Four frequently used test systems for serum cholesterol evaluated by isotope dilution gas chromatography–mass spectrometry candidate reference method

LINDA M. THIENPONT,1,3 KATRIEN G. VAN LANDUYT,1 DIETMAR STÖCKL,2 and ANDRÉ P. DE LEENHEER1*

We evaluated the performance of four frequently used cholesterol test systems, using split-sample measurements with a panel of 79 patients' specimens and isotope dilution gas chromatography–mass spectrometry (ID GC-MS) as a comparison method. The test systems were from Beckman, Boehringer Mannheim, Merck, and Johnson & Johnson Clinical Diagnostics, performed on the Synchron CX7, Hitachi 717, Mega, and Ektachem 250 analyzers, respectively. The linear regression data for the method comparison (ID GC-MS as independent variable (y) were for Beckman: slope = 1.012, intercept = 0.0243 mmol/L, dispersion (S_p) = 0.1303 mmol/L, and correlation coefficient (r) = 0.9867; for Boehringer Mannheim: slope = 1.002, intercept = 0.114 mmol/L, S_p = 0.0759 mmol/L, r = 0.9954; for Merck: slope = 1.034, intercept = -0.0613 mmol/L, S_p = 0.0886 mmol/L, r = 0.9941; and for Johnson & Johnson Clinical Diagnostics: slope = 1.007, intercept = 0.01 mmol/L, S_p = 0.15 mmol/L, and r = 0.9811. These data demonstrate excellent state-of-the-art cholesterol measurement for some of the most widely used test systems.

INDEXING TERMS: Reference Method • methods comparison • standardization

The establishment of the National Cholesterol Education Program (NCEP) by the National Institutes of Health in the US in 1987 had the purpose of convincing physicians and the population at large of the relation between high serum cholesterol and cardiovascular disease.4 Within this program, reference values for cholesterol were selected on the basis of probability of developing coronary heart disease, with related recommendations for treatment by diet or drugs. A direct consequence of this program was that precision and accuracy of cholesterol measurements were also emphasized [1–4]. In this context, the Laboratory Standardization Panel (LSP) advocated that for routine cholesterol methods both the total CV and bias be limited to ±3% within a period of 5 years [5]. To achieve this objective, collaborative efforts within the laboratory community and under the auspices of the National Committee for Clinical Laboratory Standards (NCCLS) were set up, namely, the National Reference System for Cholesterol (NRSC) and the Cholesterol Reference Method Laboratory Network (CRMLN). They have the task of standardizing cholesterol methods through manufacturers of instruments, reagents, and control materials or through calibrators commercially available to individual laboratories. In principle, standardization is achieved via a comprehensive measurement system that transfers accuracy from the definitive method based on isotope dilution gas chromatography–mass spectrometry (ID GC-MS) [5, 6] through the Abell–Kendall nationally recognized reference method [7–10] to the routine methods. Because of the possible matrix effects caused by lyophilized specimens, the accuracy transfer from the reference to the routine method has to be done via split-sample measurement of patients' samples [11, 12]. Until

4 Nonstandard abbreviations: NCEP, National Cholesterol Education Program; LSP, Laboratory Standardization Panel; NRSC, National Reference System for Cholesterol; CRMLN, Cholesterol Reference Method Laboratory Network; ID GC-MS, isotope dilution gas chromatography–mass spectrometry; SRM, Standard Reference Material; IQC, internal quality control; PNL, Precinorm L; PPL, Precipath L; IS, internal standard; and CHOD-PAP, cholesterol oxidase-p-aminophenazone.
now, cholesterol is the only analyte for which such a comprehensive measurement system has been applied. To assess its usefulness, even for cholesterol analysis in Europe, where the above performance criteria are not yet officially valid, we evaluated four frequently used cholesterol test systems [Beckman, Boehringer Mannheim, Merck, and Johnson & Johnson Clinical Diagnostics (formerly Kodak)] with an ID GC-MS candidate reference method [13, 14], using a panel of 79 patients' samples. The results of this comparative study are reported here.

