meetresultaten veelal belemmerd.

Hoofdstuk 2 bespreekt literatuurgegevens over verschillende zaken die met het meten van cholesterol zowel als met de interpretatie van de uitkomsten te maken hebben. De nadruk hierbij ligt op de Liebermann-Burchard methode (de oudste en meest gebruikte). De enzymatische techniek (de nieuwste) wordt eveneens vermeld.

Voor het interpreteren van cholesterol waarden is kennis nodig omtrent biologische variabiliteit. De eisen welke aan de analytische precisie en juistheid gesteld dienen te worden hangen daarmede ten nauwste samen. Variaties in en tussen individuen moeten worden onderscheiden. De wijze waarop voldoende bekwaamheid in het laboratorium kan worden verkregen, middels kwaliteitsbewaking en met gebruikmaking van referentiemethoden en -materialen, wordt besproken. Tevens wordt aangetoond dat het voldoen aan de WHO eisen m.b.t. de juistheid niet betekent dat laboratoriumresultaten onbeperkt kunnen worden geacht gelijk te zijn. Tevens wordt beschreven dat sinds kort een Nederlands referentielaboratorium voor cholesterol in oprichting is.

Hoofdstuk 3 behandelt de internationale standaardisatie van twee laboratoria en een referentielaboratorium. In de KRIS werd drie jaar lang cholesterol en glucose in plasma bepaald op zodanige wijze dat er slechts 2% verschil bestond tussen Rotterdam en Kaunas, Litouwen, USSR. Bovendien hadden de glucosewaarden praktisch geen systematische afwijking, terwijl die van cholesterol een constante positieve onjuistheid hadden. De precisie werd nauwlettend bewaakt en de VC binnen vooraf vastgelegde grenzen gehouden. De gebleken kwaliteit was het resultaat van het strikt aanhouden van het laboratorium voorschrift (volledig in de Appendix weergegeven), de intensieve communicatie met het referentielaboratorium (CDC) en de, zo vaak als mogelijk was, onderhouden contacten met elkaar.

Hoofdstuk 4 is in het bijzonder gewijd aan intralaboratorium kwaliteitsbewaking zoals die is gegroeid in het CKCL tijdens het verrichten van de bepalingen voor het CB Hart Project. De belangrijkste onderdelen van het systeem zijn: de procesbewaking, de 'blinde' externe controle en het standaardisatieprogramma van het WHO Regionale Europeese Referentielaboratorium in Praag (IKEM). De procesbewaking stelde ons in staat kleine fluctuaties welke over lange tijdvakken in het laboratorium optraden vast te leggen en te gebruiken om seizoensinvloeden daarvoor te corrigeren. De blinde controles vormen het meest betrouwbare middel waarlangs dupliceerbaarheid en reproduceerbaarheid worden geverifieerd. Onze VC is 2 - 4% en de hogere waarden, die af en toe voorkomen zijn i.h.a. een gevolg van sporadische uitbijters. Een ernstige complicatie bij de standaardisatie vormt de matrix van de erbij in gebruik zijnde sera, waardoor systematische afwijkingen niet dan met veel moeite kunnen worden vastgesteld.

Hoofdstuk 5 beschrijft het standaardisatieprogramma van de Stichting KCA, dat in Nederland in gang is gezet om de vele laboratoriumresultaten een basis voor vergelijkbaarheid te verschaffen die tot dusverre ontbrak. Diverse epidemiologische onderzoekingen en veel periodiek geneeskundig onderzoek geschiedt vooralsnog met niet gestandaardiseerde laboratoria.

De evaluatiefase (deel I) en de eerste bevindigen van de standaardisatiefase (deel II van het programma) worden beschreven. Een aantal praktijkproblemen, veelal betreffende het serummateriaal, moest worden opgelost. Bij de deelnemers bestaat nogal eens gebrek aan voldoende intralaboratorium kwaliteitsbewaking.

Wij constateerden dat de onderlinge vergelijkbaarheid bij gebruik van calibratieserum sterk verbetert in de 'Huanggroep'. Voor de gebruikers van enzymatische methodieken lijkt calibratieserum niet direct nodig te zijn. Humaan materiaal is voor controle en eventueel voor bevordering van vergelijkbaarheid het meest geschikte materiaal. Onderzoek in deze richting is nog gaande.

Hoofdstuk 6 sluit de presentatie van de diverse resultaten af met een bespre-

king. In het algemeen zijn klinisch chemici nog onvoldoende met standaardisatie bezig en voeren zij nog steeds vele verschillende nieuwe technieken in. Zo worden dan ook vele HDL-bepalingen uitgevoerd waarvan de uitkomsten in onvoldoende mate zijn te interpreteren door de geïnteresseerde lezer van publikaties erover.

Voor een beter begrip van de problematiek is allereerst een éénduidige terminologie nodig. Een organisatiestructuur wordt vervolgens voorgesteld voor de kwaliteitsbewaking in het laboratorium.

Materialen voor kwaliteitsbewaking zoals die veel in de handel zijn kunnen voor controle op de precisie gebruikt worden. Om de systematische afwijking van een meettechniek te bepalen zijn ze minder geschikt. Zelfs de sera die bij het WHO referentielaboratorium te Praag in gebruik zijn, behoeven op dit punt verbetering. De calibratiesera die bij het RIV voor de Stichting KCA werden bereid, kunnen wellicht in de naaste toekomst de resultaten van Nederlands bevolkingsonderzoek in goede overeenstemming brengen. De analytische precisie behoeft niet zeer goed te zijn om gemiddelde veranderingen in het serum cholesterol gehalte te bestuderen, daar de biologische variaties relatief groot zijn. Als het laboratorium op langere termijn geen 'drift' vertoont, zijn de gemiddelde steekproefresultaten goed te beoordelen.

Echter bij de aanvang van een onderzoek (bijv. bij 'baseline screening') moet de precisie hoog zijn, omdat anders heel wat individuen onjuist worden ingedeeld in diverse risicogroepen.

Seizoenvariaties blijken in de m.b.v. kwaliteitscontrole gegevens gecorrigeerde resultaten afkomstig uit het CB Hart Project niet op een jaarlijks terugkerende, voorspelbare wijze voor te komen. Aangezien het er bij diverse studies veelal om gaat individuen in te delen in risicogroepen, zal precisie daar aan hoge eisen dienen te voldoen. Wanneer dit het geval is, zal vaststelling van de juistheid weinig extra moeite behoeven te kosten. Hoewel voor dit doel nu nog onvoldoende referentieserum voorhanden is, behoeft dit voor het relatief kleine aantal laboratoria dat bij bevolkingsonderzoek is betrokken geen blijvend obstakel te vormen. Ongetwijfeld zal de selectie van laboratoria die deelnemen aan epidemiologisch onderzoek steeds meer geschieden op geleide van op voorhand aangetoonde kwaliteiten, middels een certificaat van een erkend standaardisatieprogramma.

De Appendix is een onverkorte weergave van het Laboratorium Protocol van de KRIS, waarmee een unieke internationale standaardisering werd bereikt. Het heeft zijn ontstaan grotendeels te danken aan de ervaring van Dr. A. Mather van het CDC, die de beide KRIS laboratoria adviseerde gedurende de officiële studieperiode.

CURRICULUM VITAE

De schrijver van dit proefschrift is geboren op 28 maart 1944 in Pieterburen (Gr.), doorliep aldaar de Openbare Lagere School en bezocht de Rijks HBS (b) te Warffum (1956-1961). Na een verblijf van een jaar aan de École d'Humanité (Prof. Paulus Geheeb) Goldern, Zwitserland, werd de studie scheikunde begonnen aan de Rijksuniversiteit van Groningen.

Het doctoraalexamen is in 1971 afgelegd met als hoofdrichting organische chemie (Prof. Dr. J. Strating, Prof. Dr. H. Wijnberg en Dr. A. M. van Leusen) en bijvak polymeerchemie (Prof. Dr. G. Challa). Gedurende enkele jaren werd een assistentschap vervuld en voorts is voldaan aan de eisen voor onderwijsbevoegdheid.

Na het vervullen van de militaire dienstplicht, waarvan 16 maanden in het Klinisch Chemisch Laboratorium (Kol. Drs. P. J. Prosée) van het Militair Hospitaal 'Dr. A. Mathijsen' te Utrecht, werd de opleiding tot klinisch chemicus begonnen in het Centraal Klinisch Chemisch Laboratorium (Prof. Dr. B. Leijnse) van het Academisch Ziekenhuis Rotterdam - Dijkzigt. Inschrijving in het Register van erkend klinisch chemici volgde op 1 april 1977.

De schrijver werd later, met ingang van dezelfde datum, benoemd tot klinisch chemicus. Het proefschrift is gedurende de afgelopen jaren in het CKCL bewerkt, waarbij veel gastvrijheid werd genoten in de afdeling Chemische Pathologie der Medische Faculteit.

