ANNALS OF LABORATORY MEDICINE

Practical Considerations for Clinical Laboratories in Top-down Approach for Assessing the Measurement Uncertainty of Clinical Chemistry Analytes

Hyunjung Gu, M.D., Ph.D.[1], Juhee Lee, M.T.[1], Jinyoung Hong, M.D.[1], Woochang Lee, M.D., Ph.D.[1], Yeo-Min Yun, M.D., Ph.D.[2], Sail Chun, M.D., Ph.D.[1], Woo-In Lee, M.D., Ph.D.[3], and Won-Ki Min, M.D., Ph.D.[1]

[1]Department of Laboratory Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; [2]Department of Laboratory Medicine, Konkuk University School of Medicine, Seoul, Korea; [3]Department of Laboratory Medicine, School of Medicine, Kyung Hee University and Kyung Hee University Hospital at Gangdong, Seoul, Korea

Background: The top-down (TD) approach using internal quality control (IQC) data is regarded a practical method for estimating measurement uncertainty (MU) in clinical laboratories. We estimated the MU of 14 clinical chemistry analytes using the TD approach and evaluated the effect of lot changes on the MU.

Methods: MU values were estimated using subgrouping by reagent lot changes or using the data as a whole, and both methods were compared. Reagent lot change was simulated using randomly generated data, and the mean values and MU for two IQC datasets (different QC material lots) were compared using statistical methods.

Results: All MU values calculated using subgrouping were lower than the total values; however, the average differences were minimal. The simulation showed that the greater the increase in the extent of the average shift, the larger the difference in MU. In IQC data comparison, the mean values and MU exhibited statistically significant differences for most analytes. The MU calculation methods gave rise to minimal differences, suggesting that IQC data in clinical laboratories show no significant shift. However, the simulation results demonstrated that notable differences in the MU can arise from significant variations in IQC results before and after a reagent lot change. Additionally, IQC material lots should be treated separately when IQC data are collected for MU estimation.

Conclusions: Lot changes in IQC data are a key factor affecting MU estimation and should not be overlooked during MU estimation.

Key Words: Measurement uncertainty, Top-down approach, Internal quality control

INTRODUCTION

Measurement uncertainty (MU) is a concept commonly used in various industries and engineering fields but not in clinical laboratories [1]. As the importance of standardization and traceability of test results is increasing, MU is likely to become an important issue in laboratory quality management [2-4].

Since the Guide to the Expression of Uncertainty in Measure-

ment (GUM) was published in 1996 [5], the bottom-up approach is the standard method for estimating MU. This approach involves the identification of all sources of uncertainty in the measurement procedure, estimation of their magnitudes, and calculation of the combined uncertainty according to the law of error propagation [5]. However, the MU guidelines for clinical laboratories recommend that the top-down (TD) approach is practical and particularly well-suited to closed mea-

suring systems, which are common in routine clinical laboratories [6, 7]. For most measuring systems in clinical laboratories, the most significant uncertainty contributions to the overall MU are (1) long-term imprecision data obtained for internal quality control (IQC) materials for a period sufficient to include all changes to measuring conditions (u_{Rw}, within-laboratory reproducibility); (2) uncertainty of the end-user calibrator (u_{cal}) obtained from the manufacturer or established by a laboratory with its own measuring system; and (3) bias correction, if a medically unacceptable measurement bias exists [4, 6, 7].

Identifying the sources of uncertainty may be the first step in estimating the MU of a measurement system. Various measurement factors, such as sample inhomogeneity, reconstitution procedures for lyophilized materials, reagent and calibrator instability, fluctuations in the laboratory environment, operator bias, routine instrument maintenance, lot changes for calibrators and reagents, and different operators, are common sources of MU [5-9]. It is presumed that IQC data cover all anticipated routine changes in the measuring system for an appropriate period [6, 7, 10]. When repeatability or long-term imprecision data for a well-controlled measurement procedure are plotted as a Gaussian distribution, the magnitude of the dispersion of values around the mean value can be quantified by calculating the standard deviation (SD) [6, 7, 10]. Standard uncertainty can be expressed as SD. Because SD or u values cannot be added or subtracted, relative standard uncertainties (u_{rel}) first have to be converted to their respective variances (SD^2 and CV^2) in calculations [6, 7, 10].

Among the various MU factors mentioned above, reagent lot

changes are important factors that may cause a shift in IQC values, leading to MU overestimation [7]. Therefore, it is recommended that both IQC and human sample results demonstrate similar behaviors upon a reagent lot change [7]. If IQC values obtained before and after lot change are treated as a single dataset for u_{Rw} calculation, MU may be overestimated. Practical considerations for when a shift occurs after a reagent lot change are reported in several guidelines [7]. However, more specific recommendations are needed, e.g., a "significant change" upon a reagent lot change has to be clearly defined. In addition, the extent of differences that such considerations can bring about when MU is calculated using real-world IQC data should be demonstrated.

In this regard, we estimated MU by the TD approach using long-term IQC data generated in our laboratory to demonstrate how reagent lot changes influence uncertainty.

