Clin Chem Lab Med 2008;46(6):863–871 © 2008 by Walter de Gruyter • Berlin • New York. DOI 10.1515/CCLM.2008.159

Analytical performance evaluation of the Cobas 6000 analyzer – special emphasis on trueness verification

Adriaan J. van Gammeren*, Nelly van Gool, Monique J.M. de Groot and Christa M. Cobbaert

Department of Clinical Chemistry and Hematology, Amphia Hospital, Breda, The Netherlands

Abstract

Background: Consolidation of analyzers is an emerging issue in clinical chemistry. We evaluated the analytical performance of the Cobas 6000 analyzer (Roche Diagnostics), which is considered a candidate for replacement of current Hitachi 917 analyzers and for consolidation of chemistry and immunochemistry.
Methods: The precision, accuracy, linearity and correlation with current field methods were evaluated according to Clinical and Laboratory Standards Institute protocols EP5, EP9 and EP10. A total of 31 routine chemistry assays and 18 immunoassays were studied. Accuracy and linearity were verified for 24 chemistry parameters using value-assigned trueness controls from the Dutch External Quality Assessment Scheme organizers. In addition, traceability to methods endorsed by the Joint Committee of Traceability in Laboratory Medicine was examined.
Results: All analytes met allowable precision criteria, apart from the low level for sodium and folate. Total coefficients of variation ranged between 0.6% and 4.4% for routine chemistry and between 0.8% and 5.8% for immunochemistry, apart from folate (12% at the low end). The correlation coefficients for comparison to current field methods were >0.975, except for magnesium and for six out of 18 immunochemistries. Recovery experiments indicated high recovery for most of the 24 routine chemistry assays.
Conclusions: Considering the excellent precision data and the result equivalence for most assays, it can be concluded that Cobas 6000 accommodates robust chemistry and immunochemistry, and has good potential for workstation consolidation in medium-sized laboratories.
Clin Chem Lab Med 2008;46:863–71.

Keywords: accuracy; Cobas 6000; consolidation; correlation; IVD directive; precision; traceability.

*Corresponding author: Adriaan J. van Gammeren, Department of Clinical Chemistry and Hematology, Amphia Hospital, Langendijk 75, 4819 EV Breda, The Netherlands
Phone: +31-76-5955259, Fax: +31-76-5952092, E-mail: avangammeren@amphia.nl
Received October 15, 2007; accepted January 3, 2008

Introduction

The Cobas 6000 system consolidates clinical chemistry, ion-selective electrode (ISE) methods, and immunochemistry within one analytical system. It uses the Integra cassette and Elecsys rack pack for reagent handling. It has been designed as a random-access, fully automated analyzer for the clinical chemistry laboratory with the capability of determining concentrations or activities of various substances in body fluids, such as enzymes, substrates, electrolytes, proteins, (therapeutic) drugs, cardiac markers, tumor markers, and hormones. The aim of the present study was to evaluate the Cobas 6000 system for precision, accuracy, linearity and correlation in comparison with current field methods on the Hitachi 917 and Elecsys 1010 (Roche Diagnostics, Mannheim, Germany) and Immulite 2000 (Siemens Medical Solutions Diagnostics, Deerfield, IL, USA) systems. The evaluation was performed according to Clinical and Laboratory Standards Institute (CLSI) protocols EP5, EP9 and EP10. Although the current methods are periodically subjected to an External Quality Assessment Scheme (EQAS), method comparison alone is not sufficient to verify trueness, for which EP10 was performed. The traceability of the method results was also examined. Traceability is a property of a test result that allows it to be related to stated international reference methods and reference materials with a stated level of uncertainty, through an unbroken chain of comparisons. The European Commission Directive on in vitro diagnostic medical devices (IVD 98/79/EC) requires manufacturers to establish metrological traceability of measurement results to reference measurement procedures and/or available reference materials of a higher order. Joint Committee of Traceability in Laboratory Medicine (JCTLM)-endorsed reference materials, reference procedures and reference laboratories are currently listed in the database at www.bipm.org/jctlm. We evaluated the traceability of Cobas 6000 chemistries using commutable and value-assigned trueness verification materials for 24 chemistry parameters.

Materials and methods

Analytical performance was investigated in our laboratory for 31 routine chemistry assays and 18 immunoassays. Test results were collected and compared to test results from our current field methods, i.e., Hitachi 917 (Roche Diagnostics) and Immulite 2000 (Siemens Medical Solutions Diagnostics). Addendum 1 and Addendum 2 (see Supplementary data that accompany the online version of this article at http://www.reference-global.com/WDG/loi/cclm) show the test principles for the chemistry and immunochemistry methods,

respectively. Analysis and reagent handling were performed according to Roche recommendations using a single reagent, control and calibrator lot for all tests, except for urine tests. A different reagent lot was used for total protein and albumin analyses in urine during EP5 and EP9 protocols.

Most immunoassays on the Cobas 6000 and Immulite 2000 were performed according to a two-side luminescent enzyme immunoassay (LEIMA) or competitive luminescent enzyme immunoassay (LEIA) with different capture antibodies and tracer antigens (see Addendum 2 in the online Supplementary data). Folate and vitamin B12 assays were performed with a luminescent competition protein-bound enzyme immunoassay (LCPBE). Estradiol was assayed using a radioimmunoassay (RIA, Siemens) according to the competition principle on the Cobas 6000.

Beside fresh samples from daily routine, a variety of pathophysiological samples within the Cobas measurement range were collected from leftover routine samples to obtain a broad and Gaussian measurement range. All samples for routine chemistry evaluation were collected in heparin collection tubes and pathophysiological samples were decanted into cryovials, frozen within 24 h after blood donation and stored at –70°C until use. Before analysis, samples were thawed, centrifuged at 2500×*g* for 5 min at 4°C and analyzed in duplicate within 2 h on both analyzers. All samples were decoded to disconnect medical information associated with samples from patients.