**Materials and Methods**

**Patients' Samples**

Patients' samples were kindly provided by the Hospital "Institut Moderne" of Ghent (Belgium) and treated in accordance with the ethical standards of the Committee for Medical Ethics of the University of Ghent. The 79 samples used in this study were randomly collected among a population of hospitalized individuals. All samples had been ordered by physicians for cholesterol analysis. They were divided in 400-μL portions and stored at −20 °C until the day of analysis (−2 months). All samples were shipped frozen to the laboratories participating in this study. They were thawed on the day of analysis and never reused. The cholesterol concentrations in the samples ranged from 3.62 to 8.22 mmol/L.

**Routine Methods**

The following test systems, all performed in the application laboratories of the respective manufacturers or distributors, were used: the Beckman Cholesterol Reagent (kit reorder no. 467825) in combination with the Synchron CX7 Clinical System (Beckman Instruments, Brea, CA), performed in the Laboratory of Analis (Namur, Belgium); the Boehringer Mannheim (Tutzing, Germany) colorimetric cholesterol oxidase-p-aminoephazone (CHOD-PAP) test system (art. no. 1489232), performed on a Hitachi 717 analyzer at 37 °C in the application laboratory of the manufacturer; the Merck (Darmstadt, Germany) CHOD-PAP method (art. no. 107114; special pack for multitest analyzer systems), performed on the Merck Mega analyzer in the application laboratory of Merck Belgium (Overijse, Belgium); and the Ektachem cholesterol slides (cat. no. 168290), performed on an Ektachem 250 analyzer (Johnson & Johnson Clinical Diagnostics, Rochester, NY) in the application laboratory of Johnson & Johnson Clinical Diagnostics (Ilkirch, France).

**Materials for Internal Quality Control**

Internal accuracy control of the ID GC-MS method was done with Standard Reference Material (SRM) 909a-2 from the National Institute of Standards and Technology (NIST, Gaithersburg, MD), with a certified value of 4.463 mmol/L. Internal quality control (IQC) of the routine methods was performed with the control materials recommended by the manufacturers, using the test-system-specific target values (assigned values) for Beckman and Johnson & Johnson Clinical Diagnostics and ID GC-MS target values for Boehringer and Merck. They were Beckman Decision International, levels 1–3, for Beckman Synchron CX 4/5/7/4CE/3CE systems (assigned values: 2.616, 4.118, and 5.620 mmol/L, respectively), Boehringer Mannheim Precinorm L (PNL) and Precipath L (PPL) (ID GC-MS values: 4.843 and 8.806 mmol/L, respectively), Merck Qualitrol HSN and HSP (ID GC-MS values: 2.953 and 2.979 mmol/L, respectively), and Johnson & Johnson Clinical Diagnostics Kodatrol I and II (assigned values: 3.906 and 6.109 mmol/L, respectively).

**ID GC-MS Method**

The ID GC-MS method used for the comparison was described before [13, 14]. Briefly, after addition of the internal standard (IS; 3,4-[13C2]cholesterol) and equilibration, the samples were hydrolyzed with methanolic KOH at 70 °C. After cooling, cholesterol was extracted with hexane. A fraction of the hexane phase was evaporated under a stream of nitrogen, and the residue was redissolved with hexane and derivatized with N-trimethylsilyl-N-methyltrifluoroacetamide. A portion of the derivatization mixture was injected into a Hewlett-Packard (Palo Alto, CA) combined GC-MS 5970B Mass Selective Detector equipped with a 25-m fused silica capillary column coated with methyl silicone as stationary phase (Ultra 1; Hewlett-Packard). The GC injection system was a Model II Programmable Temperature Vaporizing system from Gerstel (Mühlheim a.d. Ruhr, Germany). Mass spectrometric detection was done in the selected-ion monitoring mode at ions m/z 458.5 and 460.5, corresponding to the trimethylsilyl derivatives of, respectively, cholesterol and the labeled IS. For calibration, SRM 911a from NIST was used. The calibrators were taken through the entire analysis.