MATERIALS AND METHODS

Materials
IQC values were collected from March 2020 to February 2021. During this period, only one IQC lot (No. 28870) was used for data analysis. IQC data of 14 analytes (serum albumin, alkaline phosphatase [ALP], ALT, AST, HDL, LDL, total cholesterol, creatinine, glucose, total protein, triglyceride, potassium, sodium, and blood urea nitrogen [BUN]) were collected. We used a single chemistry measurement system (Cobas 8000 c702; Roche Diagnostics, Rotkreuz, Switzerland) and two concentration levels of IQC materials (Lyphochek Assayed Chemistry Control Lev-
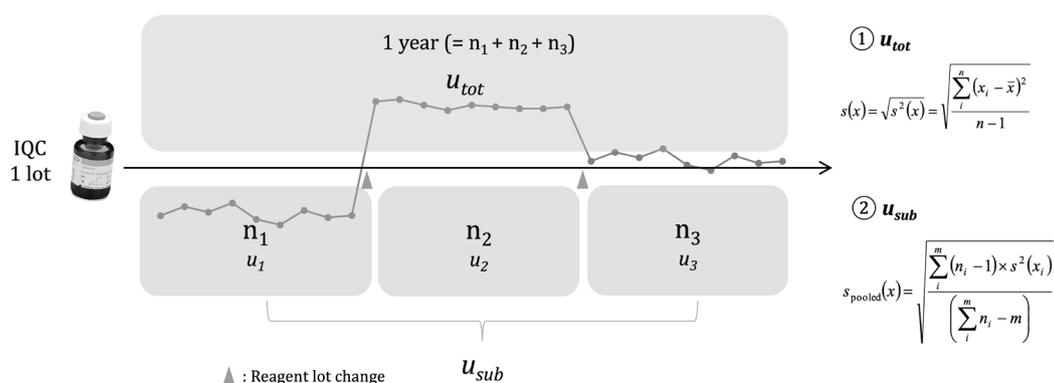
Fig. 1. Schematic illustration of MU estimation by different calculation methods. (i) Total uncertainty (u_{tot}=s(x), SD) was calculated regardless of reagent lot changes, according to equation (1). (ii) IQC values before and after a reagent lot change were collected separately by lot number, and u_{Rw} calculated in each data subgroup were combined to obtain the overall uncertainty (subgrouping uncertainty, u_{sub}=s_{pooled}(x), pooled sample SD for x organized into subgroups), according to equation (2).
Abbreviations: QC, quality control; x, measurand quantity value for a measurement; x̄, mean value of a measurand; xi, ith member of a group of values (e.g., repeated measurements of a sample); n, total number of values; ni, number of values in the ith group; m, number of groups of values.

ANNALS OF
LABORATORY
MEDICINE

Gu H, et al.
Top-down approach for clinical laboratories

els 1 and 2; Bio-Rad, Hercules, CA, USA), including multiple reagent lots during the study period (mean, 3.14; range, 2-4). Traceability and calibrator uncertainty for each analyte are summarized in Supplemental Data Table S1.

MU estimation by different calculation methods using 1-year IQC data

To demonstrate the influence of the calculation method on uncertainty, we used two methods to calculate MU (u_{Rw}) values (Fig. 1):

1) According to equation (1), the total uncertainty (u_{tot}) was calculated as the SD, regardless of reagent lot changes [6, 7].
2) IQC values before and after a reagent lot change were collected separately according to lot number, and u_{Rw} values calculated for each data subgroup were combined to obtain the overall uncertainty (subgrouping uncertainty, u_{sub}) according to equation (2) [6, 7, 10].

Basically, the calculated u_{Rw} values were combined with the uncertainty of end-user calibrator values and bias uncertainty. The combined MU values were obtained according to equation (3) [7]:

$$u_c = u(y) = \sqrt{(u_{Rw}^2 + u_{cal}^2 + u_{bias}^2)} \qquad (3)$$

where u_c is combined uncertainty, u_{Rw} is the standard uncertainty obtained by repetitive measurement, u_{cal} is the uncertainty of end-

user calibrator, and u_{bias} is the uncertainty of bias.

Bias correction is one of the most significant steps in estimating MU. We assessed the external quality assessment (EQA) results for the last two years to evaluate whether medically unacceptable bias was detected. We assumed that the end-user manufacturing process includes the correction of medically significant bias relative to the highest-order references used [4]. As there were no unacceptable results in the past two-year EQA results, we could adopt equation (4) from the ISO/TS 20419:2019 guidelines to obtain combined uncertainty [7]:

$$u_c = u(y) = \sqrt{(u_{Rw}^2 + u_{cal}^2)} \text{ (if bias is within specification)} \qquad (4)$$

The calculated u_{Rw} values were combined with the uncertainty of end-user calibrator values to obtain expanded uncertainty [6, 7]:

$$U = u(y) \times k \text{ } (k, \text{ } coverage \text{ } factor; k = 2, 95\% \text{ } confidence \text{ } interval) \qquad (5)$$

All combined uncertainty (u_c) values were multiplied by 2 (coverage factor, $k = 2$) (equation 5). The expanded uncertainty values were expressed as the standard expanded uncertainty (U) with their units and relative expanded uncertainty (%U_{rel}) [6, 7, 10].

When a new reagent lot is introduced, any change in IQC results should be matched by a change in values obtained for a panel of typical human samples of equal magnitude [7]. For new reagent lot validation, we measured three QC samples at
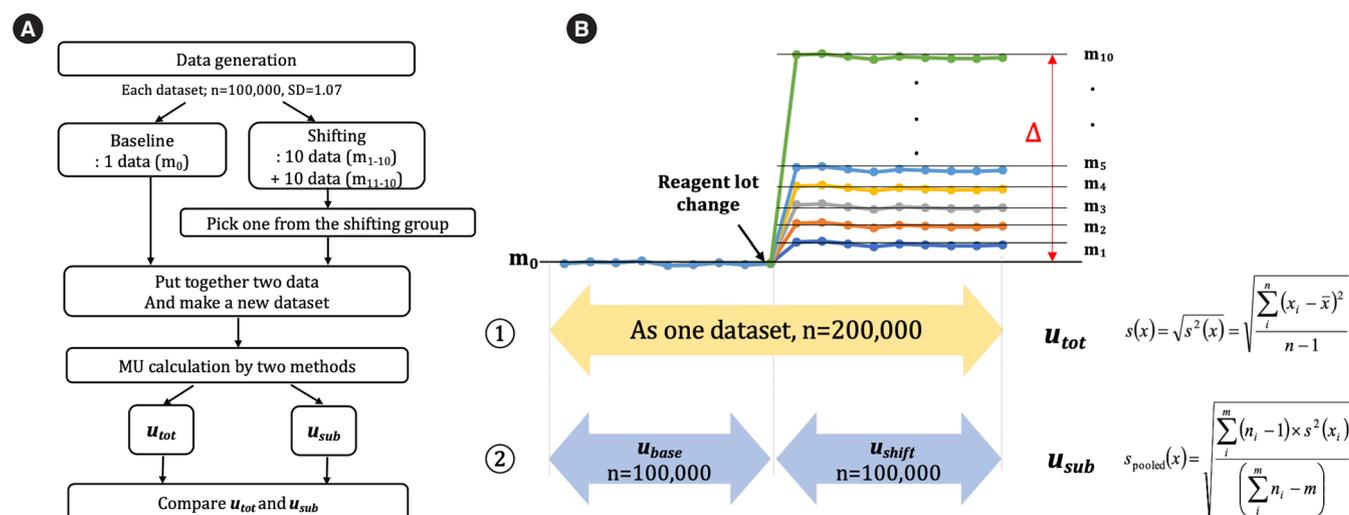
Fig. 2. Schematic illustration of reagent lot change simulation using artificial IQC data. (A) Process of generating baseline and shifted artificial IQC datasets. (B) MU estimation with the generated datasets by different calculation methods.
Abbreviations: IQC, internal quality control; MU, measurement uncertainty; m_{0-10}, arithmetic mean of baseline and shifted datasets; u_{tot}, standard uncertainty calculated regardless of a reagent lot change; u_{sub}, combined uncertainty calculated using values obtained from each subgroup; u_{base}, standard uncertainty of the baseline data; u_{shift}, standard uncertainty of the shifted data; x, measurand quantity value; n, total number of values; ni, number of values in the ith group; m, number of groups of values.