Total protein and albumin in urine were selected for evaluation of urine assays. Urine samples containing increased levels of protein were collected and stored at 4°C until analysis in duplicate on both systems.

For immunoassays, evaluation specimens were collected in serum collection tubes and stored in cryovials at –70°C until use. Frozen samples were thawed and centrifuged at 2500×*g* for 5 min at 4°C. Because vitamin B12 and folate are not stable, an exception was made for these parameters and fresh routine samples were taken.

Calibration was carried out according to Roche recommendations. Within- and between-run precision at two levels was obtained according to CSLI protocol EP5. For direct bilirubin (Dbil) and total bilirubin (Tbil), one control level (Precibil) was used. Standard deviations (SDs) obtained in CLSI EP5 were compared to SDs claimed by the vendor (SD_{VC}) and to medically required SDs (SD_{MR}) taken from Fraser (1). Calculated SD values should be $\leq SD_{VC}$ to verify the vendor's claim and $\leq SD_{MR}$ to fulfil medical requirements. Excellent within- and between-run precision is present when the total SD (SD_{tot}) is smaller than both SD_{VC} and SD_{MR}.

EP9 for immunochemistries was carried out using single measurements on the Immulite 2000 and duplicate measurements on the Cobas 6000, except for vitamin B12 and folate. Samples for vitamin B12 and folate were taken from routine samples and duplicate measurements were performed on both systems within 2 h to minimize the influence of storage.

Statistical analysis was performed using EP-Evaluator® (http://www.dgrhoads.com) and Deming regression analysis. The slopes, intercepts and correlation coefficients (R) obtained were used to deduce a proportional bias, a constant bias and correlation between the methods, respectively. According to EP9, $R > 0.975$ is considered a criterion for excellent correlation. Criteria used by EP-Evaluator® for approval or rejection of EP9 evaluations consist of within-method outlier analysis, between-method outlier analysis, a visual check for linear relationship, an adequate number of test results and an adequate range of results. The visual check for uniform scattering was not considered, since rejection due to non-uniform scattering mostly coincidences with proportional structural bias when slopes deviate from 1.00. If outliers were observed, duplicate measurements were performed with leftover sample on both systems and results were incorporated into the analysis instead of the original measurements.

For linearity and accuracy evaluations, value-assigned trueness verification materials obtained from the Dutch EQAS organization (Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek, SKML) were used to test a selected number of parameters for linearity and recovery according to EP10 and for evaluation of traceability to JCTLM-endorsed reference materials and reference methods (www.skml.nl) (2–4). The trueness verification materials, consisting of two pairs of sets with 11 different levels for each parameter, were analyzed in triplicate for each parameter involved. Accuracy was tested based on the total allowable error (TE_a) with desirable quality specifications as described by Fraser (1).

Results

EP5 results

Calculated SDs at two levels (Precinorm Universal, normal level, PNU; Precipath Universal, pathological level, PPU) for each individual assay of EP5 are summarized in Tables 1 and 2 for routine chemistries and immunochemistries, respectively. Methods approved are indicated with a ''+'' sign, while rejected methods are indicated with a ''–'' sign. If $SD_{COBAS} > SD_{VC}$ or $SD_{COBAS} > SD_{MR}$, the values are indicated in bold. The immunochemistry module is a consolidation of two Elecsys systems with two photometric cells that were evaluated separately and are indicated by C1 and C2. For each cell, two levels (L1 and L2) were investigated. Table 1 demonstrates a narrow CV range between 0.6% and 4.4% for all routine chemistries. For immunoassays, apart from the low level of folate, CVs between 0.8% and 5.8% were observed. For all assays, with the exception of sodium and folate, SD_{tot} is smaller than both SD_{VC} and SD_{MR}, which indicates excellent within- and between-run precision.

EP9 and EP10 results

Because of the inter-relation between field method correlation and accuracy, results from EP9 and EP10, presented in Tables 3 and 4 for routine chemistry, are discussed together. Recovery of chemistry parameters was determined using 11 value-assigned trueness verification materials from the SKML and median recoveries are presented in Table 4. Addendum 5 (see the online Supplementary data) shows all data for each parameter included in the EP10 evaluation. EP9 results for immunochemistry methods are summarized in Table 5. Method comparisons were approved for $R > 0.975$. Table 3 shows that this is the case for all routine chemistry tests, except for magnesium. In the case of magnesium, an inadequate range of results was indicated and a visual check indicates that the majority of results were in the range 0.7–1.0 mmol/L. Addendum 3 (see the online Supplementary data) reveals that a more equal distribution of measurements should improve the correlation

Table 1 Imprecision of routine chemistry methods on the Cobas 6000 analytical system (Roche Diagnostics) according to the CSLI EP5 protocol.