**Evaluation Protocol**

Assessment of the routine methods was performed by split-sample measurement. The routine laboratories received the sera randomly numbered. ID GC-MS and routine measurements were done with independent duplicates (named A and B for each sample). The comparison was performed on the mean of the duplicates of each sample. For internal accuracy control of the ID GC-MS method, the maximum deviation from the certified value was set to 1%. For IQC of the Beckman and Ektachem methods, the maximum deviation from the assigned target values was set by us to 2%. For Boehringer Mannheim and Merck, however, the IQC limit was set to 3% because the control sera used with these methods had been assigned an ID GC-MS target value and, therefore, could have exhibited a matrix effect. For the ID GC-MS measurements, nine samples were analyzed in one run: on the first day, the samples and standards were pipetted and prepurified; on the second day, samples and standards were injected into the ID GC-MS system. On the day of the measurements, the IQC sample was injected first in duplicate; then the standards and samples were injected (standard 1, standard 2, standard 3, sample 1A, sample 1B, sample 2A, sample 2B, sample 3A, sample 3B, standard 4, standard 5, standard 6, sample 4A, etc.); and at the end of the run, the IQC was measured again in duplicate. Only when the mean IQC value deviated <1% from the target value was the run taken into account.

For the routine methods, all samples were measured on one
day and in one run. At the beginning of the run, each IQC was first analyzed six times. When the preset criteria were fulfilled, the measurements were started in the following way: sample 1A and B until sample 10A and B, each IQC once in duplicate; sample 11A and B, etc., until sample 20A and B, each IQC once in duplicate; sample 21A and B until sample 30A and B, etc.) At the end of the run (after measurement of all 79 samples), again each IQC was measured six times. This means that each individual IQC sample was measured 28 times during the run. The mean deviations from the target values were calculated and had to fulfill the preset criteria. For method comparison, standard linear regression was used.

**Results and Discussion**

For the described method comparison, we took extreme care that the achieved results were representative for the respective test systems, because the measurements were performed by only one laboratory on one day and in one run. First, we evaluated only completely homogeneous test systems, which means that instrument, reagent, and calibrator were from the same manufacturer. In addition, we asked the different manufacturers or distributors whether they were willing to perform the measurements in their application laboratories. Consequently, the Boehringer, Merck, and Johnson & Johnson analyses were performed in the respective application laboratories of Tuttzing (Germany), Overijse (Belgium), and Illkirch (France). The Beckman analyses were performed in the laboratory of the Belgian distributor Analis in Namur (Belgium). Further, they agreed to perform the analyses with rigorous IQC as described below. Before starting the measurements of the patients’ samples, we asked the different laboratories to control whether they principally would be able to reach the values of their IQC materials within the limits preset by us. These preliminary experiments revealed that all laboratories were able to do so. The mean deviations from the low, medium (only for Beckman), and high IQC target values during analysis of the patients’ samples were: Beckman = -1.6%, -1.9%, and -1.7%; Boehringer = +0.1% and +1.7%; Merck = +0.1% and -2.5%; Johnson & Johnson = -1.4% and -0.5%. From these data, we conclude that the results for the patients’ samples are indeed representative for the potential accuracy and precision of the respective test system; however, a laboratory bias on the order of 1–2% cannot be excluded.

Another fact to be considered is that we used frozen specimens in our study. Usually, neither short-term [15] nor long-term freezing [16] considerably affects results for cholesterol. However, some methods exhibit matrix effects for some frozen serum pools [17, 18]. A positive bias (compared with fresh sera) on the order of 1–3% was observed for frozen sera with the Boehringer Mannheim method [19]. Nevertheless, despite the fact that frozen samples might not totally mimic fresh samples, we felt it advantageous that, by use of frozen specimens, we were able to investigate all methods with the same panel of specimens. (Note: For this study, fresh specimens could not be used because of logistic reasons.) However, the results indicate that the influences due to freezing were of minor relevance in our case.