Gu H, et al.
Top-down approach for clinical laboratories

ANNALS OF
LABORATORY
MEDICINE

two levels and five human samples with various concentrations encompassing the reference interval concentration of each analyte. The human samples were verified using a current (old) and a new reagent to check the consistency of the results (Supplemental Data Table S2). The study was exempted from approval from the Institutional Review Board (IRB No. S2022-0370-0001) given the study's design.

Monte Carlo simulation (MCS) of a reagent lot change using artificial IQC data

A shift in artificial IQC data due to a reagent lot change was simulated to show differences in MU depending on the degree of shift and the calculation method (Fig. 2A). MCS was used to demonstrate how a shift in IQC data may significantly affect MU evaluation. MCS uses algorithmically produced pseudo-random numbers that follow a specified probability distribution. The dispersion of random integers in a normal distribution is predefined by the stated mean and SD. The MCS method creates a ran-

dom numeric number from the probability distribution function for each input. It generates numeric values for all inputs to the known functional relationship, which are utilized to generate a single numeric value as an output. This is repeated a sufficient number of times ("trials") to provide a collection of simulated results as an output. The measurand and its standard uncertainty are then estimated using the mean and SC of the output results.

Following this procedure, multiple datasets were generated using a random number generator with the same SD, but different means. Each dataset included 100,000 random numbers. The simulation demonstrated various degrees of shift in IQC data upon a reagent lot change. First, a baseline IQC dataset (baseline; m0 = 100; SD = 1.7; N = 100,000) was generated. Then, 20 shifted IQC datasets with the same SD as the baseline data, but different mean shifts, were generated (m1 to m10: 1%, 2%, 4%, 6%, 8%, 10%, 12.5%, 15, 17.5%, 20% higher than m0 and m11 to m20: 1%, 2%, 4%, 6%, 8%, 10%, 12.5%, 15%,

Table 1. Results of MU estimation using different calculation methods

Analytes	QC level 1			QC level 2			No. of reagent lots used
	U_{rel_tot} (%)	U_{rel_sub} (%)	%Difference*	U_{rel_tot} (%)	U_{rel_sub} (%)	%Difference	
Albumin	2.944	2.881	0.063	2.565	2.479	0.086	3
ALP	4.259	4.013	0.245	4.085	3.59	0.495	3
ALT	8.01	7.995	0.015	2.591	2.573	0.018	2
AST	5.425	5.382	0.043	2.939	2.908	0.031	3
HDL	4.655	4.165	0.49	5.348	4.523	0.825	3
LDL	4.414	4.312	0.101	3.926	3.837	0.089	2
Total cholesterol	3.554	3.491	0.064	3.567	3.485	0.083	4
Creatinine	5.847	5.813	0.034	4.345	4.334	0.011	3
Glucose	2.159	2.085	0.073	2.144	2.068	0.076	4
Total protein	2.359	2.311	0.048	2.29	2.263	0.028	4
Triglyceride	2.581	2.579	0.002	2.534	2.532	0.002	3
K	1.538	1.538	7.13E-05	1.284	1.282	0.002	2
Na	1.557	1.525	0.032	1.368	1.347	0.021	4
BUN	4.366	4.306	0.06	3.477	3.458	0.02	4
Mean	4.752	4.606	0.146	3.575	3.342	0.232	2.9
Median	4.414	4.165	0.064	3.567	3.485	0.086	3
Min (analyte)	2.944	2.881	1.45E-02	2.534	2.532	0.018	2
		K			Triglyceride		
Max (analyte)	4.655	4.165	0.49	5.348	4.523	0.825	4
		HDL			HDL		

*%Difference = % U_{rel_tot} − %Sub U_{rel_sub}
Abbreviations: MU, measurement uncertainty; QC, quality control; U_{rel_tot}, expanded relative uncertainty (coverage factor, $k=2$) calculated regardless of reagent lot change, akin to CV; U_{rel_sub}, expanded relative uncertainty (coverage factor, $k=2$) with obtained values from each subgroup, akin to CV; ALP, alkaline phosphatase; K, potassium; Na, sodium; BUN, blood urea nitrogen.

ANNALS OF
LABORATORY
MEDICINE

Gu H, et al.
Top-down approach for clinical laboratories

17.5%, and 20% lower than m0 (SD = 1.7, N = 100,000 each). The baseline data and one subset from the shifted dataset were combined (baseline-shifted) to generate a new dataset. Then, we calculated and compared the uncertainties using two methods [6, 7]. In the first method, uncertainty was calculated regardless of a mean change, and in the second method, uncertainties were calculated separately according to the mean and then combined using equation (6) [6, 7, 10] (Fig. 2B).

