Analytical performance and clinical concordance of the cancer biomarkers CA 15-3, CA 19-9, CA 125 II, Carcinoembryonic Antigen, and Alpha-Fetoprotein on the Dimension Vista® System

Robert H. Christenson [a,*], Denise R. Cervelli [b], Jamie Sterner [b], Lauren M. Bachmann [d], Heather Rebuck [a], Jeffery Gray [a], Walter E. Kelley [c]

[a] University of Maryland School of Medicine, Baltimore, MD, USA
[b] Siemens Healthcare Diagnostics, Newark, DE, USA
[c] University of Maryland Medical Center, Baltimore, MD, USA
[d] VCU Health System, Richmond, VA, USA

ARTICLE INFO

ABSTRACT

Objectives: We examined analytical characteristics of new CA 15-3, CA 19-9, CA 125 II, Carcinoembryonic Antigen (CEA), and Alpha-Fetoprotein (AFP) assays on the Dimension Vista® System.

Design and methods: Imprecision studies used CLSI-EP5-A2, Limit of Blank and Limit of Detection used CLSI-EP17 and measurement ranges were determined. Method comparisons were evaluated with Passing-Bablok, least-squares regression and residual plots. Reference intervals were determined and valid specimen types, lot-to-lot variability and sample storage stability were defined. Clinical monitoring patterns for each tumor marker in patients were examined.

Results: Reproducibility for each method was <6.5%. Limits of Blank and Detection were low. Comparisons between methods showed slopes ranging from 0.89 to 1.32 with low y-intercepts and scatter. Minimal lot-to-lot variability was documented; serum/plasma specimens provide valid results; sample stability at −70 °C was >9 months. Clinical monitoring patterns correlated with established methods in >89% of cases.

Conclusions: Measurement of CA 15-3, CA 19-9, CA 125 II, CEA and AFP on the Dimension Vista® System is an attractive alternative.

© 2011 The Canadian Society of Clinical Chemists. Published by Elsevier Inc. All rights reserved.

Introduction

Biochemical markers can be useful for screening, diagnosis, prognosis, monitoring, staging and overall management of cancers of the breast, pancreas, ovary, germ cell tumors and colorectal cancer. A recent review articulates National Academy of Clinical Biochemistry practice guidelines for various biomarkers in oncology [1].

Numerous studies have confirmed that CA 15-3 is among the best available serum markers for breast cancer [1]. However CA 15-3 use for screening and diagnosis is limited by low sensitivity in early-stage disease, lack of specificity and controversy regarding whether measurement benefits outcome [2]. For use in breast cancer staging, insufficient data are available for incorporation of CA 15-3 measurements along with other variables [3]. As a result, CA 15-3 is not recommended by any oncology group for screening, diagnosis, or staging of breast cancer [1]. CA 15-3 has achieved Food and Drug Administration (FDA) clearance for monitoring recurrence in breast cancer patients with advanced disease [1]. Also, even though routine use of CA 15-3 alone for monitoring treatment response is not recommended, guidelines support use of the marker to suggest treatment failure in some cases where residual disease is not readily measureable [1].

CA 125 is the best available serum marker for epithelial ovarian cancer; guidelines generally agree that the cancer marker is useful for detecting recurrence and monitoring therapy of this disease [1]. Most oncology groups also recommend CA 125 for staging ovarian cancer [1], however the American Joint Committee on Cancer does not include CA 125 levels in their current treatment and staging system [4]. Numerous studies have established the prognostic value of the rate of decrease in CA 125 after cytoreductive surgery and during cytotoxic chemotherapy [2]. In the context of screening or diagnosis, most oncology groups do not recommend CA 125 measurements because the biomarker has not demonstrated sufficient diagnostic sensitivity and specificity. However the marker may be useful in selected populations; an NIH Consensus Statement [5] advocates screening for malignancy, particularly in a clinical trial setting in women with a strong family history of ovarian cancer [6]. Also, CA 125 appears to contribute to the differential diagnosis of pelvic masses in

* Corresponding author.
  E-mail address: RCHRISTENSON@umm.edu (R.H. Christenson).

postmenopausal women [1]. CA 125 has achieved FDA clearance for monitoring patients with ovarian cancer.

According to American Society for Clinical Oncology's recommendations for the use of Tumor Markers [7], CA 19-9 can be measured at the start of treatment for locally advanced metastatic pancreatic cancer and every 1–3 months thereafter during active treatment. If there is a rise in serial CA 19-9 determinations, this may be an indication of progressive disease and confirmation with other studies should be sought [7]. CA 19-9 as a marker for pancreatic cancer is not recommended for screening, as a sole indicator of operability, or by itself as definitive evidence of disease recurrence. Also, data are insufficient to recommend the routine use of serum CA 19-9 alone for monitoring response to treatment [7].

Alpha fetoprotein (AFP) is a well established tumor marker that is used for optimal management of patients with testicular and other germ cell tumors [1]. There is virtually no controversy about the need for measurements of this marker, largely due to the close collaboration among clinicians for more than 20 years regarding use in diagnosis, prognosis and staging, detection of recurrence and therapeutic monitoring of germ cell cancer patients [8,9]. The availability of effective and salutary treatment and the fact that temporal changes in biomarker concentrations reliably reflect and predict clinical response makes pre- and post-treatment AFP measurements a mandatory part of appropriate treatment [8]. The SOR guidelines also recommend measurement of AFP in younger women to exclude a germ cell tumor [10]. AFP has achieved FDA clearance for managing patients with non-seminomatous testicular cancer.

In colorectal cancer, currently the most relevant tumor marker is carcinoembryonic antigen (CEA). The major use of CEA is for prognosis, monitoring and assessing follow-up of treatment in colorectal cancer patients [1]. Recommendations indicate that CEA should be tested every 2–3 months [11]. Also some groups advocate utilizing CEA for monitoring measurement in ovarian cancer when CA 125 is not elevated [1]. Clearly early detection with fecal occult blood assessment and colonoscopy is the key to control of colorectal cancer; a lack of diagnostic specificity and sensitivity precludes a role for CEA measurements for either screening or early diagnosis [12].

Immunoassays produced by in vitro diagnostics manufacturers are utilized for routine CA 15-3, CA 19-9, CA 125, AFP and CEA measurement [1]. A large number of different strategies are used for these measurements; a recent survey of available instruments and technologies was recently produced by the College of American Pathologists that lists available instruments and manufacturers (http://www.cap.org/apps/docs/cap_today/surveys/0608_ImmunoSurvey.pdf).

Tumor marker assays must have appropriate characteristics to optimize use in patient care. Here we examine the analytical characteristics and examine clinical concordance in appropriate populations for new CA 15-3*, CA 19-9*, CA 125 II*, AFP and CEA assays on the Dimension Vista® System. This system utilizes Luminescent Oxygen Channeling Immunoassay (LOCI®) technology [13]. LOCI involves generation of a high-intensity chemiluminescent signal via singlet Oxygen channeling in a homogeneous assay format to achieve functional sensitivity at very low-analyte concentrations. (http://www.medical.siemens.com/siemens/nl_NLDIAG/rg_marcom_FBAs/files/perspectives/Perspectives_OUS_FINAL.pdf). *Under review by US FDA; not available for sale.