The results of the method comparison by standard linear regression are presented in Fig. 1. ID GC-MS served as the independent variable because of its proven accuracy and its superior precision. The Beckman assay (Fig. 1A) showed a slope of 1.012, an intercept of 0.0243 mmol/L, and a dispersion, S_yx, of 0.1303 mmol/L. The correlation coefficient was 0.9867. The mean bias was +1.7%. These data prove an excellent calibration and specificity of the Beckman test system. To our knowledge, the Beckman test system has not been evaluated before by a split-sample protocol and a candidate reference method. The Boehringer method (Fig. 1B) revealed an excellent correlation with ID GC-MS (r = 0.9954), a very low scatter around the regression line (S_yx = 0.0759 mmol/L), a slope of 1.002, and an intercept of 0.114 mmol/L, also demonstrating excellent specificity and calibration. For Boehringer Mannheim, a mean bias of +2.3% was observed. Our results are in agreement with the results obtained by Warnick et al. [19] (r = 0.998, S_yx = 0.0976 mmol/L, slope = 1.01, intercept = -0.0585 mmol/L), taking into account a 1–2% positive bias of the Abell–Kendall method they used for comparison, vs the ID GC-MS Definitive Method [20]. Our results cannot be compared with those of Poon et al. [21], however, because only lipemic sera were used in that study. (Because of the introduction of NCEP, we used only references after 1986 for comparison.) The Merck assay (Fig. 1C) showed excellent correlation to ID GC-MS (r = 0.9941), a slope of 1.034, an intercept of -0.0613 mmol/L, and a very low dispersion of S_yx = 0.0886 mmol/L. For Merck, a mean bias of +2.3% was observed. To our knowledge, no data comparable with ours are available. The Ektachem method (Fig. 1D) exhibited the largest scatter around the regression line (S_yx = 0.23 mmol/L), mainly due to the unusual behavior of two sera. This also resulted in a relatively poor correlation coefficient (r = 0.9557), as compared with the other test systems. The slope for Johnson & Johnson was 0.9759 and the intercept was 0.1553 mmol/L. The unusual behavior of the two sera may have been due to the fact that the sera had been frozen or had high concentrations of triacylglycerols as described recently [22]; however, this was not further investigated by us. Because of the two samples with an unusual behavior, we also applied linear regression without them, resulting in the following numbers: slope = 1.007, intercept = 0.01 mmol/L, S_yx = 0.15 mmol/L, and r = 0.9811. The mean bias was +0.4% (+0.8% without the two unusual samples). Our data are comparable with those in the Ektachem information sheet (Kodak Ektachem Clinical Chemistry slide cholesterol instruction sheet, publication no. MP2–35), describing the comparison of the Ektachem method with the Abell–Kendall Reference Method (slope = 1.02, intercept = -0.01 mmol/L, S_yx = 0.23 mmol/L, and r = 0.994).

The values for bias (+0.4 to +2.3%) observed by us might be due partly to individual laboratory bias or to the fact that frozen sera were used. For the Ektachem and Beckman systems, additionally, the calibration by the Abell–Kendall Reference Method can contribute to the positive bias, because the latter has a positive bias of +1.6% compared with the ID GC-MS Definitive Method [20].

Concerning the precision of the above test systems, we were able to investigate only the within-run CV (CV_iw), as all samples were measured on one day and in one run. The respective values
for $CV_{wr}$ ($n = 28$) were: Beckman Decision Level 1, 2, and 3, respectively, 0.9%, 0.8%, and 0.8%; Boehringer Mannheim PNL and PPL, respectively, 1.0% and 0.8%; for both Merck HSN and HSP 1.2%; Kodatrol I and II, respectively, 1.0% and 1.3%. These low values for $CV_{wr}$ also suggest values for total $CV (CV_T) \leq 3\%$, which is the goal set by the LSP.

In summary, all test systems fulfilled the NRSC criteria for bias $\leq 3\%$ and $r \geq 0.975$ \cite{4}, which can be attributed to the direct calibration of each of the systems by reference methods (Abell–Kendall for Ektachem and Beckman; ID GC-MS for Boehringer and Merck). The excellent performance of the evaluated test systems can be considered further as a direct consequence of the large efforts made by the NCEP, the LSP, and the NRSC. Therefore, the CRMLN can serve as a model for the establishment of a similar network for other analytes to achieve the success found with cholesterol.

Funding for this research was provided by the Belgian government (contract no. OOA 12050690), by the Fund for Medical Scientific Research (contract no. 33.0007.91), and by the National Fund for Scientific Research through a bursary to K.V.L. Excellent technical assistance of B. Becue and P. Verhaeghe is gratefully acknowledged.

References

1. Current status of blood cholesterol measurement in clinical laboratories in the United States: a report from the Laboratory