Assay	Mean value	SD_{Cobas}	SD_{VC} (95%)	SD_{MR} (95%)	CV_{Cobas}, %	EP5 interpretation
Albumin PNU, g/L	47.5	0.7	1.7	1.7	1.4	+
Albumin PPU, g/L	31.3	0.6	1.1	1.2	1.8	+
Albumin, urine PN-PUC, mg/L	31.3	1.4	2.3	14.1	4.4	+
Albumin, urine PP-PUC, mg/L	103.6	1.4	7.0	41.9	1.4	+
ALP PNU, U/L	86	2.2	3.0	6.5	2.6	+
ALP PPU, U/L	216	4.2	8.4	17.7	4.2	+
ALT PNU, U/L	48.4	1.0	1.7	13.9	2.1	+
ALT PPU, U/L	136.8	1.8	5.3	42.7	1.3	+
Amylase PNU, U/L	80	0.6	2.9	8.4	0.8	+
Amylase PPU, U/L	204	1.4	7.3	21.2	0.7	+
AST PNU, U/L	45.9	1.0	1.6	6.4	2.2	+
AST PPU, U/L	148.5	1.4	5.5	21.2	0.9	+
Bilirubin direct, Precibil, μmol/L	62.7	0.7	2.4	29.8	1.2	+
Bilirubin total, Precibil, μmol/L	214.6	3.3	8.0	68.0	1.5	+
Calcium PNU, mmol/L	1.96	0.03	0.07	0.05	1.3	+
Calcium PPU, mmol/L	3.23	0.04	0.12	0.08	1.1	+
Cholesterol PNU, mmol/L	2.54	0.05	0.10	0.19	2.0	+
Cholesterol PPU, mmol/L	4.96	0.08	0.18	0.37	1.5	+
CK PNU, U/L	174	2.4	6.4	48.5	1.4	+
CK PPU, U/L	533	4.6	19.5	147.9	0.9	+
Creatinine PNP, μmol/L	85	1.9	3.1	4.5	2.2	+
Creatinine PPP, μmol/L	375	6.4	13.6	19.5	1.7	+
C-reactive protein PNP, mg/L	10.4	0.4	0.7	5.1	3.4	+
C-reactive protein PPP, mg/L	47.9	0.6	2.9	24.4	1.2	+
GGT PNU, U/L	48	0.4	1.7	7.9	0.8	+
GGT PPU, U/L	235	1.5	8.5	38.7	0.6	+
Glucose PNU, mmol/L	5.1	0.05	0.19	0.41	1.0	+
Glucose PPU, mmol/L	14.2	0.16	0.51	1.10	1.1	+
IgM PNP, g/L	0.75	0.02	0.05	0.05	2.0	+
IgM PPP, g/L	1.26	0.02	0.09	0.09	1.6	+
IgA PNP, g/L	2.14	0.02	0.07	0.14	1.1	+
IgA PPP, g/L	3.18	0.05	0.23	0.18	1.4	+
IgG PNP, g/L	9.0	0.17	0.67	0.50	1.8	+
IgG PPP, g/L	14.3	0.34	1.04	0.78	2.4	+
HDL-C PNL, mmol/L	1.23	0.02	0.04	0.11	2.0	+
HDL-C PPL, mmol/L	0.88	0.03	0.04	0.08	3.8	+
Iron PNU, μmol/L	19.8	0.28	0.72	6.37	1.4	+
Iron PPU, μmol/L	30.2	0.38	1.13	9.96	1.3	+
Lactate dehydrogenase PNU, U/L	179	2.8	6.1	13.5	1.6	+
Lactate dehydrogenase PPU, U/L	278	4.6	10.1	49.3	1.6	+
Magnesium PNU, mmol/L	1.04	0.014	0.036	0.045	1.3	+
Magnesium PPU, mmol/L	1.77	0.024	0.061	0.077	1.4	+
Phosphate PNU, mmol/L	1.20	0.02	0.05	0.13	1.6	+
Phosphate PPU, mmol/L	2.13	0.03	0.08	0.22	1.2	+
Total protein PNU, g/L	67.6	1.2	2.6	2.3	1.8	+
Total protein PPU, g/L	49.4	0.8	1.9	1.6	1.6	+
Total protein, urine PN-PUC, g/L	0.164	0.004	0.012	0.082	2.5	+
Total protein, urine PP-PUC, g/L	1.394	0.011	0.101	0.664	0.8	+
Triglyceride PNU, mmol/L	1.43	0.02	0.05	0.35	1.2	+
Triglyceride PPU, mmol/L	2.36	0.02	0.08	0.59	0.9	+
Uric acid PNU, μmol/L	241	3.0	8.7	25.1	1.3	+
Uric acid PPU, μmol/L	642	7.9	23.5	67.4	1.2	+
Urea PNU, mmol/L	7.0	0.1	0.3	1.1	1.9	+
Urea PPU, mmol/L	24.9	0.5	0.9	3.8	1.8	+
Sodium PNU, mmol/L	**119**	**1.1**	2.9	**1.0**	0.9	–
Sodium PPU, mmol/L	142	0.9	3.4	1.2	0.6	+
Potassium PNU, mmol/L	3.27	0.03	0.07	0.19	1.0	+
Potassium PPU, mmol/L	6.33	0.05	0.11	0.36	0.7	+
Chloride PNU, mmol/L	82.1	0.8	2.1	1.3	1.0	+
Chloride PPU, mmol/L	114.2	0.7	2.7	1.6	0.6	+

PNU, Precinorm Universal, normal level; PPU, Precipath Universal, pathological level; PNP, Precinorm Protein; PPP, Precipath Protein; PUC, protein in urine concentration; SD_{VC}, standard deviation claimed by the vendor; SD_{MR}, medically required standard deviation. The value in bold indicates $SD_{COBAS} > SD_{MR}$.

Table 2 Imprecision of immunochemistry methods on the Cobas 6000 analytical system (Roche Diagnostics) according to the CSLI EP5 protocol.