APPENDIX I

WHO/Kaunas/Rotterdam

BEHAVIOURAL AND OPERATIONAL COMPONENTS
OF HEALTH INTERVENTION PROGRAMMES

Laboratory Protocol

CONTENTS

The KRIS began when only a typewritten version of this protocol was available. This is the 1973, modified version, which is reproduced here by permission of the WHO.

The KRIS received an important grant from the Fannie E. Rippel Foundation.

1. Units of reporting

In Western European countries a new convention of reporting units of biochemical concentration, based upon Système Internationale, is being promoted; in the Netherlands, this system became official on 1 January 1971. For purposes of documentation and future referral, all values derived for the study (cut-off or screening levels, means, variance, etc.) should be available in new as well as old units. This is a simple matter of conversion from the older system by the use of standard coefficients, and this computation could be performed in all computer print-outs, either alone or in parallel equivalents of old and new units. It was felt, however, that since: (a) the older mass-based units are familiar and will be continued in Kaunas as well as in the United States of America and other places for some time; (b) they are easier to handle within and between the two laboratories in the study, particularly with unfamiliar methodologies and methods of calibration; (c) they are less susceptible to amputation and weighing erros; and (d) they are better adapted to the problem of rapid standardization of the two laboratories; the older system should be the "working" system for the study. Moreover, since past guidelines for epidemiologic and clinical evaluation are in the old units interpretation of the data by reference to past experience will be facilitated during the difficult period of transition which may extend over several years.

Specifically, mg/100 ml concentrations will be employed for the preparation of calibration standards, control and evaluation samples, and reporting results. Screening limits for glucose and cholesterol concentrations will also be based upon these units.

2. Common sources and specifications of equipment and materials

In view of the lack of comparative information on specifications and nature of materials commonly employed in the two regions, it was considered essential to acceptable standardization that those supplies considered critical be replicated in as close counterpart as possible in the two laboratories or furnished from common sources. This included several items of equipment selected for ease and speed of analysis that either had no corresponding counterpart or that might provide unique basic data that could not be strictly compared, even though the derived quantitative results might not be impaired. For example, spectrophotometers with flow-through cuvets furnish instrument readings that are not identical with specific optical absorbance, and cannot be comparable unless the dimensions and geometry of the cuvet system in each instrument are identical; reagent dispensers of different manufacture may deliver different volumes at the same nominal setting. For control purposes, and to provide sensitive means of diagnosis of deviations when problems of interlaboratory variance occurred, those variables that were considered amenable to such standardization were covered by common equipment. These are listed in Appendix I (page 9).

To provide the necessary flexibility for meeting unexpected situations and to establish a readily available source of common lot or standard specification reagents, procurement and provision of any necessary central supply was supervised by the Rotterdam laboratory. In certain cases, specifications had to be worked out by pre-study evaluation and comparisons.

3. Collection and handling of specimens for analysis

(a) Examinee preparation for laboratory work: This is an extremely important point of standardization in any collaborative study, and often the problems of obtaining sufficient numbers, or adequate cooperation, of subjects override the desirability or even necessity of standardizing patient preparation or physiologic and diurnal variables. Thus circulating glucose levels are quite sensitive (particularly in subjects with borderline carbohydrate metabolic competence) to the immediate dietary history and probably other diurnal factors. Since a single blood specimen was to be collected after a standard glucose challenge, this provided a more standardized intake, but necessarily widened the screening limits, owing to the increased variability of response among individuals along the post-challenge time curve.

Of necessity, single point glucose values after a standardized challenge decrease the sensitivity (i.e. require wider permissible limits) of glucose analysis as a screen for diabetic tendency. Those examinations seeking to define the circulating maximum - which may occur between 0.75 and 1.5 hours post-challenge - ordinarily select a one-hour time interval; those seeking to demonstrate tolerance of the subject by a return to normal levels by a single determination at two hours post-challenge encounter similar variance in the metabolic time curves of "normal" subjects. Although little hard data exist in the literature to support their observations, some investigators believe that the tolerance of many subjects to glucose loading is decreased in the afternoon in comparison with morning examinations. On the other hand, circulating levels of cholesterol in serum or plasma are relatively stable with respect to immediate dietary history. Protocol for standardizing the preparation of the subjects and times of collection has been detailed elsewhere in the study protocol.

(b) Collection of specimens: Specimens of venous blood are collected by venipuncture employing disposable 10 ml vacuum-collecting tubes (Vacutainers) with sterile disposable needles. Normally, serum is the specimen of choice for cholesterol assay and would be acceptable for prompt analysis of glucose also, but the stability of the latter constituent depends vitally upon maintenance of sterility, and in most population surveys flexibility is required in the times and procedures of specimen processing and in scheduling of analysis that make it advisable to provide a preservative for this constituent for some days. The present study presented an additional stability requirement in the need for interlaboratory comparisons on identical samples, and periodic shipment of samples from Kaunas to Rotterdam require long and uncertain transport times. Preservation of serum glucose with sodium fluoride is the most certain as well as benign procedure, as determined by other similar studies (see the recommendations of the American Diabetes Association, J.A.M.A., 18, 299 (1969)).

Thus the alternatives for this study were: (a) collection of two blood specimens, one clotted and the other preserved; (b) splitting a serum specimen and preserving half of it for glucose assay; or (c) performing both cholesterol and glucose analyses on preserved plasma. The latter course was selected in the expectation that plasma values for each constituent would not differ significantly from those with serum, and the use of an EDTA-fluoride combination preservative (supplied in the Vacutainer) was recommended. However, this was available from the supplier of Vacutainers only on special order and since procurement and supply proved to be a difficult problem, a more common mixture of potassium oxalate and sodium fluoride was selected.

(c) Collection of specimens for other analyses: For collection of specimens for analyses requiring serum (potassium, triglycerides or total serum lipids, and various enzyme analyses contemplated by the two study sites), another Vacutainer of blood may be collected following the first. The protocol for Vacutainer collection, processing, and distribution is detailed in Appendix II (page 11).

4. Selection and detailing of analytical methods

The selection of basic analytical methods was guided by the following considerations:

(a) Highly-mechanized or automated testing presented too great a problem of maintenance and control of the equipment and supply of expendable materials, and would not be essential for the estimated workload of 40-60 specimens per day. The only automated analytical equipment seriously considered was Technicon's AutoAnalyzer-I, since experience with other possible systems is extremely limited and conflicting. As demonstrated by the experience of several studies in the United States of America and in Britain, although internal control and precision can usually be improved by automation over the use of manual methods, the maintenance of close interlaboratory comparability of analytical results is quite difficult and requires very close external monitoring and ready and frequent communication, often involving travel of monitoring personnel or analysts between laboratories. Indeed, these difficulties in maintaining comparability of lipid analyses have only recently become the subject of a long-term study sponsored

and directed by the National Institute of Health in the United States of America. In view of the distances and transit times involved in the Kaunas-Rotterdam studies, the problems of both materials supply and diagnosing and correcting deviations in results in one or the other laboratory, the difficulties with automated methods and equipment appeared to be insurmountable

(b) Similarly, complex reference manual methods are difficult to maintain in control and are highly susceptible to error in the routine laboratory, although reliable reference methods were expected to provide the target values at which the finally selected methods should aim.

(c) In spite of the dangers inherent in oversimplification of chemical analytical procedures, simplicity is a very desirable goal under the circumstances of the study: English text procedures being official for the study, translation into two different local languages is necessary; adaptation to local or regional technology must be made in one or the other laboratory; standardization of each step in detail being necessary, the number of such steps must be kept at a minimum, etc. Moreover, the availability and supply of reagents of constant quality, simplicity and reproducibility of preparation, costs, the effects of turnover in personnel on the quality of analytical results, and other factors entered into the decision.

On this basis, two manual methods were proposed and accepted, both being conducted directly on serum samples without deproteinization or solvent extraction, and both having been extensively studied and evaluated as standard procedures in a number of countries, thus having some standing as internationally accepted methods. Intensive preliminary study before the recruitment of patients for the study began revealed serious shortcomings of both methods in the ease of standardization between laboratories, both being quite sensitive to very slight variations in analytical technique. In retrospect, however, no other candidate methods that might meet the stated requirements appeared to be freer of problems, and the extensive support of WHO in ensuring that a temporary advisor and the key laboratory professional personnel of each study site spent a week of intensive study and observation at each laboratory site laid the groundwork for ironing out future problems by written communication.