Methods

CA 15-3, CA 19-9, CA 125 II, CEA and AFP measurements

These validation studies were conducted at University of Maryland School of Medicine (Baltimore, MD) and at the laboratories of Siemens Healthcare Diagnostics (Glasgow, DE). Studies assessing different specimen types were conducted at the laboratories of Siemens Healthcare Diagnostics (Glasgow, DE) and at the Virginia Commonwealth University (Richmond, VA). Appropriate Institutional Review Board approval was obtained for the study. All tumor marker measurements were quantitatively measured with the Siemens Dimension Vista® System and associated reagents (Siemens Healthcare Diagnostics Inc., Deerfield, IL) which utilize LOCI technology [13]. The strategy for immunoassay measurement of all five tumor markers is similar. An explanation of LOCI technology for detection of immunoassays analytes can be found at (http://www.medical.siemens.com/siemens/nl_NLDIAG/rg_marcom_FBAs/files/perspectives/Perspectives_OUS_FINAL.pdf).

Imprecision

Repeatability (within-run imprecision) and total reproducibility (total within-lab imprecision) were assessed for each tumor marker assay using CLSI protocol EP5-A2. Three levels of a commercial control from Bio-Rad (Hercules, CA) and one serum pool were run in duplicate, twice daily for 20 days (total $n = 80$) were included in these studies. In accordance with CLSI EP5-A2, a fully nested Model II ANOVA was performed with adjustment for estimating repeatability and total reproducibility.

Limit of blank (LOB) and limit of detection (LOD)

LOB and LOD were carried out in accordance with CLSI EP17-A. For LOB, analyte-free calibrator was used, $n = 60$ replicates. For LOD, three human samples having analyte concentrations of ~5 U/mL for CA 15-3, ~5 U/mL for CA 19-9, ~3 U/mL for CA 125, ~1 ng/mL for CEA and ~3 ng/mL for AFP were serially diluted with analyte free calibrator and run in triplicate over 5 runs ($n = 45$ for each tumor marker). LOB data were analyzed in accordance with methodology described in CLSI EP17-A.

Analytical measurement range (AMR)

AMR for the CA 15-3, CA 19-9, CA 125 II, AFP and CEA Vista assays were determined with serial dilution mixtures of high and low serum pools for each analyte, run in triplicate. The results were analyzed by least squares linear regression.

Method comparison

Measurements using serum samples for all five tumor markers were compared to an established FDA-cleared method and analyzed by Passing-Bablok regression analysis, residual plots and least squares linear regression analysis. CA 15-3, CA 19-9, and CA 125 II measurements on the Vista were compared to the ADVIA Centaur® (Siemens Healthcare Diagnostics Inc., Deerfield, IL) and the CEA assay was compared to the Beckman Access® (Beckman Coulter, Inc, Brea, CA). The AFP method was compared to the Abbott AxSym® (Abbott Park, IL). The ADVIA Centaur®, Beckman Access® and Abbott AxSym® systems were used in strict accordance with the manufacturer's instructions. No statistical or other data outliers were removed from the analysis or plots reported in this study.

Lot-to-lot variability

Variation between lots was examined for each of the five tumor marker assays using least squares linear regression. The goal for agreement between lots was 10%.

Specimen type (serum vs. plasma)

Matched serum and Li-heparin, Na-heparin or EDTA plasma samples were used for examining valid specimen types for the tumor marker assays. Both fresh and frozen samples were analyzed. Fresh samples were tested within 48 h of the date of collection; frozen

samples (−70 °C) were tested within 1 year from the date of collection. Prior to testing, samples were thawed at room temperature for 30 min and re-centrifuged at 1500 ×g for 2–5 min. Results were compared by least squares linear regression.

Upper limit (97.5th percentile of normal) reference interval

All reference interval determinations were conducted using serum samples from apparently healthy subjects >18 years of age and having no history of cancer. CA 19-9 reference intervals were estimated using fresh serum samples obtained from a commercial blood bank. CA 15-3 and CA 125 II reference intervals were estimated using fresh sera from female donors only. These collections were obtained from a commercial blood bank and women volunteers in the Glasgow, DE area. CEA reference intervals were determined in serum and were reported based on smoking status; smokers may have values up to twice the normal range for non-smokers (http://www.labtestsonline.org/understanding/analytes/cea/faq.htmls accessed 1-19-11). For the smokers, fresh samples were obtained from a commercial blood bank; for non-smokers the samples were frozen samples leftover from a reference range study conducted earlier by Siemens Healthcare Diagnostics Inc. AFP reference intervals were estimated using samples from male blood donors and were obtained from a commercial blood bank. All 97.5th percentile reference intervals were reported using non-parametric statistics.

Storage stability

Samples collected in different matrices were stored at various temperatures to examine the stability of each tumor marker. There are no standard recommendations or protocol through CLIS or a like organization for examining sample stability on storage and the various tumor markers were developed and validated in different time frames. Therefore note that the conditions examined in some cases were somewhat different. Tumor marker values were measured at baseline (day zero), followed by aliquoting into freezer vials and storage. Follow up measurement was performed for CA15-3, CA-19-9 and CA125 II after storage for at 1 year; follow up measurement for AFP and CEA was performed after 120 days. The sample matrices tested for each tumor marker, the initial baseline (day zero) value, and the temperatures examined are displayed in Table 4.

Clinical concordance for monitoring

An important application for tumor markers is monitoring the temporal pattern for assessment of stability, a rise or a fall in values. Clinical concordance of Vista measurements for monitoring CA 15-3, CA 19-9, and CA 125 II was examined by comparison with the respective ADVIA Centaur® assays. For CEA, the Vista measurements were correlated with results from the Beckman-Coulter Access® 2 CEA assay. For AFP, the Vista measurements were correlated with results from the Abbott AxSym®. A total of 70 to 75 patients were examined for each tumor marker assay; each of the patients had between three and eight serial samples over time, mean 3.5.

Tumor marker patterns for the Vista® system and comparison (predicate) technology were categorized as Rising, Falling or Stable over time. Tumor marker patterns were categorized as Rising or Falling if they exceeded the reference change value (RCV), expressed as

$$RCV = 2^{1/2} {}^{*}Z^{*}\left(CV_A^2 + CV_I^2\right)^{1/2};$$

For this equation Z is the z-score, CV_A is the total CV of the assay and CV_I is the biological variation. For CA19-9, the CV_I was 27.2% [14] and the RCV was 76.0% for the Vista and 77.8 for the Centaur; for

CA15-3, the CV_I was 6.2% [15] and the RCV was 9.2% for the Vista and 20.3% for the Centaur; for CA125, the CV_I was 24.7% [15] and the RCV was 68.8% for the Vista and 69.6% for Centaur; for CEA, the CV_I was 12.7% [16] and the RCV was 36.2% for the Vista and 36.7% for Access 2; for AFP, the CV_I was 12% [17] and the RCV was 33.7% for the Vista AFP and 37.7% for AxSym.