Assay	Mean value	SD_{Cobas}	SD_{VC} (95%)	SD_{MR} (95%)	CV_{Cobas}, %	EP5 interpretation
TSH L1 C1, mIU/L	1.34	0.02	0.12	0.33	1.3	+
TSH L1 C2, mIU/L	1.36	0.02	0.12	0.32	1.2	+
TSH L2 C1, mIU/L	8.27	0.07	0.73	2.01	0.8	+
TSH L2 C2, mIU/L	8.37	0.08	0.71	1.98	0.9	+
FT4 L1 C1, pmol/L	13.7	0.43	1.21	0.85	3.2	+
FT4 L1 C2, pmol/L	13.8	0.37	1.19	0.84	2.7	+
FT4 L2 C1, pmol/L	32.5	1.26	2.90	2.02	3.9	+
FT4 L2 C2, pmol/L	32.7	1.05	2.86	2.00	3.2	+
Adrenocorticotropic hormone L1 C1, pg/mL	90.2	1.88	7.80	11.14	2.1	+
Adrenocorticotropic hormone L1 C2, pg/mL	994.9	2.09	7.80	11.14	2.3	+
Adrenocorticotropic hormone L2 C1, pg/mL	90.6	28.1	87.3	124.7	2.8	+
Adrenocorticotropic hormone L2 C2, pg/mL	995.6	25.9	88.3	126.1	2.6	+
AFP L1 C1, kIU/L	7.9	0.18	0.70	1.20	2.4	+
AFP L1 C2, kIU/L	8.1	0.19	0.69	1.18	2.4	+
AFP L2 C1, kIU/L	100.1	1.76	8.52	14.6	1.8	+
AFP L2 C2, kIU/L	100.0	1.80	8.65	14.8	1.8	+
CEA L1 C1, µg/L	5.4	0.18	0.47	0.86	3.4	+
CEA L1 C2, µg/L	5.4	0.15	0.46	0.84	2.8	+
CEA L2 C1, µg/L	50.2	1.68	4.51	8.18	3.3	+
CEA L2 C2, µg/L	51.2	1.17	4.53	8.23	2.3	+
Cortisol L1 C1, nmol/L	331	17.6	27.6	82.3	5.3	+
Cortisol L1 C2, nmol/L	848	17.2	28.0	83.6	2.1	+
Cortisol L2 C1, nmol/L	337	15.0	72.6	216.6	1.8	+
Cortisol L2 C2, nmol/L	856	13.5	70.5	210.4	1.6	+
DHEAs L1 C1, µmol/L	5.6	0.18	0.43	0.26	3.2	+
DHEAs L1 C2, µmol/L	5.5	0.18	0.43	0.26	2.7	+
DHEAs L2 C1, µmol/L	14.6	0.44	1.20	0.71	3.0	+
DHEAs L2 C2, µmol/L	14.7	0.37	1.19	0.71	2.5	+
Estradiol L1 C1, nmol/L	0.49	0.02	0.04	0.10	4.8	+
Estradiol L1 C2, nmol/L	0.49	0.02	0.04	0.10	3.2	+
Estradiol L2 C1, nmol/L	2.21	0.06	0.18	0.45	2.8	+
Estradiol L2 C2, nmol/L	2.17	0.06	0.17	0.45	2.8	+
Ferritin L1 C1, ng/mL	10.9	0.44	0.91	1.93	4.0	+
Ferritin L1 C2, ng/mL	11.1	0.43	0.91	1.93	3.9	+
Ferritin L2 C1, ng/mL	1358	19.0	114.9	244.6	1.4	+
Ferritin L2 C2, ng/mL	1362	20.2	115.8	246.5	1.5	+
Folate L1 C1, nmol/L	**4.5**	**0.53**	**0.40**	1.37	11.7	–
Folate L1 C2, nmol/L	**4.2**	**0.50**	**0.40**	1.35	11.8	–
Folate L2 C1, nmol/L	32.9	1.66	2.66	9.12	5.1	+
Folate L2 C2, nmol/L	32.5	1.33	2.69	9.49	4.1	+
LH L1 C1, IU/L	7.9	0.15	0.66	1.38	1.9	+
LH L1 C2, IU/L	8.1	0.11	0.66	1.35	1.4	+
LH L2 C1, IU/L	48.7	0.69	4.16	8.58	1.4	+
LH L2 C2, IU/L	49.2	0.77	4.16	8.58	1.6	+
Prolactin L1 C1, IU/L	159.1	2.3	12.9	12.8	1.4	+
Prolactin L1 C2, IU/L	161.9	2.7	13.1	13.0	1.6	+
Prolactin L2 C1, IU/L	652.4	9.8	52.4	51.7	1.5	+
Prolactin L2 C2, IU/L	663.5	12.0	52.6	51.8	1.8	+
SHBG L1 C1, nmol/L	29.2	0.44	2.58	4.46	1.5	+
SHBG L1 C2, nmol/L	29.3	0.47	2.66	4.37	1.6	+
SHBG L2 C1, nmol/L	15.6	0.26	1.29	2.23	1.6	+
SHBG L2 C2, nmol/L	15.5	0.26	1.31	2.25	1.7	+
Vitamin B12 L1 C1, pmol/L	175.7	4.8	14.4	30.9	2.7	+
Vitamin B12 L1 C2, pmol/L	171.8	4.7	14.5	31.0	2.7	+
Vitamin B12 L2 C1, pmol/L	848.7	16.9	71.6	153.4	2.0	+
Vitamin B12 L2 C2, pmol/L	847.4	14.7	70.9	151.9	1.7	+
Progesterone L1 C1, nmol/L	29.2	0.65	2.65	3.78	2.2	+
Progesterone L1 C2, nmol/L	29.1	0.96	2.61	3.72	3.3	+
Progesterone L2 C1, nmol/L	66.1	1.51	5.92	8.46	2.3	+
Progesterone L2 C2, nmol/L	65.5	1.85	5.96	8.51	2.8	+
FSH L1 C1, IU/L	9.5	0.24	0.82	1.19	2.5	+
FSH L1 C2, IU/L	9.5	0.20	0.81	1.18	2.1	+
FSH L2 C1, IU/L	40.5	0.93	3.49	5.03	2.3	+
FSH L2 C2, IU/L	40.9	0.83	3.50	5.05	2.0	+
CA 125 L1 C1, IU/mL	40.5	0.90	3.57	14.90	2.2	+
CA 125 L1 C2, IU/mL	40.4	0.81	3.52	14.66	2.2	+
CA 125 L2 C1, IU/mL	119.2	2.04	10.54	43.97	1.7	+

(Table 2 continued)

