Received 11.18.05 | Revisions Received 1.26.06 | Accepted 2.2.06

Validating New Reagents: Roadmaps Through the Wilderness

Roberta A. Martindale, BSc (MLS), MT(ASCP),[1] George S. Cembrowski MD, PhD,[1,2] Lucille J. Journault,[1,2] Jennifer L. Crawford, BSc (MLS),[1,2] Chi Tran,[1,2] Tammy L. Hofer, BSc (MLS), MBA,[1,2] Bev J. Rintoul, BSc (MLS)[1,2] Jean S. Der,[1,2] Cathy W. Revers, BSc (MLS),[1,2] Cheryl A. Vesso,[1,2] Carol E. Shalapay,[1,2] Connie I. Prosser, PhD, FCACB,[1,2] Donald F. LeGatt, PhD, FCACB[1,2]

([1]Department of Laboratory Medicine and Pathology, University of Alberta Hospitals, [2]Capital Health Authority, Edmonton, Canada)

DOI: 10.1309/BRC6Y37NM3BU97WX

Abstract

One of the most frequent quality control issues faced by laboratory professionals is how to respond appropriately to a shift in quality control (QC) following a reagent lot change. Possible actions include adjusting the control range, checking for shifts in patient data, or simply ignoring the QC shift. We offer a systematic approach to shifted quality control and/or patient data following a reagent lot change. We divide laboratory tests into 3 types, (1) tests for which the analysis of QC specimens is sufficient, (2) tests which demonstrate between reagent lot shifts infrequently, and (3) tests with between lot variation. Depending on the test type, specific information is gathered about the magnitude of the shifts in either the QC and/or the patient data. The control mean is reset following an isolated quality control shift. Evaluation of the shift in patient data is initiated by the laboratory director when the shift exceeds a multiple of the allowable error.

It is well known that reagent lot changes can cause significant shifts in patient results. Mueller-Hinton agar lot changes can affect drug susceptibility results.[1] In hematology, new lots of Simplate II have caused incorrect bleeding times,[2] and new lots of hematology reagent have affected automated complete blood counts.[3] In the realm of clinical chemistry, reagent lot changes have affected immunoassays more than general chemistry tests. Lot-to-lot variation has been frequently implicated in shifts of patient and quality control prostate specific antigen (PSA) values,[4,5] probably because of the extensive application of PSA testing for screening and monitoring of prostate cancer. Variation in reagent lots has resulted in significant variation in the testing of patients' hepatitis C antibody,[6] patients' human chorionic gonadotropin (hCG),[7] and theophylline quality control.[8] Even point-of-care cholesterol tests[9,10] are affected by reagent lot variation.

In some analytical systems, reagent lot changes seem to primarily affect quality control measurements. The influence of between lot variation on control product analysis is often ascribed to "matrix effects"; differences in constituent concentration and constituents in the control product and the plasma or serum specimen that the control product is imitating. This phenomenon has been frequently observed in dry reagent systems like that of the Vitros (Ortho-Clinical Diagnostics, Rochester NY). Hill and colleagues have shown that slide generation changes in the Vitros can cause significant excursion in quality control measurements but little change in re-analyzed patient specimens.[11]

Prior to the 1992 enactment of CLIA '88,[12] reagent lot validation practices were highly variable. General guidelines[8] for measuring and responding to lot-related shifts in control and patient data were not widely applied and very few laboratories validated new reagent lots by re-analyzing retained patient specimens. Section 493.1255 of the CLIA '88 Standard specifies that whenever a complete change of reagents occurs, "the laboratory . . . demonstrate that changing reagent lot numbers does not affect the range used to report patient test results, and control values are not adversely affected by reagent lot number changes."[12] Soon after the enactment of CLIA '88, the requirement for lot validation was promulgated in the accreditation checklists of various professional organizations including that of the College of American Pathologists (CAP).

The rigor in fulfilling lot validation depends on the knowledge, confidence, and prior practices of the laboratory directors and the supervisory technologists. In preparing for our first CAP accreditation inspection last winter, our senior biochemistry staff and medical biochemistry supervisors spent well over 30 hours synthesizing 3 different, test-dependent lot validations. Being heavily influenced by our quality system, we present these validations as flow charts. For one group of tests, we determined that the analysis of quality control materials was sufficient for the workup of new reagent lots. For the other tests, the investigation depended on the history of QC and patient shifts with new reagent lots. With a stable method, quality control analysis was usually adequate; with a method that frequently demonstrated patient shifts with new reagent lots, we required the analysis of retained patient specimens as well as quality control specimens. This general approach is summarized in **Table 1**.