Glucose

Plasma or serum specimens have by now almost universally replaced the use of whole blood samples. A method approximating "true glucose" values was desirable, although analytical results in many large population studies are being based upon the ferricyanide AutoAnalyzer method, which yields values somewhat higher. Enzymatic methods are generally considered to be most specific for glucose, but their reliance upon expensive and poorly-defined commercial enzyme products, often sensitive to deterioration and presenting problems of procurement of uniform quality, weighed against their choice. Ortho-toluidine has been widely employed as a reagent relatively specific for hexoses, and in practice yields values for human serum specimens close to those with enzymatic methods. Originally this reagent was applied to protein-free filtrates of plasma or serum, but a simplified procedure involving heating a small aliquot of sample directly with the reagent yields comparable values, and interference with non-glucose constituents is quite infrequent.

The present method has been extensively studied in parallel with other "true glucose" methods by the Reference Laboratory of the Center for Disease Control, Atlanta, which is collaborating with other WHO standardization programmes.

Cholesterol

As with glucose, analytical results for serum cholesterol levels vary rather widely with the specific methodology employed, and again a simplified method was sought that promised results not too deviant from those with much more complex reference methods. A procedure listed in the N.N.I. (No. 2415) manual of recommended methods, derived from that of Huang, et al., was selected. Intensive study revealed that the analysis was critically sensitive to

the manner in which each step of the test was performed, and the final method was determined by the techniques that gave results with plasma samples that approximated Abell reference values and gave the best reproducibility.

The critical problem with the method is that of calibration, i.e. obtaining development of colour with calibration standards that exactly parallel concentrations in the plasma specimens. Since the reaction of the water in the specimen produces a great amount of heat that cannot be replicated in pure cholesterol standards prepared in organic solvents, the colour development of the former proceeds more rapidly than in the latter. Thus, an arbitrary set of conditions must be found in which these coincide. An alternative approach would be the use of analysed serum pools as calibrators, but since this involves logistics problems in providing stable, carefully analysed materials to both laboratories over the term of the studies, it was considered only as a last resort if other approaches failed.

Thus the chosen method is not one that can be applied to similar studies without expenditure of effort comparable to that taken by this project to ensure interlaboratory consistency on a long-term basis.

The analytical methods are detailed in Appendix III (page 13).

5. Standard procedures for calibration and computation

Three major points were proposed and adopted by both laboratories.

(a) Standard materials

Pure standard solutions of each material were adopted as the basis of calibration. This decision avoids the extremely troublesome problem of maintaining and characterizing serum calibration "standards", which has been the source of many difficulties in international standardization programmes. It was recognized, however, that technical differences between the two laboratories in relating the standards to the specimens might persist, leading to an unacceptable bias between their results that might be resolved only with the use of a stable serum calibration standard; in this eventuality, a common source of such calibration materials would have to be provided for both laboratories. If such differences could not be resolved in the preliminary trials, a mechanism for obtaining a common basis for the laboratory data was provided whereby analysis of all of the early specimens of the project would be analysed in one laboratory until these differences were resolved.

(b) Linearity of response

Linearity of the analytical response with concentration of the analysed substance is confirmed daily by the employment of standards at several calibration levels with each analytical run. Where problems of non-linearity are encountered, a more intensive investigation of the chemistry of the analysis is provided before proceeding. A standardized computation procedure is based upon the values of the multiple standards.

(c) Recovery of the data

Standard forms for recording O.D. readings and the computed concentration values for each run are to be used, copies of which are to be kept on file at Geneva. This provides a means of validation or correction of sporadic computational errors on subjects in which the laboratory data may be questioned.

Details of the computations, the evaluation of acceptability, and the procedures to be followed in rejecting analytical results are given in the detailed procedures in Appendix III (page 13). The rejection limit of three per cent. for variance in replicate results or in determining the calibration was based upon preliminary study of the achievable results in

both Rotterdam and Kaunas laboratories. It was estimated that these limits might result in as much as 10 per cent. repetitions of analyses during the early stages of the study but should steadily decrease as proficiency improved with practice.

6. Quality control programme

The quality control programme of the project consists of three chief categories: (i) local measures for refinement of reported values by means of replication of results and procedures for handling deviations beyond acceptable limits; (ii) local surveillance of quality and benchmark calibration by the laboratory supervisor, with provision for his response to detected problems; and (iii) external surveillance by the project coordination team, with provision for intervention and aid if and when precision in either laboratory becomes unacceptable or there is evidence of deviation from comparability of the results or from benchmark accuracy. Definitions of terms, detailed instructions, and reporting form formats are given in Appendix IV (page 19).

Control procedures incorporated into routine analyses

(a) Duplicate analyses of specimens: The precision and reliability of any determined value is increased with replicate determinations, since each determined value incorporates some analytical variance which tends to be averaged out with multiple determinations. If these replications are spread over separate analytical runs or days of analysis, an additional "between-run" variance is encountered but the averaged value should be more generally representative of the determined variable as measured in that laboratory; moreover, independent scheduling of replicate analyses for each specimen provides a means of monitoring sporadic errors (as distinguished from random variance), both in analysis and in specimen identity since specimen mix-ups increase significantly in the processing of larger workloads. However, the latter approach consumes considerably more analytical and data processing time and encounters problems of preservation and storage of all specimens.

Each subject specimen, as well as calibrators and controls, will therefore be analysed in side-by-side duplicate for as long as is deemed necessary by cost-yield considerations by the project coordinating team; the average value is used for patient evaluation, but both results are recorded on the central laboratory reporting form for study purposes. Specimens in which the duplicate results vary by more than three per cent. of the higher determined value in that pair are held over (refrigerated) and reanalysed in the following run on the same or the following day; in this case the second pair of results are reported as patient data, and the first pair of results are cancelled. The laboratory supervisor must be notified of all such unacceptable deviations, but may delegate the rescheduling and reanalysis as a routine procedure.

The use of multiple concentration levels of the calibration standards provides adequate control for handling excessive variance between duplicate standards, as detailed in the analytical protocols.

(b) Optical density readings of standards: A record is kept of the optical density readings of the standards for each run to monitor chemical stability of the analysis; a consistent drift from run to run is evidence of a change in the chemistry, and possible deterioration of either the standards or reagents. Such changes are most easily followed in the laboratory by plotting the average computation factor "F", which will vary inversely with the drift of the O.D. of the standards; these values are also stored in the central computer file of the control data.

(c) Local bench control (serum controls): to provide the analyst and laboratory supervisor with a daily check of the integrity of the analysis, the usual system of introducing serum control materials at two concentration levels (normal and elevated values) into each run is employed. The results of the controls and the linearity of the concentration curve with the

calibrating standards are examined to confirm control at the beginning of each day; a graph plot of daily results of each control serum is maintained in the laboratory, and results are transmitted and filed in the central records of the study. Preparation of control charts is considered in Appendix IV. In the initial stages of the study these pools were prepared in each laboratory; later they were replaced by sera of normal and elevated concentrations each supplied in a single "common lot" or production run of commercial lyophilized human serum.

(d) Common interlaboratory control pool (common control): a common, stable pool of analysed serum that is analysed in each analytical run provides a common ratio between the results of the two laboratories. Normally this ratio should be 1.00, but methodologic and equipment differences may show some consistent difference between these results. If such difference exists, and if external evaluation confirms a consistent bias between laboratory results, this ratio can, if desired, be used to rationalize ("correct") results of one laboratory to the other, or to a third control laboratory performing the external evaluation. It was intended that a commercially available lyophilized serum preparation be provided in sufficient quantity to last the entire study in each laboratory as a "common control".

However, procurement difficulties and problems of selection of suitable materials delayed the introduction of common controls until after the patient recruitment began in both study sites. It was then decided to employ common-lot lyophilized materials at two laboratory-prepared controls. This decision provided a source of two bases of interlaboratory comparison on common materials that can be followed throughout the remainder of the study.

Quality surveillance by each laboratory director

The laboratory director in each study site will maintain the following measures:

(i) Daily review of analytical data:

acceptability of the linearity and computation factor for each run;

acceptability of the serum control and common control results;

the quality control plots of the absorbance of standard readings;

normal and elevated control pools, and the common control;

acceptability of duplicate readings (decisions as to disposition of variance beyond limits).

(ii) Introduction into each analytical run of previously analysed specimens from the previous analytical day - selected by, and known only to him ("blind replicates"). He evaluates the degree of consistency with previous results and takes appropriate measures for repeating analyses when this is necessary.

(iii) Review of all reports and data go into the local study files and to Geneva.

Blind replicates

In order to ascertain the level of reproducibility and consistency between results on different analytical days, a systematic programme of reintroducing blind duplicate samples of examinee specimens into subsequent runs is maintained by the director. Three such quality assessment samples are analysed each analytical day. The results of this control measure are forwarded to Geneva in a form worked out with SHS, involving some estimate of the standard error.