Comparison of two IQC datasets from different QC material lots

Two sets of IQC data from different QC material lots, which were used consecutively for a year from March 2019 to February 2021, were comparatively analyzed.

Outlier elimination

Values outside 1.5 times the interquartile range were eliminated as outliers using Tukey's fences [11].

Statistical analysis

Basic calculations for MU and data analyses, including Tukey's fences, MCS, t-tests, and F-tests, were performed using Microsoft Excel 365 (Microsoft, Redmond, WA, USA). Data normality and distribution skewness and kurtosis were analyzed using RStudio (PBC, Boston, MA, USA). An absolute skewness value

≤2 or an absolute kurtosis (excess) ≤4 was used as a threshold for determining considerable normality [12, 13].

RESULTS

Comparison of MU data obtained via different calculation methods using 1-year IQC data

During the one-year study period, the mean number of IQC evaluations performed for each item was 1,087.9 (range, 1,049-1,296). All $\%U_{rel}$ values calculated by subgrouping ($\%U_{rel\_sub}$) were lower than the values calculated as a whole ($\%U_{rel\_tot}$); the mean values of $\%U_{rel\_sub}$ and $\%U_{rel\_tot}$ were 5.45 and 5.62 at level 1 and 4.95 and 5.07 at level 2, respectively (Table 1). The entire calculation process and MU values for each analyte are listed in Supplemental Data Table S3 (A-N). The mean MU differences (%, $\%U_{rel\_sub}-\%U_{rel\_tot}$) were -0.13% at level 1 (range, -0.88 to -2.43×10^{-5}%) and -0.12% at level 2 (range, -0.77 to -2.20×10^{-5}%; Fig. 3).

MCS of a reagent lot change using artificial IQC data

IQC data with varying degrees of shift were simulated. As shown in Fig. 4A, all $\%U_{sub}$ values were relatively constant, irrespective of the degree of shift. However, $\%U_{sub}$ values increased with increasing degree of shift in both directions. The mean MU differences (%, $\%U_{rel\_sub}-\%U_{rel\_tot}$) gradually increased as the mean
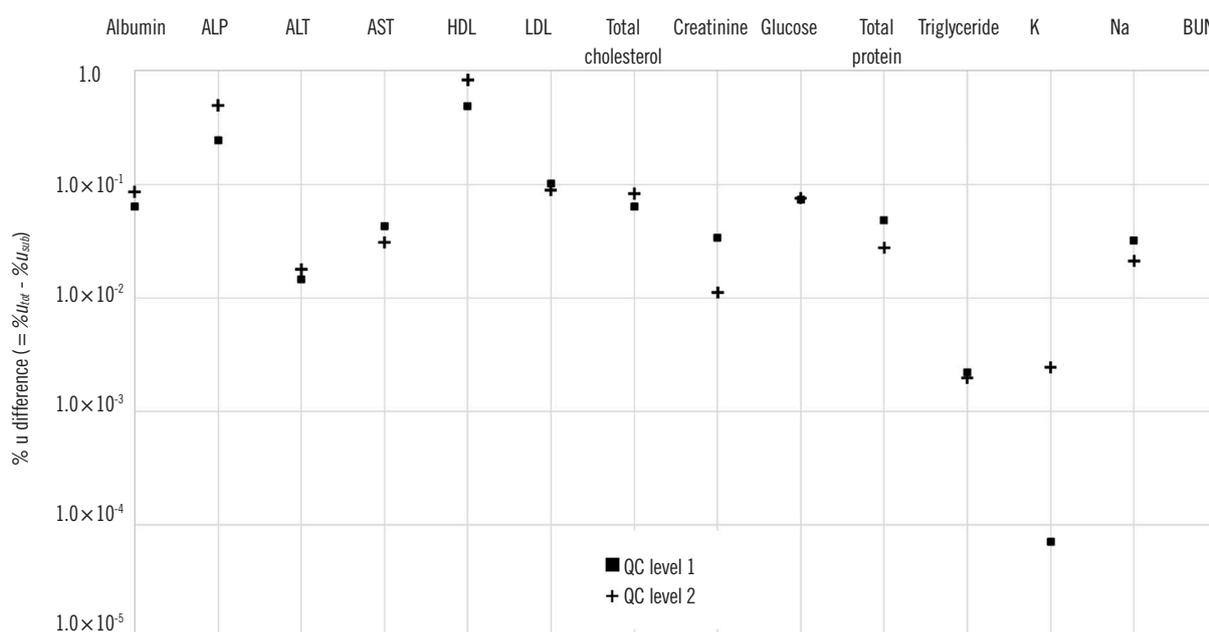
Fig. 3. Results of MU estimation using different calculation methods.
Abbreviations: MU, measurement uncertainty; u_{tot}, standard uncertainty calculated regardless of a reagent lot change; u_{sub}, combined uncertainty calculated using values obtained from each subgroup; ALP, alkaline phosphatase; K, potassium; Na, sodium; BUN, blood urea nitrogen.

differences (shifts) increased (Fig. 4B).