Multiple quality control specimens must be run in order to detect the presence of a shift. To quantify a quality control shift, we suggest that 4 measurements be made at each control level. The shift is calculated by subtracting the usual control mean from the average of the controls. We suggest that the magnitude of the QC shift be assessed by comparing the control shift to the method's stable standard deviation. If the magnitude of the QC shift exceeds 1 SD, we require recalculation and resetting of the QC mean.

To quantify the shift in patient data, we suggest that 10 different previously analyzed patient specimens be selected and re-analyzed with the new reagents. If possible, the specimens should be chosen to represent the usually measured concentration range. The shift is calculated by subtracting the average of the initial results from the average of the new results. The reanalysis of 10 specimens will result in greater than 60% of significant shifts being discovered and corrected. (Personal communication, November 17, 2005, A. Srinivasan, LifeScan, Milpitas, CA). The magnitude of the patient shift is assessed by comparing the shift to a multiple of the allowable error (AE) of the method. This allowable error may be the CLIA proficiency test (PT) limit. As CLIA proficiency test limits are not available for all analytes and are sometimes regarded as too broad, alternate allowable errors based on physiologic variation are available.[13] **Table 2** shows the allowable errors that we use in our medical biochemistry laboratory. We suggest that if the patient shift exceeds 0.5 x AE, then the method is judged as unacceptable and a new reagent lot be obtained. Some laboratory professionals may elect to use more stringent criteria for maximum allowable patient shift (eg, a maximum bias of 0.33 x AE). Some clinical laboratories may require even tighter limits (eg, clinical trial laboratories). These laboratories might attempt to use allowable error limits derived from state of the art analytical performance.

Test-Specific Algorithms for Reagent Validation

I. Tests where only QC is evaluated (Unstable analyte/ Unstable reagents/Insufficient material for retesting/Extremely tedious or time-consuming)

There are at least 4 types of tests where we assess only the shifts in quality control data, even when between lot variation in reagents may cause shifts in patient data. Some tests measure very *unstable analytes* such as ACTH, insulin, vitamin A, and vitamin E. For these analytes, it would be misleading to reanalyze previously analyzed specimens with alternate set of reagents. Some tests use *highly unstable reagents* (eg, bile acids and free fatty acids). These reagents are so unstable that they need to be reconstituted daily. In order to validate such reagents with patient specimens, we would need to measure the retained specimens daily. For other tests, there may be *little or no specimen remaining* (eg, tissue iron, tissue copper, etc). Finally, some tests are so *laborious or time-consuming* that we rarely would repeat retained specimens with new reagent lots (eg, fecal fats).

Figure 1 shows the algorithm for validating new reagent lots with only quality control analysis. This validation can occur during patient testing. Briefly, quality control specimens are run in duplicate until there are 4 control observations at each level. Trouble-shooting and/or recalibration follow out of range QC or the discovery of QC shifts exceeding 1 standard deviation (SD). The 4 replicate quality control measurements are averaged with their average compared to the usual QC average. If the reagents are deemed unacceptable, reagents are remade if prepared in-house. Otherwise, manufacturer controls may be analyzed before the manufacturer is notified.