External surveillance and proficiency testing

Periodically, at about two-month intervals, a set of 12 unknown specimens will be sent out from the Atlanta WHO Standardization Group for the evaluation of the current proficiency of the two laboratories. The performance evaluation will serve two functions: recognition and alert of analytical difficulties in either laboratory, and a means of comparison and correction between analytical levels that are not mutually consistent. For each survey (one shipment), samples will be provided for inclusion into two separate runs, with means for testing within-run and between-run variance and the current level of bias of either cholesterol or glucose analysis from a reference basis. Both known and blind replicates will be included by a system of internally related pools. Two sets of six samples each will be included in each shipment: the six vials in one set are analysed (in duplo, as indicated on the report form); the second set of six vials is similarly analysed one week later. All results are recorded on a standard form supplied with each shipment and promptly returned to Atlanta. Evaluations of the survey will be sent to SHS and a summary of performance (blinded as to individual samples) is sent to each laboratory. When a serious problem is revealed, the laboratory will be notified in detail in an attempt to remedy the problem.

Mechanisms for response to detected problems will include submission of common standard solutions if calibration appears to be a problem, written consultation, and (probably) reduction of analytical data to a common basis. In extreme cases, personnel exchange between the two laboratories, either as consultation or retaining, may be arranged to solve difficult problems. The results of the surveys will be logged into the "quality control" file for Geneva. A sample format for the receipt of this information is attached.

Base comparability by analysis in one laboratory

Although two-way exchange of the specimens of one study site with the other was considered, considerations of the difficulties of transport of specimens and the effect of possible double workloads dictated that one-way exchange of the early specimens of the Kaunas study to be analysed in Rotterdam be instituted; Rotterdam patient recruitment begins somewhat later than in Kaunas, and this would permit concentration on problems that may arise. If the results of the two laboratories diverge, methodologic variance can be investigated aimed toward better coincidence of the results; if there still remain discrepancies, the chief contribution of this exchange will be in validating constant bias between the two laboratories discovered by the external surveillance programme, so that the results may be rationalized.

APPENDIX I

LIST AND SPECIFICATIONS OF COMMON EQUIPMENT AND MATERIALS

Common-lot control sera

(1) Normal range: Metrix Normal, Armour (Chicago, USA)

(2) To be specified later.

Reagents

From: Merck, Darmstadt, Germany

Quality: pro Analyse

Reagent	Merck Cat. No.
Propanol-(2)	9 634
Sulfuric acid	731
Acetic acid	90 063
Acetic anhydride	42
Benzoic acid	136
Anhydrous sodium sulfate	6 649
d-Glucose	8 337
Thiourea	7 949

From British Drug Houses Ltd.

O-Toluidine
Cholesterol

Equipment in each laboratory

(1) Photometer

From Meyvis and Co., Pergolesilaan 4, Bergen op Zoom, Holland

Vitatron Colorimeter	UC 200-S
Flow-through cuvet in holder light path 10 mm; minimal volume 1.5 ml	UC 204/02
Three-way stop cock for flow through cuvet	UC 205/03
Suction unit + pump + reservoir + tubing	UC 205
Interference filter 618 nm	
Interference filter 623 nm	
Thermostated cuvet holder	
5 x cuvets	UC 207/02
English manual	
2 spare lamps	

Appendix I

(2) Reagent dispensers (four each)

From Salm & Kipp, Keizersgrach 642-644, Amsterdam, Holland

Citopipette, complete with uncoloured bottle for delivery of maximal 10 ml, D.B.G.M. 1942737, type 6101

(3) Mixing apparatus

2 x Vortex Genie Mixer for 220 V

(4) Water-bath 100°

G.F.L. water-bath, type 1013
8 x reagent tube racks type 1920

(5) Thermostat

Thermomix II (manufacturer B. Braun Melsungen), Cat. No. 51012
Contact thermometer 0-100°, Cat. No. 52270

(6) Balances

Mettler H 10 TW digital analytical balance,
P 1200 precision balance

(7) Deep freezer

Bosch GS 230

(8) Micro-pipets

Glucose pipets according to MacLean - fabricated by Haak
200 x 0.2 ml
200 x 0.05 ml

APPENDIX II

COLLECTION OF SPECIMENS

1. Specimens of venous blood are collected by venipuncture employing disposable 10 ml vacuum-collecting tubes (Vacutainers) with sterile disposable needles.

2. The sample is collected into a premeasured and dried quantity of anticoagulant (supplied in the Vacutainer) containing 25 mg of sodium fluoride and 20 mg of potassium oxalate. This permits the storage and transport of the specimen before centrifugation for six to eight hours if absolutely necessary, but the plasma should be separated as promptly as possible. After centrifugation the separated plasma may be stored in the refrigerator overnight or frozen for several months without loss of glucose.

3. For collection of specimens for other analyses employing serum (such as potassium, triglycerides, or total serum lipids, and various enzymes), another Vacutainer of blood may be collected following the first; this tube will be identified by a red stopper. The procedure is as follows:

(a) The first blood sample is drawn into a Vacutainer (containing fluoride-oxalate in the standard manner, allowing it to fill completely. The tube is then removed from the holder without disturbing the location of the needle in the vein; the sleeve of elastic tubing will cover the outlet of the needle to prevent loss of blood from the vein.

(b) For studies involving analysis for serum constituents (enzymes, lipids) a second red stopper Vacutainer (containing no preservative) is carefully inserted into the holder and pressed firmly so that the needle outlet penetrates the stopper of the Vacutainer. As before, the Vacutainer is allowed to fill completely. (If only glucose and cholesterol are to be analysed, the second Vacutainer is not collected.)

(c) The needle is smoothly removed from the vein with the tube and holder still attached; a sterile cotton sponge is pressed over the puncture site.

(d) The fluoride oxalate Vacutainer is mixed by inversion six times; it may be centrifuged at any time up to six hours. The second (serum) tube is left to stand at room temperature for 30-45 minutes to permit complete clotting before centrifugation. Before centrifuging a clotted tube, the clot is "rimmed" by placing a clean applicator stick down one side of the tube between the clot and the glass and drawing it around the inside wall of the tube to free the clot from the glass.

4. Centrifugation: Each tube is properly labelled by number and carefully counterbalanced in a centrifuge shield or tube holder in which it is to be centrifuged, and the tubes and holders are placed in the centrifuge in balanced pairs. The tubes are centrifuged at 1000 x G for 10 minutes.

5. Separation and transfer: With a clean 1 ml Pasteur pipet fitted with a 1 ml rubber bulb, draw approximately 1 ml of clear plasma from the fluorided Vacutainer and transfer into each of several vials:

Vial No. 1: for analyses for glucose and cholesterol;

Vial No. 1a: (during the first part of the study an extra vial (No. 1a) will be prepared at Kaunas for later shipment to Rotterdam for the interlaboratory comparison of results. Vial No. 1a will be capped and kept frozen at -20° until analysis at Rotterdam);

Appendix II

Vial No. 2: for frozen storage at -20°. This is to be used for repeating analyses (i) when a discrepancy occurs between results between Rotterdam and Kaunus results, or (ii) when the follow-up work on a screened patient is not consistent with the original results and must be checked;

Vial No. 3: if an adequate volume of plasma is available a third vial is filled with the remainder of the plasma and stored in the refrigerator for the purpose of repeating a doubtful analysis or for introduction as a blind replicate in the following day's run; if not used by the second day, it is poured into a container for collecting plasma control serum and kept frozen.

All vials are tightly capped and labelled before storage.

APPENDIX III

A. DIRECT DETERMINATION OF GLUCOSE IN SERUM BY THE ORTHO-TOLUIDINE METHOD

Reagents and equipment are specified on a separate list, Appendix I (page 9).

Reagents

Ortho-Toluidine reagent (O-T reagent, six per cent. v/v)

1.5 g of thiourea are dissolved in 940 ml of glacial acetic acid in a two-litre Erlenmeyer flask; the flask is covered and placed in a heating bath and warmed to 50° with occasional swirling to permit complete solution. After cooling to room temperature 60 ml of ortho-toluidine are added from a graduate with vigorous swirling and the solution mixed thoroughly. The reagent is stored at room temperature in a dark bottle, tightly capped and protected from light. Stable for about three months.

Benzoic acid diluent

One litre is prepared by adding 2.0 g of benzoic acid to one litre of distilled water in a flask, and heating into solution (60-70°) until dissolved; this goes in slowly. The flask is capped and left to stand overnight in a refrigerator. The clear supernatant is decanted from the precipitated benzoic crystals into a stoppered storage bottle. It keeps indefinitely at room temperature and standards prepared with it should not precipitate in the refrigerator, although this is not critical.