Tumor marker patterns were categorized as stable if they did not exceed the RCV. Concordance between categorized serial temporal pattern for the predicate assay measurements, i.e. Rising, Falling or Stable, were compared to the Vista values with corresponding 95% confidence interval (CI) for each tumor marker and cancer type.

The RCV method was performed so that the methods would be compared using a common benchmark. Although the RCV as described above was used for categorizing patterns, note that use of the RCV is not a cleared claim for the ADVIA Centaur CA 15-3, CA 19-9 or CA 125 II assays by the US Food and Drug Administration. The actual claims for the three comparison Centaur assays are as follows:

ADVIA Centaur CA 15-3   Change of >15%
ADVIA Centaur CA 19-9   Change of >21%
ADVIA Centaur CA 125   Change of ≥30% and final value >35 U/mL.

Neither the Abbott AxSym® AFP method nor the Beckman-Coulter Access® 2 CEA method make an FDA claim as to the % change that is considered significant.

Data analysis

SAS® System software version 9.1 (SAS Institute, Cary, NC) or Analyse-it® software version 1.71 (Analyse-it Software Ltd, Leeds, Eng) was utilized for analysis.

Results

Imprecision

Table 1 shows that the repeatability results for the three levels of commercial quality control demonstrated CVs that were <3% for the CA 15-3, CA 19-9, CA 125 II, AFP and CEA LOCI assays. Within-Lab reproducibility for the three commercial controls was between 1.5% and 4.0% for the tumor marker assays. The serum pool control had somewhat lower concentration values for C15-3, CA 19-9 and CA 125 II; repeatability and within-lab reproducibility for these assays were 1.0%–4.1% and 1.9%–6.4%, respectively. CEA and AFP concentrations in the serum pool tested were higher than for the commercial controls; repeatability and within-lab CVs were lower than for the commercial controls at about 1.0% and 2.5% for the assays, respectively.

All the cancer biomarkers have imprecision studies within the reference interval characterized except AFP. The 97.5th percentile of AFP values was 8.1 ng/mL where the lowest control is at the level of 12.08 ng/mL and had within-lab reproducibility of 2.4%.

LoB, LoD and AMR

Table 2 lists the LoB, LoD and AMR for the tumor marker assays. These data indicate that the markers are reliably detected at vary low values with the LOCI technology utilized, and that the assays can measure the analytes over several orders of magnitude.

Lot-to-lot variability

The following least squared linear regression equations compared the initial lot termed 'Lot 1' with a second lot termed 'Lot 2' for each of the tumor markers. For CA 15-3, Lot 1 = 1.00 × Lot2 + 1.37 U/mL, r = 0.996; for CA 19-9, Lot 1 = 1.08 × Lot 2 − 1.50 U/mL, r = 0.999; for CA 125 II, Lot1 = 1.01 × Lot2 − 1.69 U/mL, r = 0.999; for AFP, Lot 1 = 1.00 × Lot 2 +

Table 1
Reproducibility of Vista LOCI tumor marker assays.

Tumor marker	Mean value	Repeatability, (within-run) % CV	Within-lab, reproducibility % CV
CA 15-3			
Commercial Level 1	24.5 U/mL	2.8	3.9
Commercial Level 2	66.6 U/mL	2.6	3.8
Commercial Level 3	164.1 U/mL	2.4	3.8
Serum Pool	13.6 U/mL	3.0	5.3
CA 19-9			
Commercial Level 1	58.5 U/mL	2.8	4.0
Commercial Level 2	170.6 U/mL	1.3	2.6
Commercial Level 3	420.7 U/mL	1.3	2.3
Serum Pool	11.4 U/mL	4.1	6.4
CA 125 II			
Commercial Level 1	24.9 U/mL	2.0	2.7
Commercial Level 2	61.9 U/mL	1.5	3.2
Commercial Level 3	165.4 U/mL	1.8	3.4
Serum Pool	11.0 U/mL	2.6	4.0
AFP			
Commercial Level 1	12.1 ng/mL	1.1	1.9
Commercial Level 2	83.9 ng/mL	1.2	2.3
Commercial Level 3	55.8 ng/mL	1.5	2.1
Serum Pool	237.6 ng/mL	1.0	2.4
CEA			
Commercial Level 1	2.0 ng/mL	1.7	2.4
Commercial Level 2	15.9 ng/mL	1.1	1.5
Commercial Level 3	36.6 ng/mL	0.8	1.5
Serum Pool	229.3 ng/mL	1.2	1.9

0.33 ng/mL, $r = 1.000$ and for CEA, Lot1 $= 1.06 \times$ Lot2 $- 2.05$ ng/mL, $r = 0.999$. These data indicate that lot-to-lot performance for measurement of these tumor markers was within the 10% specification and therefore acceptable.

Specimen types

Comparisons of tumor marker results using serum and matched specimens collected in phlebotomy tubes containing various antico-agulants are listed in Table 3. All of the anticoagulant specimen types tested demonstrated acceptable performance compared to the serum reference.

Reference intervals

Reference intervals for the 5 tumor markers were determined. For CA 15-3, values for the all female cohort ($n = 181$) ranged from 3.3 to 42.1 U/mL. The 97.5th percentile of values was 31.8 IU/mL. The CA 19-9 reference population ranged from 0.0 U/mL to 430.4 U/mL; the 97.5th percentile of data was 29.8 U/mL. The CA 125 II reference interval population (all women; $n = 198$) ranged from 1.9 to 68.9 U/mL. The 97.5th percentile was 39.3 U/mL. For estimating the AFP reference interval, an all male cohort ($n = 231$) yielded values that ranged from 0.6 to 10.2 ng/mL. The 97.5th percentile of AFP values was 8.1 ng/mL. CEA reference interval data was determined for smokers and non-

Table 2
Analytical characteristics.

Tumor marker	Limit of blank	Limit of detection	AMR
CA 15-3	0.3 U/mL	1.0 U/mL	1.0–300 U/mL
CA 19-9	1.0 U/mL	2.0 U/mL	2.0–1000 U/mL
CA 125 II	0.5 U/mL	1.5 U/mL	1.5–1000 U/mL
CEA	0.12 ng/mL	0.2 ng/mL	0.2–1000.0 ng/mL
AFP	0.2 ng/mL	0.5 ng/mL	0.5–1000.0 ng/mL

Table 3
Examination of the tumor markers measurements in different sample types.

Tumor marker	Comparison serum vs.	Number of samples	Range	Slope	y-intercept	R
CA 15-3	Li-Heparin	66	3.0–240	0.99	0.48	0.999
	EDTA	70	3.0–240	0.97	0.54	0.999
CA 19-9	Li-Heparin	53	2.3–971.5	0.99	1.1	0.998
	EDTA	57	2.3–971.5	0.96	3.0	0.998
CA 125 II	Li-Heparin	87	1.7–921.4	0.93	0.58	0.999
	EDTA	65	1.7–896.1	1.00	−0.34	0.999
AFP	Li-Heparin	70	0.9–999.5	1.00	−0.45	0.996
CEA	Li-Heparin	54	1.2–992.6	1.00	1.45	0.997
	Na-Heparin	54	1.2–992.7	0.99	2.55	0.998

smokers because values vary according to tobacco use [18]. For non-Smokers ($n = 198$) the CEA range was 0.0–4.9 ng/mL; the 97.5 percentile was 3.9 ng/mL. For smokers ($n = 149$) the CEA range was 0.5–10.1 ng/mL; the 97.5th percentile was 7.7 ng/mL.