Review and comparative analysis of IQC data using two different QC material lots

IQC data obtained from two different IQC material lots that were used for two consecutive years (lot 1: March 2019 to February 2020; lot 2: March 2020 to February 2021) were reviewed, and the mean and SD values for each lot of IQC data were obtained (Supplemental Data Table S4). The t-test (for mean comparison) and F-test (for variance [SD] comparison) were used to analyze the significance of differences between the IQC datasets (Table 2). The mean values were significantly different at all levels for all analytes except creatinine. Additionally, SDs were significantly different for most analytes (except AST, LDL, creati-
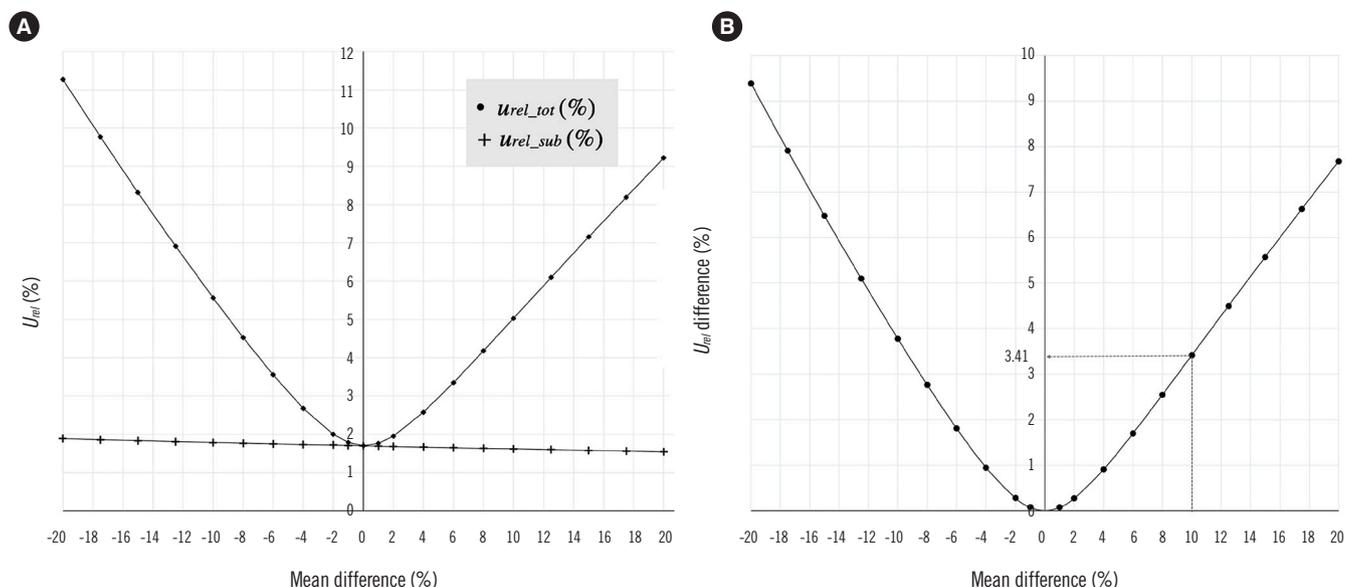
Fig. 4. Difference in MU values between two differently calculated groups. (A) As the mean differences (%) increased, u_{rel_tot} values calculated regardless of mean change showed a constant MU, whereas u_{sel_sub} values calculated separately by mean and then combined showed gradual increases in MU. (B) MU differences ($= \%u_{rel\_tot} - \%u_{rel\_sub}$) with respect to mean differences (%) are plotted. For example, a mean difference of 10% after one reagent lot change corresponded to a difference in MU of 3.41%.
Abbreviations: MU, measurement uncertainty; u_{rel_tot}, relative uncertainty calculated regardless of a reagent lot change; u_{rel_sub}, relative combined uncertainty calculated from values obtained from each subgroup; $\%u_{rel}$, relative uncertainty, akin to CV.

Table 2. Comparison of two IQC data groups from separate QC material lots, using t- and F-tests

QC level	Statistical analysis	Alb	ALP	ALT	AST	HDL	LDL	T-chol	Cr	Glu	Ptn	K	Na	TG	BUN
1	t-test	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
	P (<0.05)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.086	0.000	0.000	0.000	0.000	0.000	0.000
	F-test	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N
	P (<0.05)	0.015	0.000	0.000	0.114	0.009	0.693	0.000	0.709	0.000	0.795	0.000	0.000	0.014	0.705
2	t-test	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	P (<0.05)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	F-test	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
	P (< 0.05)	0.228	0.000	0.000	0.769	0.000	0.006	0.000	0.092	0.000	0.001	0.003	0.000	0.033	0.014

*t-test: Y (P<0.05), a significant difference exists between the mean values of two groups; N (P≥0.05), no significant difference exists between the mean values of two groups.
*F-test: Y (P<0.05), a significant difference exists between the variance (SD) values of two groups.; N (P≥0.05), no significant difference exists between the variance (SD) of two groups.
Abbreviations: MU, measurement uncertainty; IQC, internal quality control; Alb, albumin; ALP, alkaline phosphatase; T-chol, total cholesterol; Cr, creatinine; Glu, glucose; Ptn, protein; K, potassium; Na, sodium; TG, triglyceride; BUN, blood urea nitrogen.

ANNALS OF
LABORATORY
MEDICINE

Gu H, et al.
Top-down approach for clinical laboratories

nine, total protein, and BUN at level 1 and albumin, AST, and creatinine at level 2).

DISCUSSION

Guidelines issued by the International Organization for Standardization (ISO) and the CLSI recommend the TD approach for MU estimation in clinical laboratories [6-8]. In this approach, IQC data, which are easily obtainable in clinical laboratories, are required as a key component to evaluate MU, particularly, in laboratories that use closed measurement systems [14]. However, the practical issues faced in working conditions need to be addressed. For example, the recommended collection period, described as a "sufficiently long time," should be defined in detail [6, 7]. Furthermore, clear recommendations should be made for IQC data obtained from different reagent lots. The guidelines recommend collecting IQC data separately if "a significant shift" in the IQC absolute values occurs when a new lot of reagents is introduced [6, 7]. However, the range of acceptable values is not defined. If each laboratory uses different standards to calculate MU, the accuracy of the results may be compromised and/or confusion may arise.

Initially, it was assumed that MU values ($\%U_{sub}$) calculated by subgrouping of the data would be substantially lower than those calculated as a whole ($\%U_{tot}$). However, the differences between MU values obtained by the two different calculation methods were minimal (minimum difference: $7.13 \times 10^{-5}\%$, maximum difference: 0.825%), although the $\%U_{sub}$ values were lower for all analytes.