Stock glucose standard (10.0 mg/ml)

2500 mg of anhydrous glucose are carefully weighed out and dissolved in about 200 ml of benzoic diluent in a 250 ml volumetric flask; after the glucose has completely dissolved, the contents are diluted to the mark at 20-23° with diluent. This stock is kept in tightly-capped bottles in the refrigerator and is stable for up to a year. For the preparation of working standards suitable portions are poured into clean flasks and allowed to come to 20-23° before being pipetted. Note that the solid Merck glucose must be kept tightly capped in a vacuum disiccator over anhydrous calcium chloride.

Working glucose standards

The 100 mg% standard is prepared by diluting 10.0 ml of stock 10 mg/ml glucose to 100 ml in a volumetric flask with benzoic diluent. Similarly, 200 mg% (20.0 ml to 100 ml) and 300 mg% (30.0 ml to 100 ml) standards are prepared. These standards must be kept refrigerated and tightly stoppered. Suitable portions are poured into clean tubes and warmed to 20-23° before pipetting. Fresh working standards are prepared monthly.

Analytical run

A single analytical run should consist of not more than 40 tubes, including standards, controls, and patient specimens. Each sample is set up in duplo and run side-by-side. Two stable control plasma or serum pools (one in the normal and one in the elevated concentration range) are set up and analysed as unknowns in each run.

1. Switch on the Vitatron colorimeter, select and insert a filter between 620 and 630 nm, and select the red phototube by switching to the red dot. Allow the instrument to warm up for about 30 minutes. In the warming up time, the following procedures can be carried out.

Appendix III

2. Label the following 17 x 150 mm reaction tubes for the run, in duplicate:

Blank
Standards: S_{100}, S_{200}, S_{300}
Patient specimens
Controls: normal, elevated
A final S_{200} standard

Place the tubes in a rack in serial order. A "Run List" is prepared on the analysis recording form (see Appendix IV, page 21) with each standard or sample in this same numbered order. Dispense 5.0 ml of o-toluidine reagent into each tube from a Citopipette dispenser set at exactly 5.0 ml, with the reagent at 20°.

3. Use separate, clean micro-pipets for each measurement, carefully filled and adjusted to the mark and wiped clean before delivery.

To the blank tubes add 0.050 ml water; to the standard tubes add 0.050 ml of the respective working standards; to the control and patient tubes add 0.050 ml of each plasma specimen.

The micro-pipet is rinsed three times with the reagent and blown out. Immediately after each pipetting, the tube contents are mixed twice for three seconds each on a Vortex mixer.

4. When all tubes are prepared, the rack is placed in a vigorously boiling water bath for exactly 12 minutes; this period of heating is critical.

5. At the end of 12 minutes' heating, the rack is removed from the bath and placed in a cold water bath for 10 minutes. The colour is relatively stable after cooling, but the tubes should not be allowed to stand for variable periods of time before photometry. The final temperature of reading should be stabilized at the temperature of the room.

6. Before beginning the readings, adjust the Vitatron to zero optical density (O.D.) with o-toluidine reagent; remove this by suction. Final adjustment of zero is made with the first blank, and the O.D. of standards and samples are read and recorded on the form in duplicate. The cuvet is rinsed with a little o-toluidine reagent, and the second blank is poured in and read as a final check on possible instrument drift.

Notes

1. Rinsing of the flow-through cuvet is performed with half the 5 ml volume, insuring that the funnel and cuvet are washed down; this wash is removed as completely as possible with suction before the final portion is added. At least 2.0 ml is required for the final reading.

2. The measurement of 0.050 ml by micro-pipet is critical and may contribute to excessive variance if not carefully performed.

3. The delivery of the Citopipette must be adjusted to 5.0 $\pm$ 0.1 ml by dispensing 10 aliquots of o-toluidine reagent into a 50 ml volumetric flask; the quantity delivered must be within 1 ml of the 50 ml mark on the flask.

4. The blanking of the photometer is conducted similarly to the procedure given in (3) under the cholesterol method.

Computation of results

1. The computation of results is a comparison between average O.D. readings of each specimen and averaged readings of the standards, according to the general equation:

Appendix III

$$\text{Glucose concentration in specimen} = O.D._{spec} \times \frac{Conc_{std}}{O.D._{std}} = O.D._{spec} \times \bar{F}$$

2. Factor (F) is obtained by dividing the concentration of each standard in the run by its corresponding O.D. reading. The calculated factor for each duplicate standard at any one concentration should not vary by more than three per cent. If this holds true, an average factor at each standard concentration (100 mg%, 200 mg%, 300 mg%, and the final 200 mg%) is obtained. If one set of duplicates varies by more than three per cent., then the factor for the result which lies farthest from the average factors of the remaining three standard duplicates is discarded (see instructions in Appendix V, page 25).

3. Overall computational factor ($\bar{F}$): The spread of the average factors at each standard concentration level, corrected for any outlier, should not exceed three per cent., in which case all of the accepted factors are averaged to obtain factor $\bar{F}$, for computation of results in the above equation.

4. Using this average $\bar{F}$, compute the average concentration values of all samples, both standards and specimens, from the average O.D. value for each, and record in the proper column on the reporting form. Duplicate values for each specimen should not vary by more than three per cent. from each other. Samples exceeding this allowable variance are re-analysed in following runs; if this reanalysis must be delayed to a following day, the suspect samples is frozen until such analysis.

5. For concentrations of specimens appreciably exceeding 300 mg%, adequate answers may be obtained by diluting one volume of the coloured product of the specimen with one volume of the o-toluidine reagent and reading this mixture at the end of the run; this result is multiplied by 2.0 to obtain final concentration. In this way no dilution and resetting of the reagent blank is necessary.

B. DIRECT SERUM METHOD FOR CHOLESTEROL ANALYSIS

Source reagents and equipment are specified on a separate list, Appendix I.

Reagents

Lieberman-Burchard reagent

1. Measure by graduate cylinder 300 ml of glacial acetic acid into a two-litre Erlenmeyer flask From a graduate, add portionwise and with constant mixing 600 ml of acetic anhydride. The flask is left to chill to 5° or below in an ice-bath.

2. 100 ml of concentrated sulfuric acid are measured in a 100 ml graduate cylinder and also chilled in ice.

3. With constant stirring or swirling of the large flask in the ice-bath, pour the sulfuric acid in portions of about 10 ml at a time into the acetic acid-anhydride mixture, so that the temperature does not rise above 20°.

4. After five minutes standing in ice, the flask is removed, 20 g of anhydrous sodium sulfate are added, the flask is covered with a glass plate, and the mixture is left to stand at room temperature with occasional shaking to dissolve the salt, which may require some hours. When solution is complete, store in a screw-capped bottle or reagent dispenser in the refrigerator. The reagent is stable for several weeks; if a faint yellow colour appears, it is discarded. The volume required for each day's work must be brought to 20-23° before dispensing the 5.0 ml aliquots.

Appendix III

Stock cholesterol standard (1000 mg/100 ml)

Weigh 2500 mg ± 2 mg of cholesterol standard in a preweighed weighing trough. Holding the trough with clean forceps, rinse the powder into a funnel in a 250 ml volumetric flask five times with 5 ml (approximately) volumes of isopropanol. Wash down funnel with about 100 ml of isopropanol; stopper the flask and place in a 37° water-bath; swirl occasionally till cholesterol is completely dissolved. Transfer the flask to a bath held at 20° ± 2° and dilute to the 250 ml mark with isopropanol. Stopper and invert flask 10 times for complete mixing of the standard.

This concentrated standard is stable if kept tightly stoppered in a refrigerator (5°-8°). Note that any evaporation of solvent will alter the concentration of this stock and this must be avoided.

Working cholesterol standards (prepared each month)

Note: Standard concentrations at two levels (200 and 400 mg/100 ml) will be employed routinely. However, during the preparatory studies of the analysis, and at any time thereafter when problems of non-linearity arise four levels should be employed: 100, 200, 300 and 400 mg/100 ml. The two additional standards are to be prepared by dilution from the S_{200} and S_{400} working standards. All solutions are to be brought to 20° ± 2° before any samples are removed or dilutions made.

S_{400} (400 mg/100 ml standard): 20.0 ml of Stock 1000 mg/100 ml standard are measured from a 20 ml volumetric pipet into a 50 ml volumetric flask. The contents of the flask are diluted to 50.0 ml with isopropanol at 20° ± 2° and the flask is stoppered tightly and inverted 10 times.

S_{200} (200 mg/100 ml standard): From the same 20 ml pipet 20.0 ml of Stock 1000 mg/100 ml standard are measured into a 100 ml volumetric flask and similarly diluted to mark and mixed.

Since these two standards are prepared each month, they need not be refrigerated but should be tightly capped and stored in a dark cabinet. Sufficient quantities of each are poured into clean tubes and brought to 20° ± 2°.

S_{100} (100 mg/100 ml standard): 2.00 ml of S_{200} plus 2.00 ml of isopropanol are pipetted from a 2.0 ml pipet into a clean test-tube, mixed and used immediately.