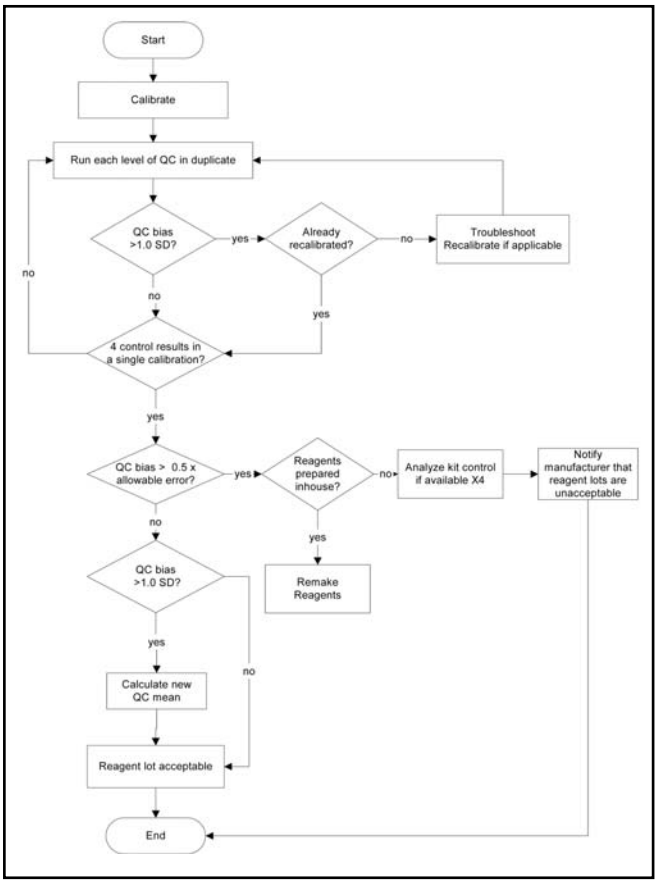

Figure 1_Tests where only QC is evaluated: Unstable analyte/Unstable reagents/Insufficient material for retesting/Extremely tedious or time-consuming.

II. QC-Based Reagent Lot Validation for Methods with Usually Clinically Unimportant Lot-to-Lot Variation

Some chemistry methods rarely demonstrate between lot variation for either control or patient results. We propose the following criterion for defining a method to be independent of lot variation: a method is stable if the QC bias is less than 1 SD for 3 successive reagent lot changes. The history of lot stability must be documented. Examples of stable methods include electrolytes, calcium, phosphorous, total protein, albumin, urea, and

Table 1_General Approach to the Investigation of New Reagent Lots

Test Method	QC Analysis	Retained Patient Specimen Analysis
Laborious/Infrequently performed/Unstable analyte	4 control replicates, each level	None
Stable, generally no QC or patient shift with new reagent lots	4 control replicates, each level	10 retained patient specimens if control shift discovered
Sometimes demonstrates clinically important lot to lot variation in patient specimens with new reagent lot	4 control replicates, each level	10 retained patient specimens