It is common to observe matrix effects in IQC materials that produce different results than human serum samples during the reaction with reagents [15]. We attempted to identify how large mean differences between IQC and patient sample results are before and after a reagent lot change (Supplemental Data Table S2). The mean differences in the two groups were within a narrow interval (in IQC data, up to 3.73% in absolute value; in patient sample data, up to 2.5% in absolute value). We therefore presumed that a mean change of <4% in IQC data may not cause a marked difference, depending on the consideration of a reagent lot change during MU estimation. This may be due to the good management of IQC activities in the laboratory.

To demonstrate the effect of a significant shift in IQC data after a reagent lot change, we conducted a simulation with artificial IQC datasets consisting of random numbers and considering one reagent lot change. As the degree of IQC data shift gradually increased, the difference between MU results in-

creased according to the calculation method. For example, in a dataset generated with a 10% shift from the mean, the difference in MU values was ~3.41% (U_{rel}, relative expansion uncertainty, $k=2$) (Fig. 4B). When we comparatively evaluated reagent lot changes in the laboratory, the predefined allowable total error was used as an acceptable performance criterion [16]. If we presume that the mean difference after a lot change was 8%, which is within the acceptable interval, the new lot would be used without further evaluation. However, in MU estimation, a significant difference was observed depending on the calculation method used. The MU value calculated regardless of the shift of 8% was higher ($u_{tot}=4.34$) than that calculated considering the shift ($u_{sub}=1.7$), which led to a highly overestimated MU value.

As observed in the third analysis, a shift in IQC data may indicate that all lots of IQC materials should be treated as different materials. In addition, an SD change in the IQC data may show various uncertainty factors related to the measurement system at the time and/or the IQC material lot change. Therefore, an IQC material lot change may be accompanied by changes not only in the IQC material substances but also in the measurement system over time.

If MU estimation was performed using the combined results of multiple IQC material lot changes, the MU values would be overestimated due to the effects of shifts and other influences from the measurement system over time (6, 7). Therefore, to obtain stable MU values using the TD approach, we suggest that one QC lot should be used for at least six consecutive months.

Lot changes of the calibrator used in MU estimation were not considered in accordance with the ISO guideline, which stipulates that separate collection and calculation of IQC data are not required unless the calibrator manufacturer introduces significant changes, such as a setpoint change [6, 7]. Furthermore, the calculation using the IQC data as a single set based on the calibrator lot change will capture the variability of human sample results due to this change as a random error [6, 7]. The means changed in different patterns over several reagent lot changes, which may indicate that the effects of the mean change were weakened. This weakening effect was not considered, and further studies may be needed.

MU estimation in clinical laboratories universally requires more detailed discussions and revisions by expert groups. The results of this study may provide basic, but practical, considerations in clinical laboratories for conducting MU estimation using a TD approach. In conclusion, reagent lot changes should

be considered when the TD approach is applied to IQC data, and data from a single lot of IQC materials are recommended to obtain stable and reliable MU values using the TD approach.

AUTHOR CONTRIBUTIONS

Gu H designed the study, analyzed the data, and wrote the draft; Lee W conceived the study, analyzed the data, and finalized the draft; Lee WI conceived the study and reviewed the manuscript; Yun YM discussed the data and reviewed the manuscript; Chun S and Min WK provided perspectives on the study concept; Lee J supported the study data collection; and Hong J discussed the data. All authors have read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

ORCID

Hyunjung Gu	https://orcid.org/0000-0001-8673-7326
Juhee Lee	https://orcid.org/0000-0002-1832-7787
Jinyoung Hong	https://orcid.org/0000-0001-8855-7900
Woochang Lee	https://orcid.org/0000-0003-3956-6397
Yeo-Min Yun	https://orcid.org/0000-0002-5485-8331
Sail Chun	https://orcid.org/0000-0002-5792-973X
Woo-In Lee	https://orcid.org/0000-0002-9409-2959
Won-Ki Min	https://orcid.org/0000-0002-5158-2130

REFERENCES

1. Ćelap I, Vukasović I, Juričić G, Šimundić AM. Minimum requirements for the estimation of measurement uncertainty: recommendations of the joint Working group for uncertainty of measurement of the CSMBLM and CCMB. Biochem Med (Zagreb) 2017;27:030502.
2. Infusino I and Panteghini M. Measurement uncertainty: friend or foe? Clin Biochem 2018;57:3-6.
3. Braga F, Infusino I, Panteghini M. Role and responsibilities of laboratory medicine specialists in the verification of metrological traceability of *in vitro* medical diagnostics. J Med Biochem 2015;34:282-7.
4. Braga F and Panteghini M. The utility of measurement uncertainty in medical laboratories. Clin Chem Lab Med 2020;58:1407-13.
5. BIPM. Evaluation of measurement data — Guide to the expression of uncertainty in measurement, GUM 1995 with minor corrections. JCGM, 2008;100.
6. CLSI. Expression of measurement uncertainty in laboratory medicine, approved guideline. CLSI EP29-A. Wayne, PA: Clinical and Laboratory Standards Institute, 2012.
7. ISO. Medical laboratories — Practical guidance for the estimation of measurement uncertainty. ISO/TS 20914:2019 Technical Specification: The International Organization for Standardization, 2019.
8. EURACHEM/CITAC. Measurement uncertainty arising from sampling: a guide to methods and approaches, 2019.
9. Braga F, Infusino I, Panteghini M. Performance criteria for combined uncertainty budget in the implementation of metrological traceability. Clin Chem Lab Med 2015;53:905-12.
10. Ellis AD, Gross AR, Budd JR, Miller WG. Influence of reagent lots and multiple measuring systems on estimating the coefficient of variation from quality control data; implications for uncertainty estimation and interpretation of QC results. Clin Chem Lab Med 2020;58:1829-35.
11. Johansen MB and Christensen PA. A simple transformation independent method for outlier definition. Clin Chem Lab Med 2018;56:1524-32.
12. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth 2019;22:67-72.
13. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013;38:52-4.
14. Ceriotti F. Deriving proper measurement uncertainty from Internal Quality Control Data: an impossible mission? Clin Biochem 2018;57:37-40.
15. Kim SY, Chun S, Lee W, Min WK. Commutability of proficiency testing (PT): status of the matrix-related bias in general clinical chemistry. Clin Chem Lab Med 2013;51:e169-73.
16. Thompson S and Chesher D. Lot-to-lot variation. Clin Biochem Rev 2018;39:51-60.