Assay	Mean value	SD_{Cobas}	SD_{VC} (95%)	SD_{MR} (95%)	CV_{Cobas}, %	EP5 interpretation
CA 125 L2 C2, IU/mL	120.8	1.74	10.45	43.97	1.4	+
CA 15-3 L1 C1, IU/mL	18.8	0.67	1.74	1.54	3.6	+
CA 15-3 L1 C2, IU/mL	18.8	0.72	1.77	1.56	3.8	+
CA 15-3 L2 C1, IU/mL	90.2	4.16	8.90	7.89	4.6	+
CA 15-3 L2 C2, IU/mL	89.8	5.20	8.93	7.91	5.8	+
tPSA, total L1 C1, ng/mL	3.3	0.08	0.30	0.76	2.3	+
tPSA, total L1 C2, ng/mL	3.3	0.05	0.30	0.76	1.5	+
tPSA, total L2 C1, ng/mL	34.7	0.69	3.14	8.13	2.0	+
tPSA, total L2 C2, ng/mL	35.2	0.83	3.02	7.82	2.3	+

SD_{VC}, standard deviation claimed by the vendor; SD_{MR}, medically required standard deviation. Values in bold indicate $SD_{COBAS} > SD_{VC}$.

to $R > 0.975$, although a minor constant bias of 0.1 mmol/L still remains. The structural bias may be caused by the method difference (Xylidolblue vs. Chlorophosphonazo III).

The sodium method demonstrated a slope of 1.071 and intercept of –11.8 mmol/L which can also be explained by a narrow measuring range: 47 out of 54 measurements are clustered in the range 135–150 mmol/L, which has a significant influence on the slope and intercept if small changes in the results are introduced. For chloride a similar effect is observed.

In spite of excellent method correlation for glucose, a relatively negative bias was observed for concentrations >15 mmol/L. For concentrations <15 mmol/L, exchangeable results were observed. EP10 revealed glucose recovery of 102.0%–97.4% for concentrations of 3.8–10.4 mmol/L. The recovery decreased to 96.5%–95.4% for concentrations of 13.8–37.0 mmol/L, indicating some non-linearity, which is allowed within the EP10 criteria.

The high-density lipoprotein-cholesterol (HDL-C) EP10 results encompassed two outliers, with lower values measured by the Cobas 6000 with a third-

Table 3 Method comparison of 31 routine chemistry parameters on the Cobas 6000 system (Y) compared to the Hitachi 917 method (X) according to the CLSI EP9 protocol.

Assay	Slope	Intercept	Correlation, R	Range tested Cobas	EP9 interpretation
Albumin, g/L	1.004	–0.87	0.996	10.8–49.7	+
Albumin, urine, mg/L	1.067	–2.70	0.998	0–354.2	+
ALP, U/L	0.996	–0.50	1.000	41–969	+
ALT, U/L	1.002	–0.35	1.000	7.5–458.5	+
Amylase, U/L	1.065	1.10	1.000	21–1353	+
AST, U/L	1.075	–1.99	1.000	9.6–685.5	+
Bilirubin direct*, μmol/L	0.905	0.23	0.998	2.08–145.7	+
Bilirubin total, μmol/L	0.977	–0.82	1.000	2.5–420.9	+
Calcium, mmol/L	0.943	0.00	0.982	1.32–3.27	+
Cholesterol, mmol/L	1.051	–0.04	0.998	1.6–10.10	+
CK, U/L	0.979	–2.20	1.000	13–1958	+
Creatinine, μmol/L	1.018	–3.90	1.000	45–1244	+
CRP*, mg/L	1.047	0.02	1.000	0.66–120.3	+
GGT, U/L	0.987	0.40	1.000	9–1190	+
Glucose, mmol/L	0.990	0.03	1.000	3.81–34.8	+
IgM, g/L	0.950	0.02	0.999	0.24–3.81	+
IgA, g/L	0.963	0.12	0.999	0.56–5.76	+
IgG, g/L	1.145	–2.05	0.982	3.93–45.79	+
HDL-C, mmol/L	0.991	–0.07	0.976	0.2–2.07	+
Iron, μmol/L	0.995	–0.17	1.000	1.01–55.67	+
Lactate dehydrogenase, U/L	1.048	5.90	0.999	143–912	+
Mg, mmol/L	1.017	0.10	**0.971**	0.57–1.39	–
Phosphate, mmol/L	1.002	0.00	1.000	0.48–5.29	+
Total protein, g/L	0.970	0.67	0.998	18.1–97.8	+
Total protein in urine, g/L	0.947	–0.015	0.997	0.044–1.65	+
Triglyceride, mmol/L	1.038	–0.027	0.996	0.45–7.3	+
Uric acid, μmol/L	0.956	0.01	1.000	0.133–0.94	+
Urea, mmol/L	0.996	–0.02	1.000	1–46.9	+
Sodium, mmol/L	1.071	–11.80	0.977	113–169	+
Potassium, mmol/L	1.000	0.015	0.998	2.09–7.29	+
Chloride, mmol/L	1.036	–3.46	0.988	72.1–132	+

*Preliminary result because of insufficient pairs of measurements: 34 for CRP, 29 for bilirubin. The value in bold does not meet the criterion $R > 0.975$.

Table 4 Median recovery (%) and trueness verification of routine chemistries tested according to the CLSI EP10 protocol using 11 pairs of trueness verification samples.