S_{300} (300 mg/100 ml standard): 2.00 ml of S_{200} and 2.00 ml of S_{400} are pipetted into a clean tube, mixed and used immediately.

Analytical run

A single analytical run should consist of not more than 40 tubes, and the precision and accuracy of the analysis depends vitally upon exact timing and replication of the pipetting and mixing steps.

1. Switch on the Vitatron Colorimeter for about 30 minutes before use. During the warming-up time, the following procedures are carried out:

2. Select and insert the proper filter (625 ± 5 nm), and select the proper (red) phototube by turning the switch to the red dot.

3. Label the following 17 x 150 mm reaction tubes for the run, in duplicate:

Blank
Standards: S_{200}, S_{400}
Patient specimens
Controls: normal, elevated
A final S_{200} standard

Place the tubes in a rack in serial order. A "Run List" is prepared on the analysis recording form (see Appendix IV, page 19) with each standard or sample in this same numbered order.

4. Dispense 5.0 ml of Lieberman-Burchard reagent into each tube from a Citopipette dispenser, set at exactly 5.0 ml with the reagent at 20°.

5. With the stop-watch or timer started and using a clean micro-pipet for each sample delivery, add 0.200 ml of sample (water for the blank, respective working standards, or patient plasma specimens) at a timed schedule of 30 seconds. The sample is carefully measured and adjusted, and the pipet tip wiped off; the pipet is inserted into the bottom of the tube under the reagent, and the sample is blown out forcibly from the pipet, followed by air for mixing. Rinse the micro-pipet quickly by drawing up and expelling reagent from the tube three times. Quickly mix the contents of the tube by Vortexing three times on the mixer, being careful to hold the tube close to the top; insert the tube in a rack in the 25° water-bath. Not more than 10 seconds must be consumed between the first delivery of sample and insertion of the well-mixed tube into the bath.

6. Each tube is read at 25 minutes from the addition of sample. Before beginning the readings, adjust the Vitatron to sero O.D. (optical density) with L-B reagent. If any water has been left in the cuvet, this must be carefully removed by suction to prevent damage to the cuvet by the heat of reaction with the L-B reagent.

7. At exactly 25 minutes from addition of the blank, read the first blank and adjust the colorimeter to zero O.D. if necessary. Now read the developed colours of all subsequent tubes on a 30-seconds time schedule, rinsing the cuvet carefully with half of the 5 ml reaction mixture, carefully removing the wash by suction, and inserting the last half of the sample. After the last sample, rinse the cuvet with L-B reagent and read the second blank as a check on instrument drift.

8. Record the O.D. reading for each tube as it is obtained on the reporting form.

Notes

1. The temperature of the reagent at the addition of sample should be at a constant temperature, which may not be the case if room temperature varies greatly from 20° and the tubes have stood for excessive times after delivery at 20°. They are, however, not placed in a constant temperature bath before the addition of sample, in order to prevent more serious errors from water dropping into the tubes during manipulations. The heat of reaction within the tube exceeds 65° with the serum samples, much lower with the isopropanol standards, so that exact timing and rapid processing in order to get each tube into the 25° bath within 10 seconds determines the precision and reproducibility of the final result.

2. The order of addition was determined by the use of micro-pipets, from which all of the sample must be rinsed out in the reagent to ensure complete delivery. Both the order and nature of addition of sample and reagent affect the relative responses between the isopropanol standards and serum or plasma samples. This has been extensively studied, and the final procedure (although not ideal) has been derived as producing the best controlled results.

Appendix III

3. Usually the O.D. of the blank is zero when compared with the initial zeroing of the colorimeter with L-B reagent; the advantage of the preliminary setting is the availability of adequate volumes for rinsing the cuvet free of water. If the blanc O.D. changes appreciably from L-B O.D., fresh reagent should be prepared. If the final blank reading varies appreciably from zero, its plus or minus deviation should be estimated, and a linear plot of the standard curve should be made to ascertain whether the first or second blank falls on the extrapolated curve. If the best straight line through the average O.D. readings of the three standard levels goes through the origin (the original blank zero), ignore the second blank. If the second blank reading is nearer the standard line and it intersects the concentration axis beyond $\pm$ 3 mg%, the instrument was not zeroed properly, and the majority of the results will exceed the three per cent. deviation limit. In this case, the second blank O.D. is subtracted from all standard and sample O.D. readings for the following computations.

4. The delivery of reagent from the citopipette must be adjusted to 5.0 $\pm$ 0.1 ml by checking volume delivery of 10 aliquots of L-B reagent into a 50 ml volumetric flask at 20° (see note 3 under glucose analysis).

Computation of results

A factor (F) is calculated for each standard by dividing the O.D. of each by its respective concentration in mg%:

$$F = \frac{\text{Concentration of the standard}}{\text{O.D. of that standard}}$$

If one of the factors calculated from the four S_{200} results differs by more than three per cent. from the three others, this one standard is omitted, and the mean value of the remaining S_{200} factors is obtained by averaging. This mean should not differ by more than three per cent. from the mean factor derived from the S_{400} standards, in which case all acceptable values are averaged for $\bar{F}$. (see Appendix V. page 25).

If the differences between S_{200} and S_{400} factors is greater than these limits, or more than one standard shows unacceptable deviation, a new experimental run of standards + controls only should be made with a complete series of standards, from S_{100} through S_{400}. If non-linearity persists, prepare new reagent and working standards, and compare new and old standards with the new reagent.

The concentrations of the controls and patient samples are calculated by multiplying the average O.D. of each sample by the average factor, $\bar{F}$. Also calculate and record computed values for the standards by this equation; this indicates the deviation of each standard from the mean factor.

APPENDIX IV

Cholesterol Analysis revised 12.2.1973

Cholesterol analytic run no. ____ (1-4)

Laboratory ---------- code ____ (5)* Date: day ____ (6-7) month ____ (8-9) year ____ (10-11)

Analyzed by -------- ____ (12) code Supervised by -------- ____ (13) code Calculation factor ____,____ (14-17)

	Sample Identification					Results			
	number	subject number	exam code	sample/ pool	test no	O.D.	O.D.	mg%	test acceptable yes=1;no=2
col.no.	18-19	20-23	24-25	26	27	28-30	31-33	34-36	37
		standard 200 mg%							
	02	9999	99	1	1				
		standard 400 mg%							
	04	9999	99	1	1				
	05			1	1				
	06			1	1				
	07			1	1				
	08			1	1				
	09			1	1				
	10			1	1				
	11			1	1				
	12			1	1				
	13			1	1				
	14			1	1				
	15			1	1				
	16			1	1				
	17			1	1				
	18			1	1				
	19			1	1				
	20			1	1				
	21			1	1				
	22			1	1				
	23			1	1				
	24			1	1				
	25			1	1				
	26			1	1				
	27			1	1				
	28			1	1				
	29			1	1				
		normal Metrix control							
	30	9998	99	1	1				
		elevated Metrix control							
	31	9999	99	1	1				
		standard 200 mg%							
	32	9999	99	1	2				

* code 1 = Kaunas: code 2 = Rotterdam

Appendix IV

Glucose Analysis

revised 12.2.73

Glucose analytic run no. 8 (1-4)

Laboratory ---------- code ___ (5)* Date: day ___ (6-7) month ___ (8-9) year ___ (10-11)

Analyzed by -------- ___ (12) code Supervised by -------- ___ (13) code Calculation factor ___,0 (14-17)

	number	subject number	exam code	sample/ pool	test no	O.D.	O.D.	mg%	test acceptable yes=1;no=2
col.no.	18-19	20-23	24-25	26	27	28-30	31-33	34-36	37
		standard 100 mg%							
	01	9991	99	1	1				
		standard 200 mg%							
	02	9992	99	1	1				
		standard 300 mg%							
	03	9993	99	1	1				
	05			1	1				
	06			1	1				
	07			1	1				
	08			1	1				
	09			1	1				
	10			1	1				
	11			1	1				
	12			1	1				
	13			1	1				
	14			1	1				
	15			1	1				
	16			1	1				
	17			1	1				
	18			1	1				
	19			1	1				
	20			1	1				
	21			1	1				
	22			1	1				
	23			1	1				
	24			1	1				
	25			1	1				
	26			1	1				
	27			1	1				
	28			1	1				
	29			1	1				
		normal Metrix control							
	30	9998	92	1	1				
		elevated Metrix control							
	31	9999	99	1	1				
		standard 200 mg%							
	32	9992	99	1	2				

* code 1 = Kaunas: code 2 = Rotterdam

Appendix IV

revised 12.2.73

INSTRUCTIONS FOR COMPLETION OF RECORD FORM FOR LABORATORY ANALYSIS

Each form will represent one analytic run of up to thirty samples, controls and standards, each run in duplicate. The identification of the analytic run will be the heading of the form.