ANNALS OF
LABORATORY
MEDICINE

Gu H, et al.
Top-down approach for clinical laboratories

Supplemental Data Table S1. Information of traceability and calibrator uncertainty

Analyte	Reagent kit	Calibrator	Reference method	Reference material	Calibrator value	Uncertainty (expanded uncertainty; $k=2$)	Unit	Routine method
Albumin	ALBP	C.f.a.s.		ERM-DA470k/IFCC	2.82	0.0264	g/dL	BCP
ALP	ALP2	C.f.a.s.	Original formulation IFCC5 (Tietz, 1983), manual measurement		211	1.25	U/L	IFCC Gen. 2
ALT	ALT	C.f.a.s.	Original formulation IFCC5 (2002) modified, manual measurement		96.2	0.638	U/L	IFCC without pyridoxal phosphate activation
AST	AST	C.f.a.s.	Original formulation IFCC5 (2002) modified, manual measurement		105	0.859	U/L	IFCC without pyridoxal phosphate activation
HDL	HDLC4	C.f.a.s.	According to the CDC reference method ([pPrecipitation by dextran sulphate and Abell–-Kendall])		65.3	1.17	mg/dL	Eenzymatic colorimetric Gen. 4
LDL	LDLC3	C.f.a.s.	According to CDC beta quantification method ([pPrecipitation by heparin-manganese (ultracentrifugation) and Abell–-Kendall])		137	1.55	mg/dL	Gen. 3
Total cholesterol	CHOL2	C.f.a.s.	ID-MS		158	1.35	mg/dL	CHOD-PAP Gen. 2 stand. ID/MS
Creatinine	CREJ2	C.f.a.s.	ID-MS		4.06	0.0576	mg/dL	Jaffé rate-blanked and compensated Gen. 2 serum, plasma
Glucose	GLUC3	C.f.a.s.	ID-MS		193	1.60	mg/dL	HK
Protein	TP2	C.f.a.s.		SRM 927d	5.05	0.0269	g/dL	Biuret Gen. 2
TG	TRIGL	C.f.a.s.	ID-MS		59.5	0.700	mg/dL	GPO-PAPGlycerine phosphate oxidase peroxidase
K	ISE	ISE Standard Low/High	ISE Standard: pPrimary calibrators prepared gravimetrically from p.a. quality salts		3.00/7.00	0.0122/0.0285	mmol/L	ISE
Na	ISE	ISE Standard Low/High	ISE Standard: pPrimary calibrators prepared gravimetrically from p.a. quality salts		120/160	0.439/0.640	mmol/L	ISE
BUN	UREAL	C.f.a.s.	ID -MS		47.3	0.880	mg/dL	Urease/GLDH glutamate dehydrogenase serum, plasma

*All reagent kits and calibrators were manufactured by Roche Diagnostics (Roche, Rotkreuz, Switzerland).

Abbreviations: C.f.a.s, calibrator for automated system; IFCC, the method authorized by the International Federation of Clinical Chemistry and Laboratory Medicine; ID-MS, isotope dilution-mass spectrometry; ERM, European Reference Materials; CDC, Centers for Disease Control; ISE, ion-selective electrode; p.a., pro analysis (for analysis); ALP, alkaline phosphatase; K, potassium; Na, sodium; TG, triglyceride; BUN, blood urea nitrogen.

Supplemental Data Table S2. Reagent lot-to-lot validation results with QC materials and patient samples for each test item

Analyte	No. of reagent lots used		%Mean difference ((new value − old value)/(old value) × 100)			%U_{rel} ($k=2$)		
	No. of lot changes		1	2	3	U_{rel_sub}	U_{rel_tot}	U_{rel_tot}-U_{rel_sub}
Albumin	3	QC level 1	−0.75	0.00		2.88	2.94	0.06
		QC level 2	0.49	−1.96		2.47	2.57	0.09
		Patient	0.40	−1.20				
ALP	3	QC level 1	1.53	−2.73		4.01	4.26	0.25
		QC level 2	0.76	−2.27		3.59	4.09	0.50
		Patient	−0.30	−2.50				
ALT	2	QC level 1	3.32			8.00	8.01	0.01
		QC level 2	1.28			2.57	2.59	0.02
		Patient	1.10					
AST	3	QC level 1	2.70	−3.51		5.38	5.43	0.04
		QC level 2	1.70	3.32		2.91	2.94	0.03
		Patient	0.60	0.40				
HDL	3	QC level 1	0.46	0.95		4.17	4.66	0.49
		QC level 2	1.82	0.34		4.52	5.35	0.83
		Patient	0.70	0.40				
LDL	2	QC level 1	−2.12			4.31	4.41	0.10
		QC level 2	−2.26			3.84	3.93	0.09
		Patient	−1.30					
Total cholesterol	3	QC level 1	−1.16	−0.95		3.49	3.55	0.06
		QC level 2	0.61	−0.45		3.49	3.57	0.08
		Patient	0.90	0.20				
Creatinine	3	QC level 1	−0.08	0.55		5.81	5.85	0.03
		QC level 2	−3.67	3.29		4.33	4.35	0.01
		Patient	−1.00	0.20				
Glucose	4	QC level 1	0.73	0.85	−1.20	2.09	2.16	0.07
		QC level 2	0.18	0.36	−0.04	2.07	2.14	0.08
		Patient	0.20	0.01	0.10			
Total protein	4	QC level 1	−1.16	−0.95	0.12	2.31	2.36	0.05
		QC level 2	0.61	−0.45	1.20	2.26	2.29	0.03
		Patient	0.90	0.20	0.40			
Triglyceride	3	QC level 1	0.90	−0.69		2.58	2.58	0.002
		QC level 2	0.16	1.40		2.53	2.53	0.002
		Patient	0.40	0.10				
BUN	4	QC level 1	−3.73	−0.65	−2.61	4.31	4.37	0.06
		QC level 2	−0.42	−2.49	0.64	3.46	3.48	0.02
		Patient	−0.70	0.40	0.30			
Minimum (absolute difference)		QC level 1	−0.08		(cCreatinine)	2.07 (gGlucose)	2.14 (gGlucose)	0.002 (tTriglyceride)
		QC level 2	−0.04		(gGlucose)			
		Patient	0.01		(gGlucose)			
Maximum (absolute difference)		QC level 1	−3.73		(BUN)	7.995 (ALT)	8.01 (ALT)	0.825 (HDL)
		QC level 2	−3.67		(cCreatinine)			
		Patient	−2.50		(ALP)			
Mean		QC level 1	−0.41					0.125
		QC level 2	0.17					
		Patient	0.04					