Assay	Approved*	EP10 interpretation	Range tested	Median recovery, %
Albumin, g/L	4	–[b]	29.3–60.4	102.6
ALP, U/L	11	+	86–271	98.1
ALT, U/L	11	+	10–81	99.7
Amylase, U/L	11	+	42–407	100.7
AST, U/L	0	–[a]	20–316	107.4
Calcium, mmol/L	2	–[b]	1.73–3.54	97.7
Chloride, mmol/L	0	–[a]	83.6–117	96.8
Cholesterol, mmol/L	0	–[a]	4.65–6.81	104.2
CK, U/L	6	–[c]	70–454	85.8
Creatinine, μmol/L	11	+	56–796	100.5
GGT, U/L	11	+	27–201	100.3
Glucose, mmol/L	3	–[a]	3.8–37.0	96.5
HDL-C, mmol/L	5	–[a]	0.93–1.79	105.2
Iron, μmol/L	11	+	19–73.2	99.3
Lactate dehydrogenase, U/L	11	+	97–953	100.8
Magnesium, mmol/L	4	–[b]	0.64–2.05	102.5
Phosphate, mmol/L	11	+	0.80–3.08	100.7
Potassium, mmol/L	5	–[a]	3.24–8.38	98.2
Sodium, mmol/L	0	–[a]	116.0–167.0	97.3
Bilirubin total, μmol/L	9	–[a]	4.8–103	92.3
Total protein, g/L	8	–[b]	46.6–80.5	100.4
Triglycerides, mmol/L	9	–[a]	0.71–6.68	109.1
Uric acid, μmol/L	0	–[a]	0.22–0.58	91.5
Urea**, mmol/L	8	+	4.3–48.2	102.0

*In total, 11 trueness verification values were used. **For urea, eight trueness verification controls were measured because three concentrations were outside the measurement range of the Cobas 6000 system for direct measurements. All eight trueness verification values were approved. [a]Can be approved if calibration adjustment is performed. [b]Variable recovery across the measuring range, calibration adjustment should be focused on the clinically relevant range. [c]CK shows under-recovery at the higher activity range owing to instability and light sensitivity of the EQAS material itself.

generation reagent compared to the Hitachi 917 with a second-generation reagent.

α-Fetal protein (AFP), carcinoembryonic antigen (CEA), dehydroepiandrosterone sulfate (DHEAs), ferritin, follicle-stimulating hormone (FSH), luteinizing hormone (LH), progesterone, prolactin, prostate-specific antigen (PSA), sex hormone-binding globulin (SHBG), thyroid-stimulating hormone (TSH) and vitamin B12 showed R-values >0.975, which are approved (Table 5 and Addendum 4 in the online Supplementary data). For cancer antigen (CA) 15.3, CA 125, cortisol, folate, free thryoxine (FT4) and estradiol,

Table 5 Method comparison of 18 immunochemistry parameters on the Cobas 6000 system (Y) compared to Immulite 2000 (X) and RIA (Siemens) methods according to the CLSI EP9 protocol.

Assay	Slope	Intercept	Correlation, R	Range tested Cobas	EP9 interpretation
TSH, mIU/L	0.876	0.20	0.986	0.484–42.7	+
FT4, pmol/L	0.805	–0.07	**0.912**	9.67–40.63	–
AFP, kIU/L	0.975	0.97	0.981	0.73–64.57	+
CEA, ng/L	0.645	1.02	0.997	0.87–234.7	+
Cortisol, nmol/L	0.993	43.48	**0.962**	153.1–1167	–
DHEAs, μmol/L	1.177	0.31	0.986	0.58–17.14	+
Estradiol, nmol/L	0.689	0.081	**0.617**	0.035–1.74	–
Ferritin, ng/mL	0.932	1.26	0.997	4.65–572	+
Folate*, nmol/L	0.741	5.31	**0.861**	7.37–34.63	–
LH, IU/L	1.149	0.54	0.992	0.211–73.0	+
Prolactin, mIU/L	1.054	14.6	0.997	56.43–800	+
SHBG, nmol/L	1.130	–0.26	0.996	10.21–182	+
Vitamin B12, pmol/L	1.013	28.36	0.987	133.8–705	+
Progesterone, nmol/L	1.149	–0.90	0.993	1.43–116.8	+
FSH, IU/L	0.996	–0.41	0.996	0.272–115	+
CA 125, IU/mL	2.084	–69.48	**0.682**	7.9–4274	–
CA 15-3, IU/mL	0.846	–3.97	**0.920**	4.8–266	–
tPSA, total, ng/mL	0.724	0.04	0.985	0.151–15.7	+

*Six out of 40 folate results were obtained using a second lot because of reagent shortage of the first lot. Values in bold do not meet the criterion R>0.975.

R-values are <0.975. Collecting samples distributed over the entire measurement range was difficult for AFP, CEA, prolactin, estradiol and PSA. For the latter, the range tested was limited to the range indicated in Table 5. In addition, national EQAS data for the immunochemistry assays from Roche and Siemens (SKML no. 2007.3) were taken into consideration. These EQAS data, based on pooled human serum, were used as adequate comparison material for putting the relatively poor correlation for CA 15.3, CA 125, cortisol, folate, FT4 and estradiol in our study into perspective.

For CA 15.3, it has been observed that the dispersion diverges with increasing concentration (Addendum 4, see the online Supplementary data). EQAS data show a 10%–13% decrease in CA 15.3 values for the Roche methodology compared to the Siemens/DPC methodology. This is in line with the slope of 0.846 in our method comparison for CA 15.3. For CA 125, relative differences are mainly observed for quantities >250 U/mL. Excluding measurements >250 U/mL improved the R-value from 0.682 to 0.930. The EQAS data also reveal a relatively large positive bias of 13%–20% for the Roche methodology, which is in line with the relatively large positive bias in our comparison experiments.

The slope of 0.993 for the cortisol assay indicates equal results for both methods. The method comparison data (Table 5 and Addendum 4, see the online Supplementary data) reveal a constant bias of 43 nmol/L. The national EQAS data are in accordance and reveal an increased value at 280 nmol/L, whereas equal results are obtained at 1010 nmol/L. The R-value of 0.962 is just below the approval level of 0.975.

For folate, an R-value of 0.861 indicates poor correlation, which can be partly attributed to the relatively high CV observed in EP5. Six folate samples of the total set for EP9 were added using a second reagent lot. The bias plot shows a small positive bias for folate concentrations <20 nmol/L and a negative bias for folate concentrations >20 nmol/L, resulting in a positive intercept of 5.31 nmol/L and a slope of 0.741. National EQAS results (survey no. 2007.3) show peer-group mean values of 19.5 and 30 nmol/L for the Roche method and 19.4 and 35 nmol/L for Immulite 2000, indicating a similar trend to that observed in our folate method comparison.