Table 2_Allowable Error (AE) Criteria

Analyte	CLIA '88	Alternate allowable error limit based on physiologic variation
Alpha-1 Antitrypsin	±3 SD	±10%
Acetaminophen		±15%
Acetone		±15%
Alpha-fetoprotein	±3 SD	±15%
Alanine aminotransferase	±20%	±30%
Albumin	±10%	±4%
Aldosterone		±40%
Alkaline phosphatase	±30%	±12%
Amikacin		±20%
Amitriptyline		±20%
Ammonia		±30%
Amylase	±30%	±16%
Androsteredione		±25%
Angiotensin converting enzyme		±20%
Apolipoprotein A1		±10%
Apolipoprotein B		±15%
Aspartate aminotransferase	±20%	±15%
Beta 2 microglobulin		±10%
Beta hydroxybutyrate		±15%
Bilirubin, direct		±45%
Bilirubin, total	±6.84 µmol/L or 20%	±35%
CA 125		±40%
CA 15-3		±20%
CA 549		±20%
CA 19-9		±50%
Calcium	±0.2495 mmol/L	±2.5%
Calcium, urine		±30%
Carbamazepine	±25%	±20%
Carcinoembryonic antigen		±25%
Carotene		±45%
Ceruloplasmin		±10%
Chloride	±5%	±1.5%
Cholesterol, HDL	±30%	
Cholesterol, LDL		±15%
Cholesterol, total	±10%	±10%
Complement C3	±3 SD	±10%
Complement C4	±3 SD	±20%
Copper		±10%
Cortisol	±25%	±30%
C-peptide		±15%
Creatine kinase	±30%	±30%
Creatine kinase, MB	±3 SD	±30%
Creatinine	±26.52 µmol/L or ±15%	±10%
Creatinine, urine		±30%
C-reactive protein		±70%
Cysteine		±10%
Desipramine		±20%
Dehydroepiandrosterone sulfate		±12%
Digoxin	±0.26 nmol/L or ±20%	±15%
Estradiol		±25%
Ethosuximide	±20%	±20%
Ethanol	±25%	±15%
Ethylene glycol		±15%
Ferritin		±20%
Folate, rbc		±30%
Free testosterone		±50%
Follicle stimulating hormone		±15%
Gentamicin	±25%	±20%
GGT		±25%
Glucose	±0.33 mmol/L or ±10%	±10%
Haptoglobin		±30%
HbA1C		±4%
Homocysteine		±20%
Human chorionic gonadotropin	±3 SD or Pos/Neg	
IgA	±3 SD	±15%
IgD	±3 SD	
IgE	±3 SD	
IgG	±25%	±10%
IgM	±3 SD	±20%
Imipramine		±20%
Insulin		±35%
Ionized calcium		±2%
Iron	±20%	±30%
Isopropanol		±15%
Lactate		±30%
Lactate dehydrogenase	±20%	±15%
Lead, blood	±0.19 µmol/L or ±10%	
Luteinizing hormone		±20%
Lipase		±30%
Lipoprotein (a)		±30%
Lithium	±0.3 mmol/L or ±20%	±20%
Magnesium	±25%	±5%
Methanol		±15%
Methotrexate		±20%
Myoglobin, Urine		±40%
N telopeptide		±30%
N-acetylprocainamide		±20%
Nortriptyline		±20%
Oxalate - urine		±50%
Phenobarbital	±20%	±20%
Phenytoin	±25%	±20%
Phosphate		±10%
Potassium	±0.5 mmol/L	±6%
Prealbumin		±15%
Primidone	±25%	±20%
Procainamide /N-acetyl procainamide	±25%	±20%
Prolactin		±22%
Protein, total	±10%	±5%
Parathyroid hormone		±22%
Pyruvate		±20%
Quinidine	±25%	±20%
Rheumatoid factor		±15%
Salicylate		±20%
Selenium		±15%
Sex hormone binding globulin		±25%
Sodium	±4 mmol/L	±1%
T_3 uptake	±3 SD	±7%
Thyroid binding globulin		±10%
Testosterone		±16%
Theophylline	±25%	±20%
Thyroxine	±12.9 nmol/L or ±20%	±10%
Thyroxine, free	±3 SD	±10%
Tobramycin		±20%
Transferrin		±5%
Tricyclic antidepressants		±20%
Triglycerides	±25%	±30%
Triiodothyronine	±3 SD	±15%
Troponin		±25%
Thyroid stimulating hormone	±3 SD	±25%
Urea nitrogen	±0.71 mmol/L or ±9%	±17%
Uric acid	±17%	±15%
Valproate	±25%	±20%
Vancomycin		±20%
Vitamin B_{12}		±16%
Vitamin C		±30%
Zinc		±15%

creatinine measured by the Beckman LX-20 and TSH measured by the Bayer ADVIA Centaur.

For these reagent lot-independent methods, we recommend that only the quality control results be evaluated with the new reagent before a new reagent lot is placed into service. For these tests, it is acceptable to analyze quality control material just after reagent lot change and then in the usual schedule. If a persistent QC shift is detected in the 48 hours following reagent lot change we then recommend that the magnitude of the patient bias also be evaluated.

Figure 2 shows the algorithm used to validate such stable tests. Because of the infrequent occurrence of significant lot dependent variation, we suggest that this validation occur during patient testing. Briefly, quality control specimens are run in the usual manner after reagent lot change. Trouble-shooting and/or recalibration follows the violation of any quality control rule. In the case of unsuccessful trouble-shooting, 10 retained patient specimens or previously analyzed retained PT materials (in the absence of retained patient samples) are reanalyzed. If the patient or PT bias is too large, the clinical biochemist or pathologist is consulted. Otherwise, the QC mean is reset if the QC has shifted by more than 1 SD.

III. Patient-Based Reagent Lot Validation for Methods With Significant Lot to Lot Variation

Some chemistry methods can demonstrate large variations in quality control and/or patient specimen testing with alternate reagent lots. Some of our core chemistry laboratory assays demonstrate this variation: troponin, hCG, and folate on the Bayer ADVIA Centaur and the enzyme and turbidimetric tests on the Beckman LX-20. Each section in medical biochemistry maintains a list of tests that usually demonstrate large variation with alternate lots of reagents. For these methods, before being placed into service, new reagent lots must evaluated for significant shifts in patient and quality control results.