Heading

Analytic run no. (col.1-4) A four-digit number where the first digit will represent the type of analytic run (7 = cholesterol, 8 = glucose) followed by the order number of the run within its category: separate numbering order for cholesterol and glucose.

Laboratory code (col.5) The one-digit code of the laboratory is entered according to the code number assigned to the performing laboratory: Kaunas = 1; Rotterdam = 2

Date (col.6-7, 8-9, 10-11) The date when the analytic run is performed is entered as three two-digit numbers in the appropriate boxes for day, month, and year.

Analysed by (col.12) Enter the one-digit code assigned to the laboratory assistant who performs the analytic run. (see "Codes for Laboratory Analysis")

Supervised by (col.13) Enter the one-digit code assigned to the supervisor of the analytic run. (see "Codes for Laboratory Analysis").

Calculation factors (col.14-17) Enter the calculation factor which is to be used for the calculation of the results of this analytic run as a three-digit number with any decimal rounded off. (A zero is printed in the fourth column to indicate this.) Example: 200mg% / .320 O.D. = 622.0.

SAMPLE SPECIFICATION

The identification of the standard and control samples are pre-recorded on the forms. The readings and calculated results of these samples should be recorded in the boxes on the lines assigned for each standard and control samples. All other lines are reserved for samples from subjects in the study.

Number (card no.) (col.18-19) The two-digit number of the card of the sample within an analytic run will be pre-recorded on the form (see "Codes for Laboratory Analysis - Control Sera")

Subject no. (col.20-23) The four-digit subject number assigned to the donor is recorded here.

Appendix IV

revised 12 2.73

Examination code	(col. 24-25) Enter the two-digit code of regular examination (as it appears on the subject's laboratory sheet for regular exams) as follows: 51 = initial screening 52 = rescreening 54 = start of trial examination 12 = final examination The regular check-ups are identified by the number of months since subject's start of trial examination. NOTE: If the sample is taken at an irregular examination, the examination code should be recorded at 00.
Sample no.	(col. 26) The one-digit sample number indicates the order number of the sample from same donor for same examination, e.g. if two blood samples were taken the first sample will be coded as 1 and the second sample will be coded as 2. If a sample is re-analysed for control purposes, the sample number does not change. (A temporary identification number is given by the supervisor to the sample while being analysed and the "true" identification is entered on the form after the control analysis has been finalized.)
Test no.	(col. 27) The one-digit number indicates the order number of tests made of the same sample. A repeat test for control purposes, or because the result of the first test was unacceptable, will always have a test number of a higher value than 1.
O.D.	(col. 28-30) The first reading of a sample's optical density is recorded here as a three-digit integer. (Example: 0.320 is encoded as 320) (col. 31-33) The second reading of a sample's opitcal density is recorded here as a three-digit integer.
mg%	(col. 34-36) The calculated result of the analysed sample is recorded as an integer of three digits.*
Result acceptable	(col. 37) State if the results of the test were acceptable or not acceptable. Enter 1 if result is acceptable or enter 2 if result is not acceptable and a new test is necessary.

* Examples: 127 mg% is encoded as 127
93.7 mg% is encoded as 094

Appendix IV

revised 12 2.73

CODES FOR LABORATORY ANALYSIS

Form No:		Kaunas	Rotterdam
	Cholesterol run no.	7000-7299	7000-7299
	Glucose	8000-8299	8000-8299
	Laboratory code	1	2
	Analysed by:	Z. Sakeliene = 1	Los = 1
		E. Cibauskaite = 2	van Lier = 2
		_____ = 3	_____ = 3
		_____ = 4	_____ = 4
		_____ = 5	_____ = 5
		_____ = 6	_____ = 6
		_____ = 7	_____ = 7
		_____ = 8	_____ = 8
		_____ = 9	_____ = 9
	Supervised by:	A. Toleikis = 1	C. Bartels = 1
		L. Margeviciene = 2	Blijenberg = 2
		_____ = 3	G. Boerma = 3
		_____ = 4	_____ = 4
		_____ = 5	_____ = 5
		_____ = 6	_____ = 6
		_____ = 7	_____ = 7
		_____ = 8	_____ = 8
		_____ = 9	_____ = 9

Control Sera (col.18-19)	Card no.		Subj.no.	Exam.no.
	01	Standard 100 mg%	9991	99
	02	Standard 200 mg%	9992	99
	03	Standard 300 mg%	9993	99
	04	Standard 400 mg%	9994	99
	30	Normal Metrix Control	9998	99
	31	Elevated Metrix Control	9999	99
	32	2nd Standard 200 mg%	9992	99

Note: Only data on standards, controls and patients in the study are to be transmitted to Geneva. Lines containing data for other controls or survey specimens included in a patient run are to be deleted by drawing a line through all data on such a line number.

APPENDIX V

A. ACCEPTABILITY OF ANALYTICAL RESULTS

The following procedures will be incorporated into the final laboratory protocol for checking the validity and acceptability of analytical results and runs.

Check of duplicate readings

1. Duplicate pâtient specimens: Any duplicate pair of O.D. readings yielding a deviation between the two values of greater than three per cent. of the higher value is reanalysed in a subsequent run. For this case, record both values on the report form but do not record an average mg% value. Place a "2" in the "Test Acceptable" column 37, indicating that it is not to be used and that a second set of results will follow for that specimen.

2. Duplicate standard O.D. readings: Determine whether O.D. readings of any pair of standards exceeds three per cent. of the higher reading.

(a) IF two or more pairs of standards differ by more than three per cent., REPEAT THE RUN. Do not fill in a report form for a discarded run.

(b) IF only one pair deviates by more than three per cent. and less than six per cent., proceed; record a "1" indicating that both values are acceptable.

(c) IF only one pair deviates by six per cent. or more, proceed, but indicate that this deviation is unacceptable by recording a "2". Record the O.D. for each duplo.

Calculate separate F factors for each standard O.D. reading, e.g. each pair of standards will have a pair of F values.

(a) Average all F values that are accepted (i.e. all but the unaccepted pair).

(b) Reject the F value of that standard of the unaccepted pair that lies farthest from this first average F.

(c) Reaverage (obtain final $\bar{F}$) including the accepted F values and the remaining F value of the pair from which one was rejected. Note that there must be at least five finally accepted F values for a cholesterol analysis, and seven F values included in the average for any glucose run.

Check of linearity of standards

1. Calculate the average mg% value of each standard pair (three for cholesterol and four for glucose) from the averaged O.D. of each pair of standards; if one of a pair of duplicate values has been discarded and the other is acceptable, calculate from the single remaining value.

2. If the calculated value of the 200 mg% varies by greater than 6 mg% ($> 3\%$) from 200 mg% (i.e. below 194 or above 206), reject the run and repeat analysis of all specimens in another run.

3. If both 200 mg% standard averages lie within 194-206, and all other standards vary from their nominal values by less than three per cent., proceed and calculate values of specimens and controls from the determined averaged $\bar{F}$.

(Example: Average 100 mg% standard lies within 97-103
Average 300 mg% standard lies within 291-309
Average 400 mg% standard lies within 388-412).

4. Otherwise reject and repeat the run after investigating the causes of the non-linearity with new standard solutions and controls only.

Acceptability of control serum results

The third test of acceptability of a run is that both normal and elevated serum control averages must lie within ± six per cent. of their respective assigned values (previously determined by at least 10 runs in duplicate). If this is not the case, the run is rejected, not reported, and is repeated completely. Note that no value is rejected for deviation from its duplicate; all must be used.

This criterion will be introduced with the new lyophilized common lot control serum materials. For the first 10 runs (20 single values for each control), the values are recorded and accepted until a mean determined value is obtained with the methods as they are used in the laboratory. THE LABEL VALUES MAY BE COMPARED BUT ARE NOT USED AS TARGET VALUES; these products will be analysed at CDC by the comparison methods employed for the external survey samples.

After the average of the first 20 values is established, the above criteria of runs "in control" will apply. However, a more refined value will be recalculated at the end of 10 more runs from 40 values in all. This new value will remain the target for determining whether runs are "in control" (within ± six per cent.) for the remainder of the study. Any gradual drift from these target values should be investigated with the object of correcting any deviations in procedure or calibration that may account for such drift. Note: if the value for either control serum seems to have altered to a new and stable value which produces "out of control" results for more than five per cent. of the runs, this new value may be used as the target for acceptability. The drift of alteration of values will, however, be followed serially in the records of the study to document the existence of such changes.

B. GUIDELINES FOR ALLOWABLE VARIANCE

Duplicate variance

The three per cent. limit imposed for the discarding of individual variant results in a duplo standard, or in the computation of the over-all factor from the multiple standard levels, or in the duplo results of a single specimen was obtained from evaluation of experience with the method in both laboratories before the study began. When the average computational factor at each standard level falls outside this limit, judgement must be exercised by the laboratory director in the acceptance of the analysis.