The low R-value of 0.912 for FT4 is attributed to an inadequate range of data results. Most data are in the concentration range between 10 and 30 pmol/L. A broader range is expected to improve the correlation. The slope of 0.805, indicating a relative result difference of 20% between the methods, could not be confirmed by the national EQAS survey data. A total of 13 additional samples with low FT4 concentrations were collected and analyzed on both the Immulite 2000 and Cobas 6000 systems. These data were used to obtain an impression of the results at the low end of the range and are not included in the method comparison. The Cobas data show relative differences ranging between –24% and +34% compared to the Immulite data for concentrations between 6.04 and 12.87 pmol/L, indicating poor correlation in the low range.

A correlation coefficient of 0.617 and slope of 0.689 for estradiol indicate a very poor correlation and totally different results. Results <1.0 nmol/L show an R-value of 0.955 and slope of 1.960. Results >1.0 nmol/L show much lower values compared to the RIA assay. Samples >1.0 nmol/L were reanalyzed, and fresh patient and EQAS samples with estradiol concentrations >1.0 nmol/L were included. For these samples, differences similar to those for the consensus values of the Siemens/DPC and Roche peer-groups in the national EQAS survey were observed. The estradiol EQAS data reveal ~50% result differences within method groups and significant result differences between method groups, which highlights the poor performance of current estradiol immunoassays.

For the chemistry methods evaluated, the traceability of methods manufactured by Roche Diagnostics has been verified. JCTLM-endorsed methods and materials have been inventoried from the website of Bureau International des Poids et Mesures (BIPM) and are summarized in Addenda 1 and 2 for routine chemistry and immunochemistry, respectively. For most routine chemistry assays, an international reference material and/or method is registered at JCTLM, and Roche Diagnostics methods are stated to be traceable to these standards of higher order. For AFP, CEA, tPSA, SHBG, TSH, progesterone, cortisol and estradiol, international standards are available (www.bipm.org/jctlm) and are indicated in Addendum 2 (see the online Supplementary data). Notwithstanding availability, commutability has rarely been documented. For other immunochemistry methods, traceability can only be established to the manufacturer's own in-house reference systems, because international reference methods or materials are not yet available.

Discussion

Considering the precision, method comparison and recovery data for the chemistry parameters, it can be concluded that the Cobas 6000 exhibits robust chemistry and immunochemistry and has good potential for workstation consolidation. Routine chemistry CVs, ranging between 0.6% and 4.4%, are tight. Our data show that the SDs for all assays apart from sodium and folate fulfill the EP5 criteria ($SD_{tot} < SD_{VC}$ and $SD_{tot} < SD_{MR}$). For sodium, SD_{tot} is slightly greater than SD_{MR}, which can be attributed to the borderline performance of the current state of the art compared to medical requirements. For folate, SD_{tot} at the lower end of the reference range does not meet the vendor claim, but fulfills the SD_{MR} criterion. The value of commercial internal QC materials is limited to precision evaluation because of a lack of commutability with clinical species. Our evaluation is limited to a single reagent lot. It should be remembered that lot variation

affects precision but does not affect the systematic analytical error.

Considering the chemistry method comparison and the $R > 0.975$ criterion, all routine chemistries, except for magnesium, correlate well with our current Hitachi 917 methods. Nevertheless, different calibrations and/or calibration procedures may explain the EP9 slope differences for the enzymatic aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), amylase, creatine kinase (CK) and γ-glutamyltransferase (GGT) assays. The current comparative field methods (Hitachi 917) have been calibrated according to Calibration 2000 using a fixed factor instead of an enzyme calibrator (2–4). Total protein, calcium and magnesium showed poor recovery only at the low end of the concentration range and demonstrated excellent recovery within the normal part of the range tested. For these assays, calibration adjustment should be focused on the clinical reference range.

For HDL-C, adequate recovery was obtained for normo-triglyceridemic samples. Outliers observed in EP10 and EP9 protocols appeared to be samples with relatively high triglyceride concentrations. A multicenter evaluation revealed that the second-generation Roche reagent can overestimate HDL-C concentrations compared to the third-generation reagent because of stronger interference from triglycerides (5). This is confirmed by EP10 results, which show relatively poor recovery only for the hypertriglyceridemic samples of the 11-set EQAS sample series.

Concerning the immunochemistry assays, 12 out of 18 assays show $R > 0.975$ in the method comparisons and were approved. Less satisfactory results were obtained for method comparisons of CA 15.3, CA 125, cortisol, folate, FT4 and estradiol (Addendum 4 in the online Supplementary data). Notwithstanding the good within- and between-run precision for CA 15.3, CA 125, and FT4 in EP5, it appears that there are large inter-method differences in biological samples. Apparently the assay results are influenced by method-related matrix effects, the use of different capture or tracer antibodies and the application itself. Moreover, it has to be taken into account that calculated mass units used by different manufacturers are not necessarily the same. Harmonized results cannot be obtained by calibration adjustment alone. The strong need for better FT4 and steroid (immuno)assays is generally recognized (6–8). For a better understanding of the performance of these methods, further follow-up and/or trueness verification in national EQA surveys is urgently needed. Linearity and accuracy based on desired specifications were verified for 24 routine chemistry parameters using commutable trueness verifiers (Table 4). Linearity met the EP criterion for all assays involved (Addendum 5 in the online Supplementary data). For nine out of 24 chemistry assays, accuracy could be directly approved. For AST, chloride, cholesterol, glucose, HDL-C, potassium, sodium, bilirubin total, triglyceride and uric acid, calibration adjustment is needed to align the recovery results with target specifications. Calibration adjustments needed for the other chemistry parameters, according to over- or under-recovery observed in the EP10 protocol, are roughly in line with the slope and intercepts results for the EP9 protocol (Tables 3 and 4).