Method comparison

Figs. 1 through 5 display Passing-Bablok plots (Panel a) and residual plots (Panel b), i.e. the difference between the measured point and the regression line for method comparison of each tumor marker assay.

Fig. 1 shows CA 15-3 method comparisons, which included 234 samples ranging from 0.80 to 193.9 U/mL. The Passing-Bablok slope was 0.91 (95% Confidence Interval [CI]: 0.89–0.93) with a constant bias of 2.92 U/mL (95%CI: 2.61–3.29). The residual plot (Fig. 1, panel b) shows that the measured sample pairs were distributed approximately equally about the regression line, with about 10% of samples showing variability between the measured points. The

Table 4
Table stability data for the five tumor markers.

		Baseline (day zero) values	Recovery (stability)	
CA 19-9			−20 °C	−70 °C
	Serum	18.0, 317.8, 628.3 U/mL	100–112%	100–107%
	Li-Heparin	7.1, 19.3, 24.2, 318.1, 667.8 U/mL	97.0–104%	96.0–104%
	EDTA	8.3, 13.5, 18.9, 272.6, 652.6 U/mL	94.0–103%	91.0–106%
CA15-3			−20 °C	−70 °C
	Serum	19.96, 22.17, 28.89, 81.0, 185.9 U/mL	92.0–106%	93.0–103%
	Li-Heparin	14.35, 20.91, 34.20, 77.7, 183.9 U/mL	91.0–104%	92.0–105%
	EDTA	13.72, 17.98, 22.02, 79.1, 175.3 U/mL	91.0–100%	91.0–103%
CA 125 II			−20 °C	−70 °C
	Serum	4.12, 7.05, 16.82, 243.0, 648.1 U/mL	90.0–101%	92.0–102%
	Li-Heparin	3.33, 6.29, 12.34, 245.8, 635.5 U/mL	89.0–101%	90.0–100%
	EDTA	7.86, 12.66, 13.30, 246.8, 636.8 U/mL	93.0–103%	94.0–102%
AFP			4 °C	−20 °C
	Serum	1.4, 3.1, 6.2, 90.6, 453.2, 807.7 ng/mL	93.7–104.4%	99.3–107%
	Li-Heparin	1.3, 3.1, 6.1, 92.4, 453.7, 799.3 ng/mL	93.7–104.4%	99.3–107%
CEA			−20 °C	−70 °C
	Serum	1.0, 2.0, 2.5, 65.8, 327.5, 881.7 ng/mL	98.6–110.2%	101.1–107.9%
	Li-Heparin	1.0, 2.0, 2.5, 68.0, 324.3, 889.0 ng/mL	98.6–110.2%	101.1–107.9%
	Na-Heparin	1.0, 1.9, 2.5, 60.0, 331.9, 875.9 ng/mL	98.6–110.2%	101.1–107.9%
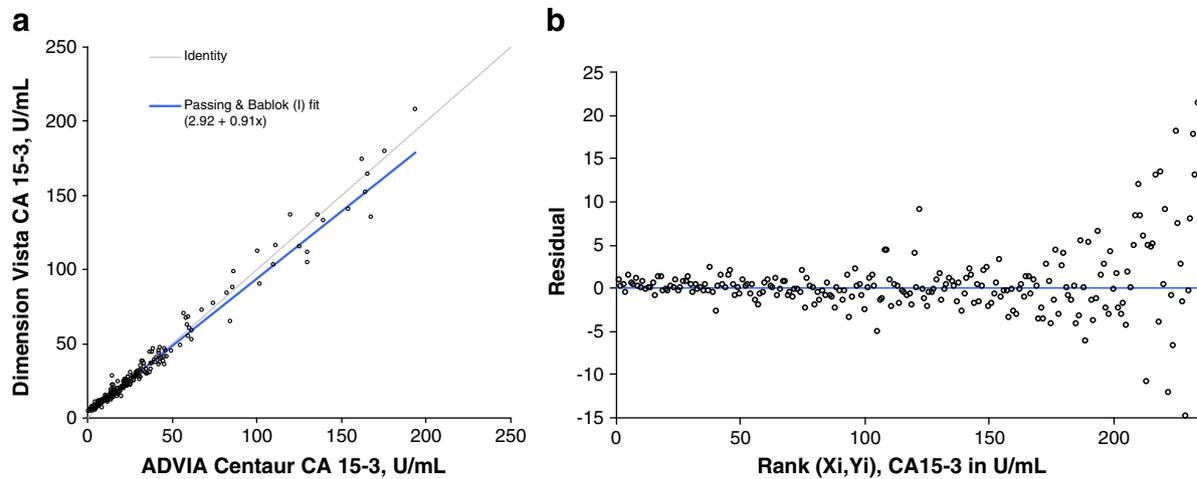
Fig. 1. Panel a shows the comparison of CA 15-3 values with Passing Bablok regression. Panel b shows the residual plot for the variation of paired sample values about the Passing Bablok regression line.

least squares linear regression equation was Vista = 0.95*ADVIA Centaur + 2.18 U/mL, Sy/x = 5.01 U/mL, r = 0.989.

Fig. 2 shows the CA 19-9 method comparison, which included 163 samples ranging from 4.04 to 653.91 U/mL. The Passing-Bablok slope was 1.07 (95% CI: 0.97–1.19) with a constant bias of −5.38 U/mL (95%CI: −7.00 to −3.89). The residual plot (Fig. 2, panel b) shows that the measured sample pairs were distributed approximately equally about the regression line, with about 10% of samples showing variability between the measured points. The least squares linear regression equation was Vista = 1.02*ADVIA Centaur + 5.26 U/L, Sy/x = 71.41 U/mL, r = 0.769. It is noteworthy that 5–10% of individuals do not express CA 19-9; we did not specifically exclude these individuals, however none of the samples showed evidence of a lack of CA 19-9 expression.

Fig. 3 shows the CA 125 II method comparison, which included 242 samples ranging from 2.2 to 518.9 U/mL. The Passing-Bablok slope was 1.32 (95%CI: 1.27–1.37) with a constant bias of 0.88 U/mL (95%CI: 0.44–1.37). The residual plot (Fig. 3 panel b) shows that the measured sample pairs were distributed approximately equally about the regression line with about 15% of samples showing variability between the measured

points. The least squares linear regression equation was Vista = 1.16*ADVIA Centaur + 4.05 U/mL, Sy/x = 18.42 U/L, r = 0.983.

Fig. 4 shows the AFP method comparison included 115 samples ranging from 2.0 to 984.8 ng/mL. The Passing-Bablok slope was 0.92 (95% CI: 0.91 to 0.93) with a constant bias of −0.45 ng/mL (95%CI: −0.58 to −0.25). The residual plot (Fig. 3, panel b) shows that the measured sample pairs were distributed approximately equally about the regression line with about 10% of samples showing variability between the measured points.. The least squares linear regression equation was Vista = 0.89*ADVIA Centaur + 1.82 ng/mL, Sy/x = 18.75 ng/mL, r = 0.997.