Figure 3 shows the algorithm used to validate such unstable tests. After the reagent lot change, quality control specimens are run in duplicate to more rapidly obtain 4 replicate observations at each level. If the QC has not shifted, 10 retained patient specimens are analyzed with the new reagent lot and the size of the patient bias assessed. If the QC has shifted, trouble-shooting and/or recalibration are attempted. If these procedures do not restore the original QC means (indicating a QC shift), the retained patient specimens are analyzed with the new reagent lot. If the patient bias is too large, the clinical biochemist or pathologist is consulted. Otherwise, the QC mean is reset if the QC has shifted by more than 1 SD.

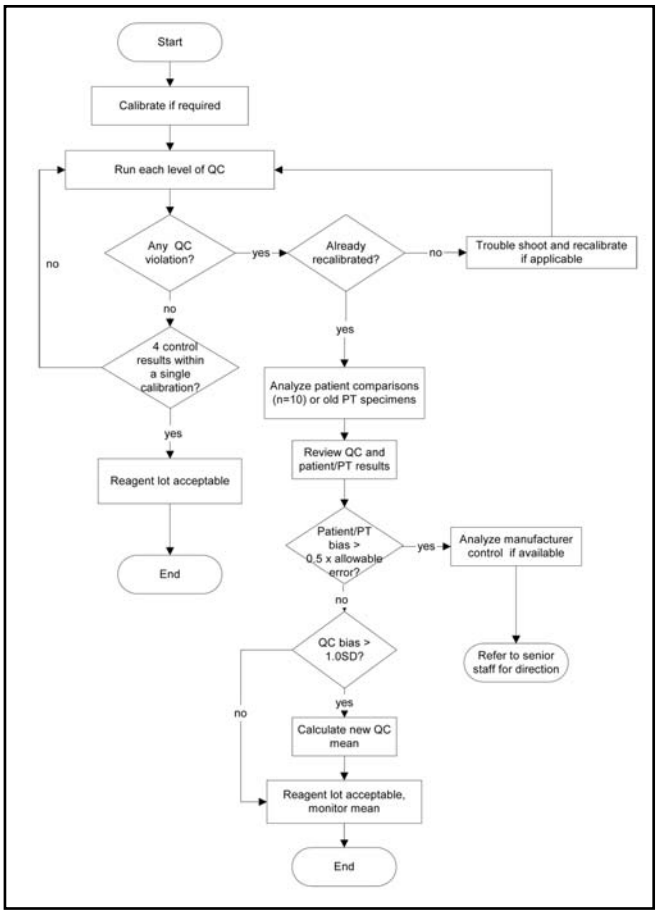

Figure 2_QC-based reagent lot validation for methods with usually clinically unimportant lot to lot variation.