Considering traceability (Addendum 1 in the online Supplementary data), routine chemistry methods with the exception of ALP, direct bilirubin, phosphate and total protein are traceable to the highest metrological standard (SI units). For chemistry methods with a reference method and/or material available at www.bipm.org/jctlm, Roche fulfils the IVD requirements, as their methods are stated to be traceable to the JCTML-endorsed metrological standards.

Addendum 2 (see the online Supplementary data) shows that for a number of immunochemistries tested (AFP, tPSA prolactin, CEA, progesterone, cortisol and estradiol) an international standard (WHO, IRMM, NIST or NIBSC) is recommended. These standards are also used by Roche, except for progesterone, cortisol and estradiol. For other immunoassays tested, traceability can be established at best to an in-house reference system. For harmonization there is a strong need to develop international standards for analytes for which such standards are not yet available.

Because of the availability of JCTLM-endorsed methods and reference materials for chemistry methods, accuracy can be examined and monitored in EQAS surveys. For a number of immunochemistry assays, no reference methods and/or reference materials are available and harmonization is targeted by aligning measurement results to all-laboratory total mean (ALTM) values. The accuracy of a number of immunoassay methods requires substantial improvement and EQAS organizers should adopt absolute target values instead of peer group means whenever reference systems are in place. Some EQAS providers regularly assess recovery using international standards where they exist (9). Currently, HPLC-tandem mass spectrometry technology and applications to metrological measurements of analyte concentrations have increasingly become available to EQAS providers and reference laboratories. This promising technique has potential as an excellent reference method to bring about high accuracy for FT4, steroids and even low-molecular-weight proteins (8–13).

For non-steroids, in spite of considerable improvements resulting from increased assay automation, major inter-method differences remain. Reasons for this include different calibration procedures, antibody specificity, assay design and vulnerability to clinically relevant interferences. An improvement in comparability will require a clear definition of the measurement and accurate calibration with highly purified standards, broad recommendations regarding the most appropriate antibody combinations, and increased awareness of clinically relevant interferences (14).

We conclude that the transition from Hitachi 917 to Cobas 6000 routine chemistry was smooth, leading to exchangeable and well-standardized results for all 31 chemistry parameters. The transition from Immulite 2000 to Cobas 6000 resulted in harmonized results for 12 out of 18 immunochemistries. Thus, additional

clinical evaluations and overlapping patient monitoring are required for the six less satisfactory methods before future implementation.

Acknowledgements

The authors thank Roche Diagnostics for reagent supply for this evaluation study.

Supplementary data that accompany the online version of this article are available at http://www.reference-global.com/WDG/loi/cclm

References

1. Fraser CG. Biological variation: from principles to practice. Washington, DC: AACC Press, 2001.
2. Cobbaert CM, Weykamp C, Baadenhuijsen H, Kuypers AW, Lindemans J, Janssen RT. Selection and characterization of commutable frozen human serum pools as potential secondary reference materials for lipid and apolipoprotein measurements: study within the framework of the Dutch project ''Calibration 2000''. Clin Chem 2002; 48:1526–38.
3. Baadenhuijsen H, Kuypers AW, Weykamp C, Cobbaert CM, Janssen RT. External quality assessment in The Netherlands: time to introduce commutable survey specimens. Lessons from the Dutch ''Calibration 2000'' project. Clin Chem Lab Med 2005;43:304–7.
4. Baadenhuijsen H, Steigstra H, Cobbaert CM, Kuypers AW, Weykamp C, Janssen RT. Commutability assessment of potential reference material using a split-patient-sample between-field methods (Twin-study) design: Study within the framework of the Dutch project ''Calibration 2000''. Clin Chem 2002;48:1521–5.
5. Grawitz B, Cole TG, Fresi P, Hoefner D, Leary ET, Lindemans J, et al. Multicenter evaluation of a new HDL-cholesterol reagent on Roche/Hitachi and COBAS INTEGRA® systems, Roche Diagnostics.
6. Herold DA, Fitzgerald RL. Immunoassays for testosterone in women: better than a guess? Clin Chem 2003;8: 1250–1.
7. Van Uytfanghe K, Stöckl D, Ross HA, Thienpont LM. Use of frozen sera for FT4 standardization: investigation by equilibrium dialysis combined with IDMS and immunoassay. Clin Chem 2006;52:1817–21.
8. Thienpont LN. Standardization of steroid immunoassays: in theory an easy task. Clin Chem Lab Med 1998; 36:349–52.
9. Sturgeon CM, Ellis AR. Standardization of FSH, LH and hCG: current position and future prospects. Mol Cell Endocrinol 2007;2:260–2.
10. Tai SS, Welch MJ. Development and evaluation of a candidate reference method for the determination of total cortisol in human serum using isotope dilution liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. Anal Chem 2004;76:1008–14.
11. Tai SS, Xu B, Welch MJ. Development and evaluation of a candidate reference measurement procedure for the determination of progesterone in human serum using isotope-dilution liquid chromatography/tandem mass spectrometry. Anal Chem 2006;78:6628–33.
12. Tai SS, Welch MJ. Development and evaluation of a reference measurement procedure for the determination of estradiol-17beta in human serum using isotope-dilution liquid chromatography-tandem mass spectrometry. Anal Chem 2005;77:6359–63.
13. Tai SS, Xu B, Welch MJ, Phinney KW. Development and evaluation of a candidate reference measurement procedure for the determination of testosterone in human serum using isotope dilution liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. Anal Bioanal Chem 2007;388: 1087–94.
14. Sturgeon CM, Seth J. Why do immunoassays for tumor markers give differing results: a view from the UK National External Quality Assessment Schemes. Eur J Chem Clin Biochem 1996;34:755–9.