Fig. 5 shows the CEA method comparison that included 120 samples ranging from 0.50 to 938.33 ng/mL. The Passing-Bablok slope was 1.04 (95% CI: 1.02 to 1.07) with a constant bias of 0.04 ng/mL (95%CI: −0.28 to 0.62). The residual plot (Fig. 3, panel b) shows that the measured sample pairs were distributed approximately equally about the regression line, with about 40% of samples showing variability between the measured points. The least squares linear regression equation was Vista = 1.02*ACCESS + 5.60 ng/mL, Sy/x = 21.17 ng/mL, r = 0.995.
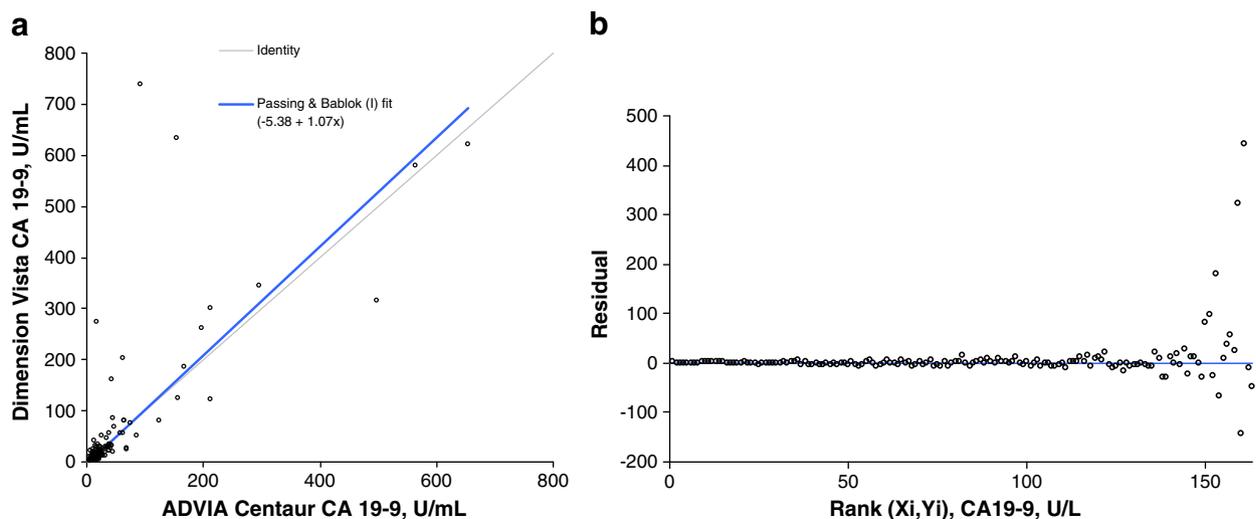
Fig. 2. Panel a shows the comparison of CA 19-9 values with Passing Bablok regression. Panel b shows the residual plot for the variation of paired sample values about the Passing Bablok regression line.

Fig. 3. Panel a shows the comparison of CA 125 II values with Passing Bablok regression. Panel b shows the residual plot for the variation of paired sample values about the Passing Bablok regression line.
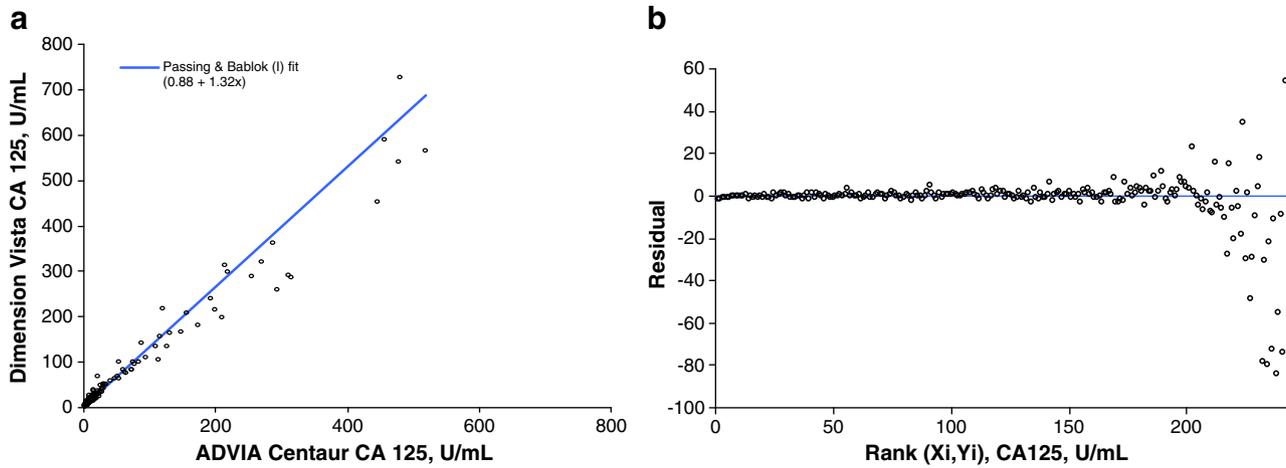
Sample stability

CA 15-3 sample recovery (stability) after 12-month storage at −70 °C was 93–103% for serum, 92–105% for Li-Heparin, and 91–103% for EDTA plasma. CA 15-3 recovery (stability) in samples maintained at −20 °C was 92–106% for serum, 91–104% for Li-heparin, and 91–100% for the EDTA plasma samples after 12 months of storage.

CA 19-9 sample recovery (stability) after 12-month storage at −70 °C was 100–107% for serum, 96–104% for Li-Heparin and 91–106% for EDTA plasma. CA 19-9 recovery (stability) for samples maintained at

Fig. 4. Panel a shows the comparison of AFP values with Passing Bablok regression. Panel b shows the residual plot for the variation of paired sample values about the Passing Bablok regression line.
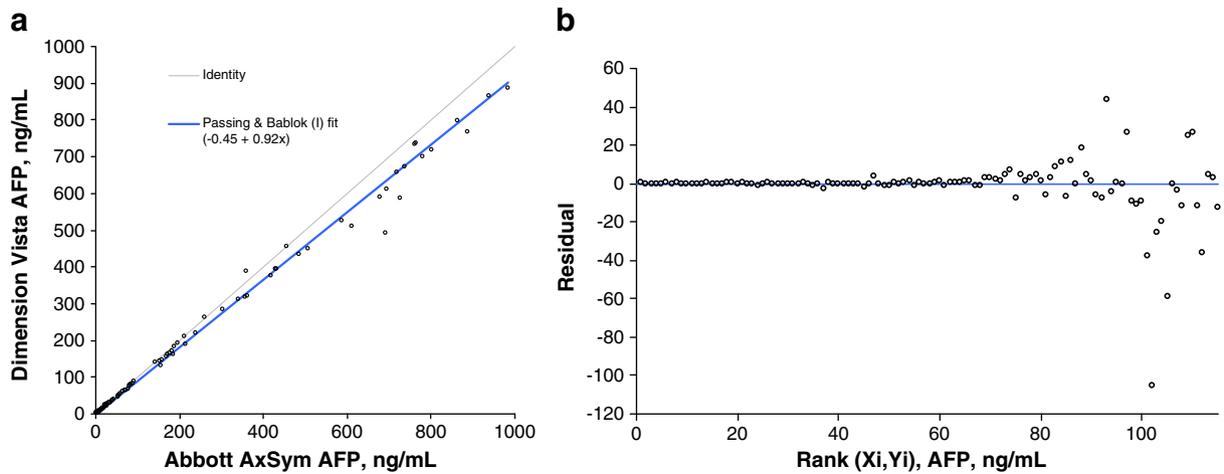