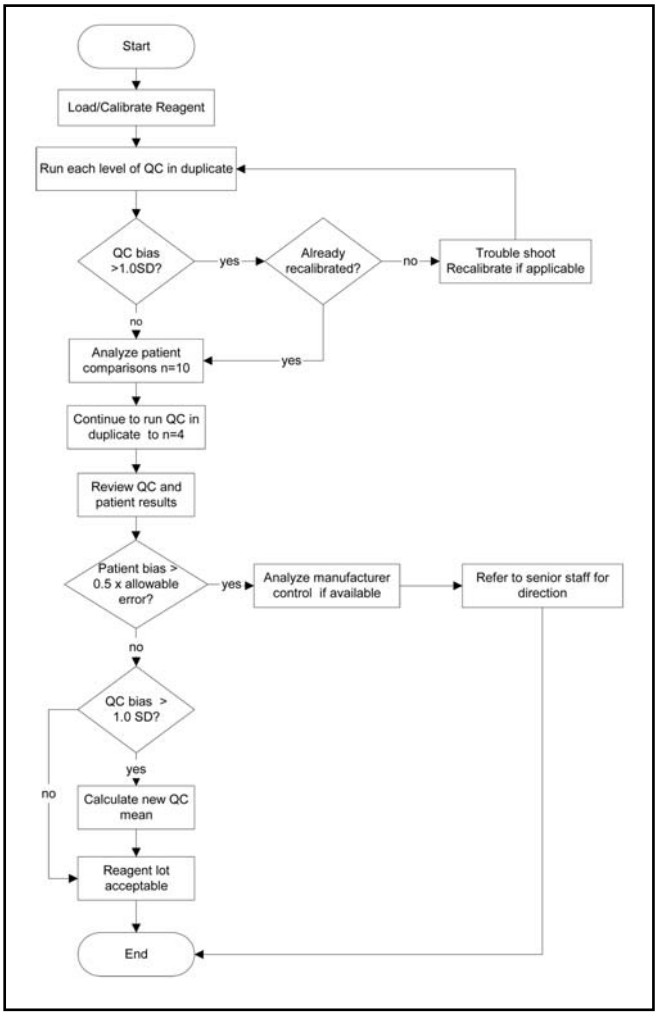

Figure 3_Patient-based reagent lot validation for methods with significant lot to lot variation.

Calculations and Record-Keeping

We designed a Microsoft Access program for entering, analyzing, and retrieving the required quality control and patient replicate values. **Figure 4** shows an output screen from this program that summarizes the reagent information as well as the requisite QC replicate data. **Figure 5** shows an example of an input screen for the entry of CK patient data (current and new lot). The "drop-down" menu of analyte names is linked with the allowable errors shown in **Table 2** and permits easy evaluation of the magnitude of the patient shift.

Discussion

We have been using this new lot qualification system for the last 16 months. For the LX-20 systems, we evaluated 103 new reagent lots. Two lots of acetaminophen were judged unacceptable; the manufacturer has just narrowed its acceptance criteria for acetaminophen variation. With respect to the other LX-20 analytes, all were acceptable; the following tests required quality control adjustments: alkaline phosphatase (2 lots), alanine transaminase (1), amylase (2), C-reactive protein (1), CSF protein (4), gamma glutamyl transferase (1), lipase (1), microalbumin (2), salicylates (1), theophylline (2), and triglycerides (1). Until recently, only troponin and vancomycin were tested with the lot qualification system; troponin required quality control adjustments with 2 lots. We are adding more Centaur immunoassay tests to our reagent qualification system.

Despite these efforts, we were unable to detect successive increases in gamma glutamyl transferase patient results. Another city laboratory discovered our higher results with a patient comparison study. We are now scrutinizing our gamma glutamyl transferase lot validations.

The system has provided us with defensible and logical criteria for remaking new reagents and even returning "defective" reagent lots. We have found that our QC adjustments are made with increased confidence. As our chemistry laboratory is large and run by 4 laboratory scientists, the flowcharts presented here have standardized the evaluation of new reagents. The Access data base program has greatly simplified our record keeping of validation of new reagent lots. LM

Acknowledgement: We wish to acknowledge the initial guidance that was provided by Ms. Elsa Quam of the University of Wisconsin Hospital Clinical Laboratories, Madison, WI.

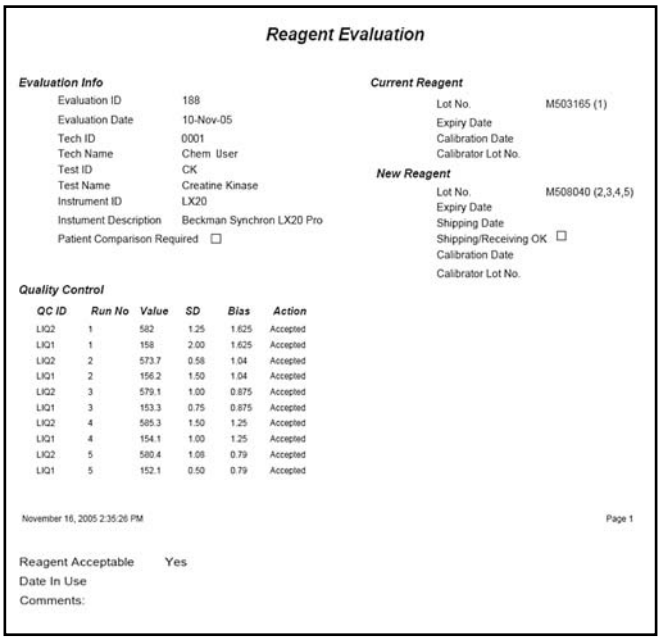

Figure 4_Output screen showing evaluation of the control data.