Duplo values for a specimen: When the difference between duplo values for a specimen falls beyond three per cent. of the higher value, the sample is reanalysed in the next run; if the next run is scheduled for the next day, the sample is refrigerated; if the sample must be kept longer, it is kept frozen. The results for that sample are cancelled from the run in which the variance occurs, and the duplo values in the re-run are reported.

Deviations in the standards: Normally, the duplo-optical density results for each standard is divided into its concentration (in mg%) to obtain separate computation factors, and these are averaged to obtain an average computation factor ($\bar{F}$); if one pair of duplo values deviates by greater than three per cent., F is computed for each, and the value that deviates least from $\bar{F}$ is averaged into a final $\bar{F}$. At least seven values for F are obtained for glucose (the 100, 200, 300 initial standards and the final 200 mg% levels); for cholesterol, at least five F values are obtained (200, 400 and final 200). These are ordinarily adequate for the exercise of judgement, and the following guidelines cover most problems that may be encountered.

Appendix V

(a) Non-linearity: The colorimetric reactions should maintain linearity through the highest concentration level, and when they do not, an inspection of the chemistry and re-run of samples is necessary. For glucose, non-linearity will most often be indicated by good consistency between the values for the 100 and 200 mg% standards with a proportionately lower absorbance (higher F) for the 300 mg% standard. In this case, the analysis may be saved by computing its concentration from either the 200 or 300 mg% standard F, whichever level it is nearest. With cholesterol, however, any deviation from linearity in the 200 and 400 mg% standards requires investigation with all four standard levels.

Non-linearity involving a consistent decrease or increase in the slope of optical density (i.e. progressive increase or decrease of F) indicates a more serious problem or potential problem. The most likely cause of an abrupt shift is an erroneous setting of the instrument with blanc, but gradual and progressive drift of F values may arise from instrument malfunction or a change in integrity of the reagents. The control chart of the 200 standard optical density from day to day is the most sensitive indicator of this trouble.

(b) Deterioriation of one standard: When the F value for one of the standard concentrations deviates consistently from the others, this is strong evidence that its concentration is in error, either in its preparation or from deterioration. A new set of working standards is prepared carefully, and these are compared with the old set in a special investigative analytic run.

(c) Decreased precision: Most often increased variability is seen when a new analyst takes over; practice runs in parallel with the regular analyst are recommended. Other sources may be a deterioriation from the standard technique at some point or improper cleaning of glassware.

Deviation in the serum controls (normal, elevated, common)

After the mean value of 20 analyses of a new serum control pool has been determined, the limits of acceptable control are six per cent. on either side of that mean. However, a constant shift of values to a region between three per cent. and six per cent. on one side of the mean indicates a change in base of the analysis or the pool. If this occurs with only one of the serum controls the actual concentration in that pool may have changed; if all controls shift in the same direction, some change in analytical technique or reagents has occurred and the source must be investigated, since most probably all analysed serum specimens show similar changes.

C. PREPARATION OF CONTROL CHARTS

1. Normal and elevated control sera

Plots of average values in each run for normal and elevated controls are made on coordinate paper as in the examples attached, and are posted daily for all runs during a month. If less than 20 runs are performed each month, the data from two months may be accumulated on one chart. This is a laboratory control record for guidance of the laboratory supervisor and staff. After the average concentration levels have been obtained from 10 or 20 runs, the five horizontal lines are drawn opposite suitably plotted concentrations on the abscissa:

a solid line for the average concentration of each pool;
two dashed lines at three per cent. above and below this average value;
two solid lines at six per cent. above and below this average value.

These lines are used as guides for determining whether a run is in control.

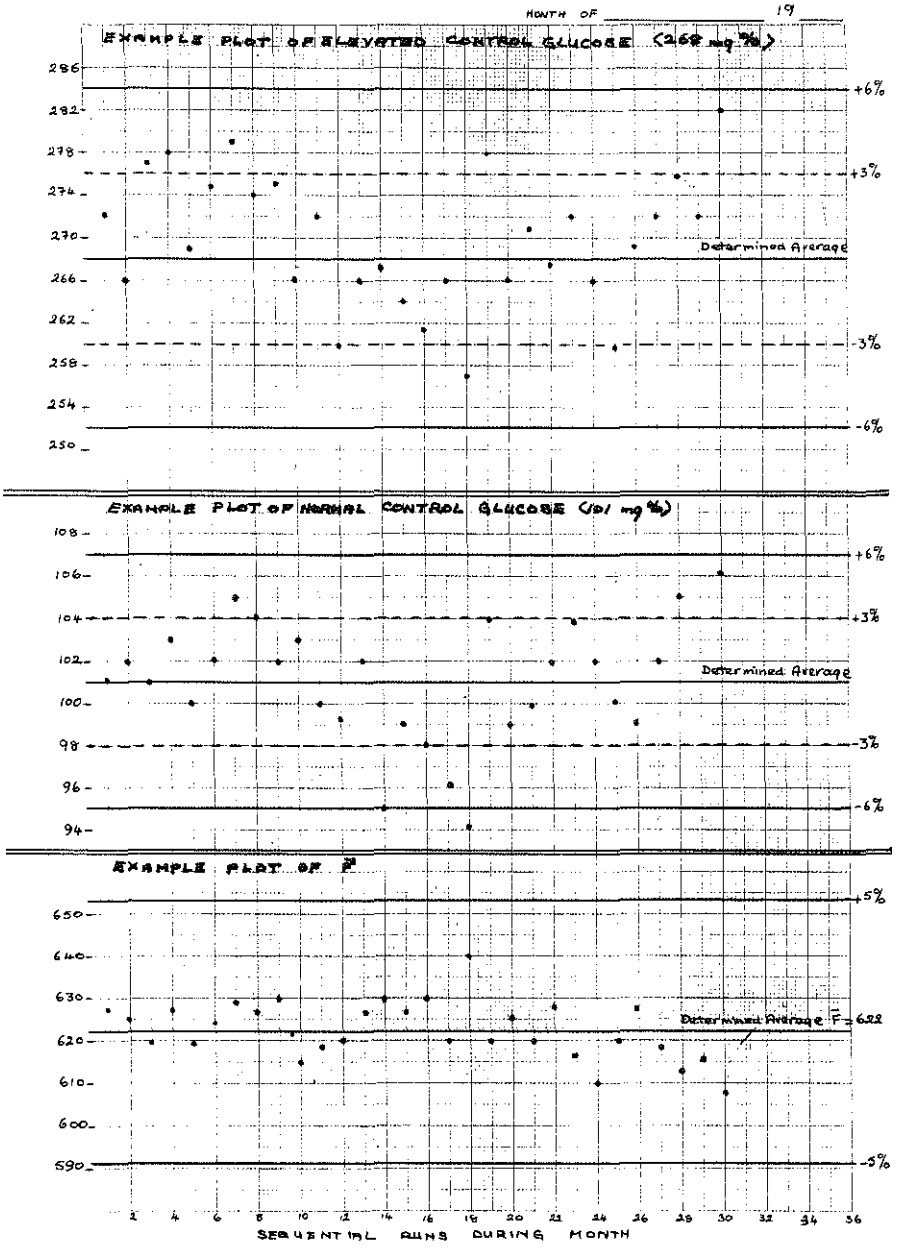
Appendix V

At the completion of each chart (one or two months) a new average value for each control is computed from the data of that period and is compared with the original mean value to determine whether a significant shift has occurred; see the previous sections for guidelines in handling such shifts.

2. Chart of average computation factors

A similar chart of the $\bar{F}$ value of each run is kept. The original average of the first 20 runs is used as a mean for the study, and lines five per cent. above and below this mean are drawn for guidance of the analyst. Whenever a factor falls outside these limits, the run is investigated and repeated if necessary (that is, the control serum values are also out of control). As mentioned in the previous section, when greater than five per cent. of runs yield $\bar{F}$ values that fall beyond these limits in one direction (either all above or all below), obviously the mean has shifted, and a new average is determined and new charts prepared. This assumes that no cause or remedy can be found for the shift in F or optical density of the standards.

Appendix V


MONTH OF 19
EXAMPLE PLOT OF ELEVATED CONTROL GLUCOSE (268 mg %)
286
282
278
274
270
266
262
258
254
250
+6%
+3%
Determined Average
-3%
-6%
EXAMPLE PLOT OF NORMAL CONTROL GLUCOSE (101 mg %)
108
106
104
102
100
98
96
94
+6%
+3%
Determined Average
-3%
-6%
EXAMPLE PLOT OF F
650
640
630
620
610
600
590
+5%
Determined Average F = 622
-5%
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
SEQUENTIAL RUNS DURING MONTH