Fig. 5. Panel a shows the comparison of CEA values with Passing Bablok regression. Panel b shows the residual plot for the variation of paired sample values about the Passing Bablok regression line.
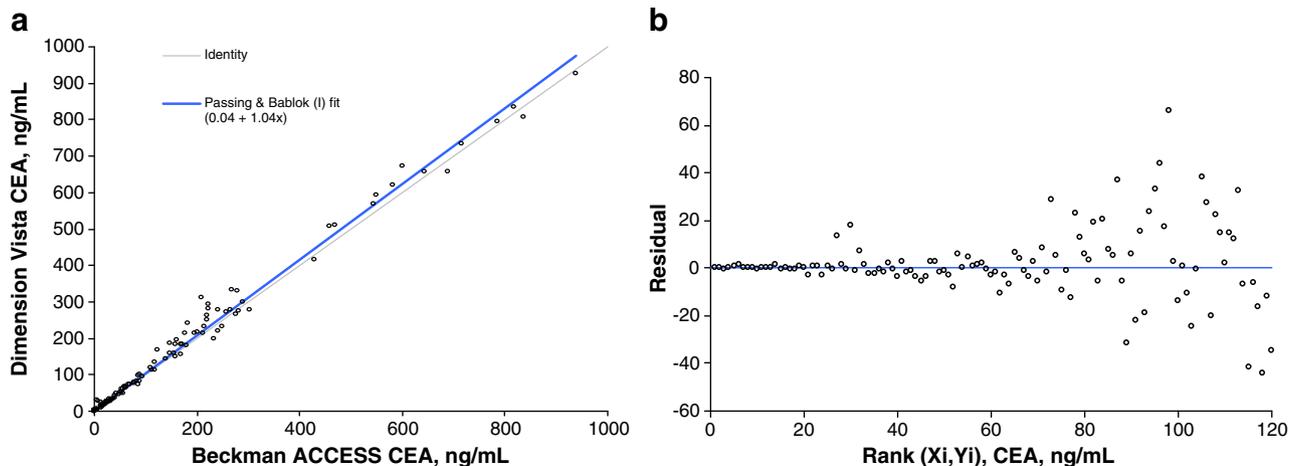
−20 °C was 100–112% for serum, 97–104% for Li-heparin and 94–103% for four of the five EDTA plasma samples after 12 months. The fifth EDTA plasma with a value of 13.5 U/mL showed recovery of 81% compared to the baseline value.

CA 125 II sample recovery (stability) after 9-month storage at −70 °C was 92–102% for serum, 90–100% for Li-Heparin, and 94–102% for EDTA plasma. CA 125 II recovery (stability) for samples maintained at −20 °C was 90–101% for serum, 89–101% for Li-heparin and 93–103% for the EDTA plasma samples after 9 months.

AFP recovery (stability) after 7 days at 4 °C and −20 °C was 93.7% to 104.4% for both serum and Li-heparin samples. AFP sample recovery (stability) for samples stored at −20 °C for 120 days was 99.3%–107%.

CEA recovery (stability) after 7-day storage at 4 °C, −20 °C and −70 °C was between 96.7% and 103.8% for the serum, Li-heparin and Na-heparin samples tested. CEA recovery (stability) after 120-day storage at −20 °C was 98.6% to 110.2% for the serum, Li-heparin and Na-

heparin samples tested. At −70 °C, CEA recovery was between 101.1% and 107.9% for the serum, Li-heparin and Na-heparin samples tested.

Patient concordance

Fig. 6, panels a–e display examples of temporal measurements with both the Vista and comparison method in samples from patients with breast, pancreatic, ovarian, non-seminomatous testicular and colorectal cancer, respectively. Fig. 6 shows that the Vista® and comparison method show similar temporal patterns that are either rising, falling or stable for each of the patients.

Table 5 lists the overall data for cohorts of patients with various cancers. Overall, monitoring patterns agreed for CA 15-3 in 89% of breast cancer patients, for CA 19-9 in 92% of the pancreatic cancer patients, for CA 125 II in 98% of ovarian cancer patients, for AFP in 90%
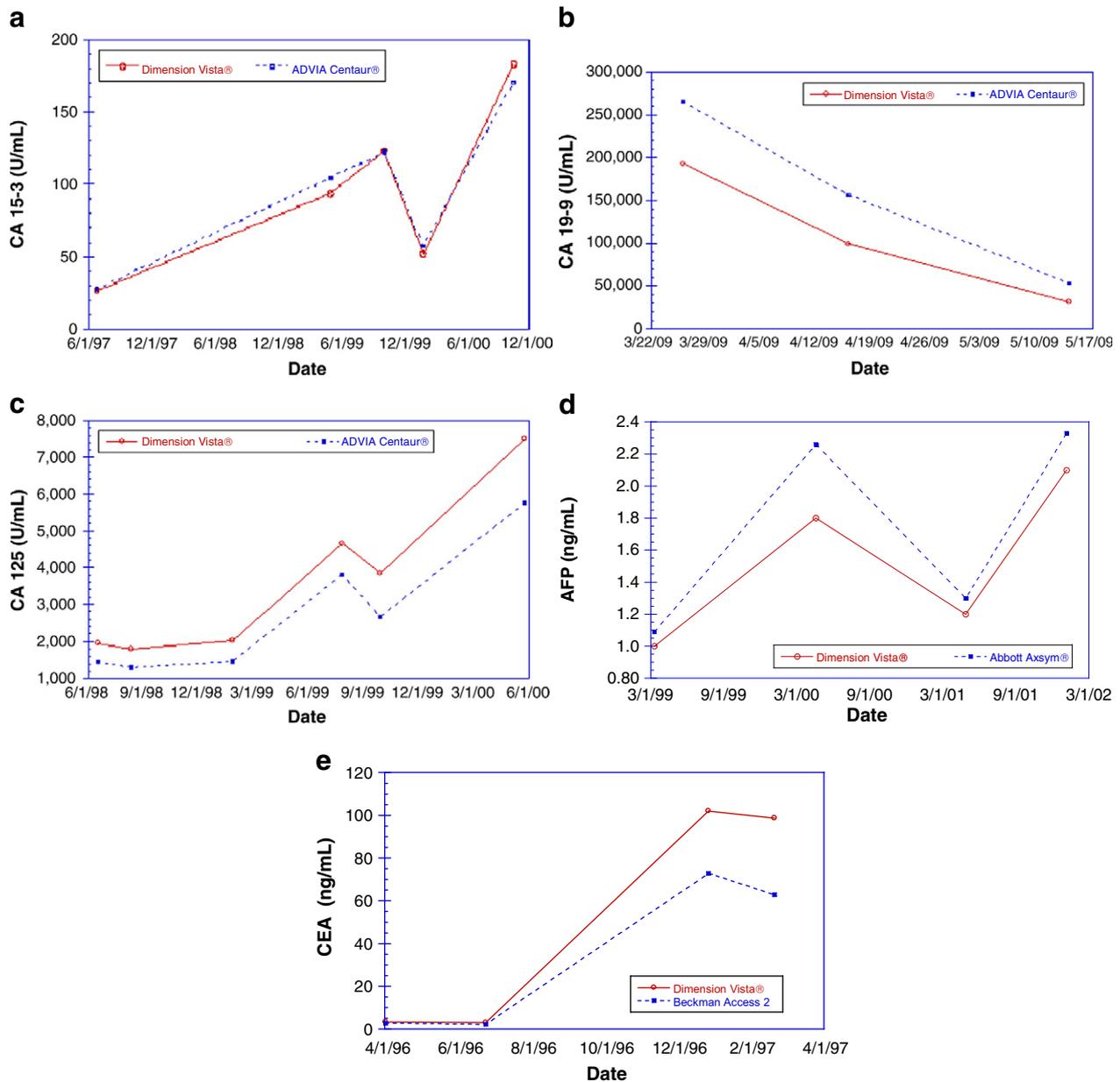