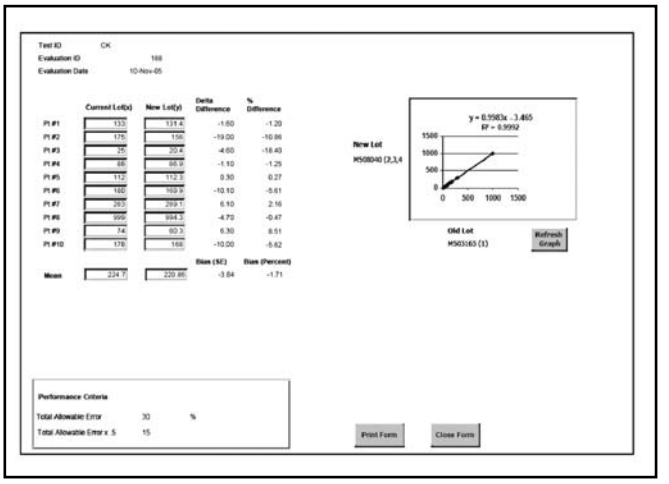

Figure 5_Input screen showing results of reagent lot comparisons using patient specimens.

1. Fuchs PC, Barry AL, Brown SD, et al. Intrepretive criteria and quality control parameters for testing of susceptibilities of *Haemophilus influenzae* and *Streptococcus pneumoniae* to trimethoprim and trimethoprim-sulfamethoxazole. *J Clin Microbiol.* 1997;35:125-131.
2. Houdijk WP. Warning: Simplate II, lack of standardization in standardized bleeding time devices—reply: Lot-to-lot variation in ejection force has been rectified. *Thromb Haemost.* 1991;66:625-626.
3. Anonymous. COULTER lyse S III diff: Reagent lot gives erroneous results. *Health Devices.* 1991;20:445-446.
4. Carey RN, Frye RM, Cook JD, et al. Between-lot/between-instrument variations of the Abbott IMx method for prostate-specific antigen. *Clin Chem.* 1992;38:2341-2343.
5. Wener MH, Daum PR, Brawer MK. Variation in measurement of prostate-specific antigen: importance of method and lot variability. *Clin Chem.* 1995;41:1730-1737.
6. Dufour DR. Lot-to-lot variation in anti-hepatitis C signal-to-cutoff ratio. *Clin Chem.* 2004;50:958-960.
7. Goodman DB, Bulley M, Hendricks M, et al. Assessment of the Abbott IMx assay system for the measurement of human chorionic gonadotropin levels in the treatment of ectopic pregnancy. *Arch Pathol Lab Med.* 1993;117:701-703.
8. Cembrowski GS, Carey NR. *Laboratory Quality Management.* Chicago: ASCP Press, 1989;102-103.
9. Gottschling HD, Reuter W, Ronquist G, et al. Multicentre evaluation of a non-wipe system for the rapid determination of total cholesterol in capillary blood, Accutrend cholesterol on Accutrend GC. *Eur J Clin Chem Clin Biochem.* 1995;33:373-381.
10. Boerma GJ, Gelderland J, van Gorp I, et al. Use of the Reflotron system for cholesterol assay in capillary blood, venous blood, and serum—evaluation of accuracy and lot-to-lot reagent comparability. *Clin Chem.* 1988;34:2117-2119.
11. Hill SA, Heathcote JC, McQueen MJ. The significance of matrix effects on the measurement of lactate dehydrogenase (LD) activity using Kodak dry slide technology in the Ontario Laboratory Proficiency Testing Program. *Clin Biochem.* 1990;23:179-181.
12. Department of Health & Human Services Health Care Financing Administration Public Health Service 42 CFR Part 405, et al. Clinical Laboratory Improvement Amendments of 1988; final rule. Federal Register Friday February 28, 1992.
13. Ricos C, Alvarez V, Cava F, et al. Current databases on biologic variation: pros, cons and progress. *Scand J Clin Lab Invest.* 1999;59:491-500. (updated in http://www.westgard.com/biodatabase1.htm)