Fig. 6. Panel a. CA15-3 serial values from one breast cancer patient over a 3-year period. Panel b. CA19-9 serial values from one pancreatic cancer patient over a 2 month time period. Panel c. CA125 serial values from one ovarian cancer patient over a 2 year period. Panel d. AFP serial values from one non-seminomatous testicular cancer patient over a 3 year period. Panel e. CEA serial values for a colorectal cancer patient over a 1-year time period.

Table 5
Comparison of concordance of the LOCI tumor markers with predicate assay (see methods) in cohorts of cancer patients.

Variable	CA 15-3	CA 19-9	CA 125 II	CEA	AFP
Number of patients	75	75	75	74	70
Type of cancer	Breast	Pancreatic	Ovarian	Colorectal	Non-seminomatous testicular
Stage (# of patients)	I (11)	I (8)	I (12)	I (6)	I (20)
	II (39)	II (21)	II (4)	II (10)	II (20)
	III (21)	III (10)	III (44)	III (27)	III (16)
	IV (2)	IV (24)	IV (11)	IV (28)	IV (13)
	Unstaged (2)	Unstaged (12)	Unstaged (4)	Unstaged (3)	Unstaged (1)
Treatment[a]					
Chemotherapy	70	59	73	63	64
Radiation	63	43	11	21	11
Surgery	66	46	54	45	36
Hormone/immune	103	47	28	–	–
Other	59	49	27	–	13
No treatment		2	2	5	–
Not available					2
Age in years	22 to 76	41 to 86	20 to 86	36 to 86	1 to 54
Samples for monitoring	258	196	255	217	244
Number rising (concordant, %)	115 (106, 92%)	26 (22, 85%)	66 (64, 97%)	64 (55, 86%)	34 (34, 100%)
Number falling (concordant, %)	69 (58, 84%)	5 (4, 80%)	39 (38, 97%)	29 (24, 83%)	66 (54, 82%)
Number stable (concordant, %)	74 (65, 88%)	165 (155, 94%)	150 (147, 98%)	124 (116, 94%)	144 (131, 91%)
Overall concordance (95% CI)	89% (84.9–92.6)	92% (87.7–95.6)	98% (95.0–99.1)	90% (85.1–93.5)	94% (90.1–96.5)

[a] Individual patients received multiple treatments.

of non-seminomatous testicular cancer patients and for CEA in 94% of colorectal cancer patients.

Discussion

Use of automated immunoassay analyzers for tumor marker measurements (and other commonly requested tests) confers responsibility for analytical quality largely on the diagnostics industry [19]. From the perspective of laboratory professionals, studies examining and validating the characteristics of new tumor marker assays from commercial manufacturers represents vitally important information. Here we report key performance characteristics for new assays for measurement of CA 15-3, CA 19-9, CA 125 II, AFP and CEA that utilize LOCI technology and is available on the Vista system. The characteristics of technology reported here demonstrate that this technology is an attractive alternative for tumor marker measurements. It is noteworthy that due to variability among tumor marker measurements, the same method should be used for individual patients. When changing methods, thorough studies to re-baseline individual patients should be conducted in collaboration with clinicians due to possible methodological differences.

Imprecision is a central performance variable for clinical assays and was rigorously characterized according to CLIS EP5-A2 for all five tumor marker assays using both commercial controls and a serum pool that covered a wide range of concentrations. The repeatability (within run CV) for the tumor markers was between 0.8% and 4.2% and within laboratory reproducibility (total CV) for all of the assays were with in the range of 1.5–6.5% for all concentrations tested (Table 1). These imprecision levels are well within acceptable levels for immunoassays and indicate appropriate performance for clinical use.

Reference intervals for appropriate populations are important to define for tumor marker assays. Here the upper reference limit for each marker was defined as the 97.5th percentile for each of the assays and were similar to previously established values [18]. For CA 15-3 the upper limit of the reference interval of 31.8 IU/mL where a standard text lists <30 IU/mL [18]; for CA 19-9 we found an upper limit of 29.8 U/mL and a standard text listed <35 U/mL [18]; the CA

125 II reference interval determined was 39.3 U/mL versus <35 U/L [18]; for AFP the upper limit of the reference interval of an all male cohort was 8.1 ng/mL versus <15 ng/mL [18] and for CEA the upper reference interval for non smokers was 3.3 ng/mL versus <5 ng/mL [18]; for smokers the upper CEA reference limit was 8.8 ng/mL versus a 97.5th percentile of 5.1–10 ng/mL [18]. As with any laboratory test, values may vary between populations and each laboratory should validate their local reference intervals for these tumor markers.

The LOCI technology used for detection in the CA 15-3, CA 19-9, CA 125 II, AFP and CEA assays demonstrated low LoB and LoD values that corresponded to a small proportion of the upper limit of the reference limit for each respective assay (Table 2). The tumor marker assays also demonstrated a wide measurement range (without sample dilution); the AMRs were 8- to 100-fold higher than the upper reference limit for the tests. These LoB, LoD and AMR characteristics indicate that the assays are appropriate for clinical use.

Laboratories may encounter a number of different sample types. Therefore validating various sample types is important for measurement of tumor markers and other analytes. Validation studies showed that serum, Li-heparin plasma or EDTA plasma are appropriate for measurement of CA 15-3, CA 19-9 or CA 125 II (Table 3). Serum or Li-heparin plasma is appropriate for measurement of AFP and either serum, Li-heparin or Na-heparin plasma are appropriate for measurement of CEA. This flexibility in sample type may be useful for routine measurement in laboratory.

Laboratories performing tumor marker testing may also encounter several lots of reagent and there is need to characterize the lot-to-lot difference in tumor marker values. The acceptable difference between lots was 10%; minimal differences between reagent lots were found for each of the CA 15-3, CA 19-9, CA 125 II, AFP and CEA tumor markers examined here. Minimal lot-to-lot difference is particularly important for tumor marker assays since serial long-term temporal monitoring is typical for these markers.

Agreement of the Vista tumor marker measurements was compared to established FDA-cleared technologies using three approaches: Passing Bablok regression, residual plots and least squares linear regression analysis. Each approach has its advantages; Passing Bablok regression allows for error in both the comparison and

tested methods. On the other hand, least squares linear regression analysis is the most conservative approach in that this technique attributes all of the error in the analysis to the new method (plotted on the *y*-axis). Residual plots allow clear visual inspection of the data distribution around the regression line from Passing Bablok analysis. The method comparisons for CA 15-3, CA 19-9, AFP and CEA all showed values for slope (proportionality) that were within 10% of the comparison method for both Passing Bablok and least squares linear regression techniques. CA 125 II showed a somewhat higher slope of 1.32 by Passing Bablok and 1.16 by least squares regression. With the exception of CEA, all of the tumor markers had a constant bias (Passing Bablok) and *y*-intercept values (least squares linear regression) that were small relative to the reference interval for each marker. The bias differences and outlier points observed between manufacturers of tumor marker assays, and indeed all immunoassays, are complicated to explain. Factors that can contribute to constant bias may be the fact that no reference method for tumor markers exist; also no pure homogenous analyte materials for any of the cancer marker methods are available. Differences in calibration scheme/style also may also contribute. Most likely all of these factors (and others) contribute to method differences.

Although all of the assays showed acceptable agreement by Passing Bablok and least squares regression analysis, it is noteworthy that the residual plots showed that between 10% and 40% of the samples indicated variability (Fig. 1–5, panel b). This information underscores the need for using the same methodology when monitoring tumor marker values in patients. When changing methods for a tumor marker used for monitoring, it is critical to perform the old and new methods in parallel for patients to establish temporal patterns.

Laboratories may archive samples for tumor marker measurement for later performance. Sample stability for the CA 15-3, CA 19-9, CA 125 II, AFP and CEA assays examined was evaluated at 4 °C, −20 °C, and −70 °C. CA 15-3 and CA 19-9 were stable when stored at −20 °C or −70 °C for 12 months; CA 125 II was stable at −20 °C or −70 °C when stored for 9 months. Both AFP and CEA were stable when stored at 4 °C for 7 days; AFP and CEA measurements were stable after storage at −20 °C for 120 days, and CEA was shown to be stable for 120 days when stored at −70 °C.

Monitoring patients with serial tumor marker measurements over time can provide important clinical insight into the disease recurrence, remission or stability [1]. The data presented here show that the LOCI technology for CA 15-3, CA 19-9, CA 125 II, CEA and AFP demonstrated acceptable clinical correlation with comparison methods in patients with regard to rising, falling, and stable temporal monitoring patterns. Direct comparison of temporal patterns between the Vista assays and the established tumor marker assays agreed in 89% to 98% of patients (Table 5). Although these data indicate high concordance between the assays, the same method should be used for monitoring individual patients. When changing tumor marker assays, parallel performance to establish the biomarker values for individual patients is important to assure that clinicians understand the behavior of new assays relative to previous measurements.

In conclusion, measurements of CA 15-3, CA 19-9, CA 125 II, AFP and CEA with LOCI technology on the Vista® analyzer demonstrated acceptable imprecision, low LoB and LoD, a wide AMR and an acceptable lot-to-lot stability. The upper reference interval for each tumor marker was defined, as was the acceptable specimen types for each of the assays. The five Vista assays were compared to established methods finding acceptable correlation with each. Specimen stability on refrigerated or frozen storage was defined for each assay. Temporal samples for individual patients were examined in parallel with the Vista method and established method, finding acceptable correlation with regard to rising, falling or stable patterns.

References

[1] Sturgeon CM, Duffy MJ, Stenman U-H, et al. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. Clin Chem 2008;54(12): e11–79.

[2] Anon. Practice guidelines and recommendations for use of tumor markers in the clinic.

[3] Yarbro JW, Page DL, Fielding LP, Partridge EE, Murphy GP. American Joint Committee on Cancer prognostic factors consensus conference. Cancer 1999;86(11):2436–46.

[4] Yarbro JW, Page DL, Fielding LP, Partridge EE, Murphy GP. American Joint Committee on Cancer prognostic factors consensus conference. Cancer 1999;86(11):2436–46.

[5] Anon. NIH consensus conference. Ovarian cancer. Screening, treatment, and follow-up. NIH Consensus Development Panel on Ovarian Cancer. JAMA 1995;273(6):491–7.

[6] Menon U, Jacobs IJ. Recent developments in ovarian cancer screening. Curr Opin Obstet Gynecol 2000;12(1):39–42.

[7] Locker GY, Hamilton S, Harris J, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. J Clin Oncol 2006;24(33): 5313–27.

[8] Thomas G, Jones W, VanOosterom A, Kawai T. Consensus statement on the investigation and management of testicular seminoma 1989. Prog Clin Biol Res 1990;357:285–94.

[9] Vastag B. Testicular cancer seen as "stalking horse" for other cancers. J Natl Cancer Inst 1999;91(21):1801–2.

[10] Kerbrat P, Lhommé C, Fervers B, et al. Ovarian cancer. Br J Cancer 2001;84(Suppl 2): 18–23.

[11] Duffy MJ. CA 19–9 as a marker for gastrointestinal cancers: a review. Ann Clin Biochem 1998;35(Pt 3):364–70.

[12] Smith RA, Cokkinides V, von Eschenbach AC, et al. American Cancer Society guidelines for the early detection of cancer. CA Cancer J Clin 2002;52(1):8–22.

[13] Ullman EF, Kirakossian H, Singh S, et al. Luminescent oxygen channeling immunoassay: measurement of particle binding kinetics by chemiluminescence. Proc Natl Acad Sci USA 1994;91(12):5426–30.

[14] Erden G, Barazi AO, Tezcan G, Yildirimkaya MM. Biological variation and reference change values of CA 19–9, CEA, AFP in serum of healthy individuals. Scand J Clin Lab Invest 2008;68(3):212–8.

[15] Anon. Desirable Specifications for Total Error, Imprecision, and Bias, derived from intra- and inter-individual biologic variation. Available at: http://www.westgard.com/biodatabase1.htm.

[16] Ricós C, Alvarez V, Cava F, et al. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59(7):491–500.

[17] Trapé J, Botargues JM, Porta F, et al. Reference change value for alpha-fetoprotein and its application in early detection of hepatocellular carcinoma in patients with hepatic disease. Clin Chem 2003;49(7):1209–11.

[18] Burtis C, Ashwood E. Tietz textbook of clinical chemistry. 3rd ed. Philadelphia, PA: W.B. Saunders Company; 1999.

[19] Sturgeon C, Diamandis EP. Use of Tumor Markers in Testicular, Prostate, Colorectal, Breast, and Ovarian Cancers. Available at: http://www.aacc.org/members/nacb/LMPG/OnlineGuide/PublishedGuidelines/major/Documents/TumorMarkersMajor10.pdf#page=1. Accessed May 12, 2011.