*The data for potassium and sodium are not included in this table because of the lack of comparison data with patient samples.
Abbreviations: IQC, internal quality control; QC, quality control; U_{rel}, relative expanded uncertainty; U_{rel_tot}, expanded relative uncertainty (coverage factor, $k=2$) calculated regardless of reagent lot change, akin to CV; U_{rel_sub}, expanded relative uncertainty (coverage factor, $k=2$) calculated from values obtained from each subgroup, akin to CV; ALP, alkaline phosphatase; BUN, blood urea nitrogen.

ANNALS OF LABORATORY MEDICINE

Gu H, et al.
Top-down approach for clinical laboratories

Supplemental Data Table S3. Worked-out examples of MU estimation of two chemistry analytes (albumin and creatinine)

(A) Albumin

Component (aAnalyte)	Albumin (Alb)			
Measurement unit	g/dL			
Reference range	3.5–5.2			
Measurement method	Colorimetric assay-BCP			
Measurement system	Roche Cobas 8000			
Calibrator uncertainty				
Assigned value	2.82			
U_{cal} ($k=2$)	0.0264			
U_{rel} (cal) ($k=2$)	0.0094			
Long-term precision	Level 1		Level 2	
Total period	2020.03.09–2021.02.28			
N, enrolled data	1,060		1,060	
N, outliers	16		20	
N, total data	1,044		1,040	
N, used reagent lot	3		3	
Data cellectioncollection per reagent lot	Subgrouping	Total	Subgrouping	Total
Period 1	2020.03.09–2020.07.08	SD (u_{rw})	2020.03.09–2020.07.08	SD (u_{rw})
N	325	$=0.0357$	325	$=0.0485$
Mean	2.57		4.08	
$u\,1$	0.0385		0.0485	
Period 2	2020.06.11–2020.11.30		2020.06.11–2020.11.30	
N	459		468	
Mean	2.55		4.05	
$u\,2$	0.0315		0.0440	
Period 3	2020.12.01–2021.02.28		2020.12.01–2021.02.28	
N	260		247	
Mean	2.56		4.06	
$u\,3$	0.0358		0.0489	
$u\,Rw$, U/L (pooled average)	0.0348		0.0466	
Total mean	2.56		4.06	
Uncertainty estimation				
Standard (Rw)				
u_{Rw}	0.0348	0.0357	0.0466	0.0485
U_{Rw} ($k=2$)	0.0697	0.0714	0.0932	0.0969
U_c at IQC mean $=\sqrt{(U2cal+U2Rw)}$ ($k=2$)	0.075	0.076	0.097	0.100
Relative (rel)				
$u_{rel(Rw)}$	0.0136	0.0140	0.0115	0.0119
$U_{rel(Rw)}$ ($k=2$)	0.0272	0.0279	0.0230	0.0239
$U_{rel,\,c}=\sqrt{(U2rel(cal)+U2rel(Rw))}$ ($k=2$)	0.0288	0.0294	0.0248	0.0256
$U_{rel,\,c}$ ($k=2$)	2.88	2.94	2.48	2.56
Applied to patients' resultsdata (95% confidencet interval)	2.56 ± 0.0745 g/dL or (2.88%)	2.56 ± 0.0761 g/dL or (2.94%)	4.06 ± 0.0969 g/dL or (2.48%)	4.06 ± 0.100 g/dL or (2.56%)

Abbreviations: MU, measurement uncertainty; BCP, bromocresol purple; U_{cal}, expanded uncertainty of end-user calibrator; $U_{rel(cal)}$, relative expanded uncertainty of end-user calibrator; ui, standard uncertainty in the ith group; u_{Rw}, standard uncertainty obtained by repetitive measurement; U_{Rw}, expanded uncertainty obtained by repetitive measurement; U_c, expanded combined uncertainty; $u_{rel(Rw)}$, relative uncertainty obtained by repetitive measurement; $U_{rel(Rw)}$, relative expanded uncertainty obtained by repetitive measurement; $U_{rel,\,c}$, relative expanded combined uncertainty.

(B) Creatinine

Component (aAnalyte)	Creatinine (Cr)			
Measurement unit	mg/dL			
Reference range	0.7–1.4			
Measurement method	Jaffé method, kinetic colorimetric assay			
Measurement system	Roche Cobas 8000			
Calibrator uncertainty				
Assigned value	4.06			
U_{cal} (k=2)	0.0576			
$U_{rel(cal)}$ (k=2)	0.0142			
Long-term precision	Level 1		Level 2	
Total period	2020.03.09–2021.02.28			
N, enrolled data	1,065		1,060	
N, outliers	41		29	
N, total data	1,024		1,031	
N, used reagent lot	3		3	
Data cellectioncollection per reagent lot	Subgrouping	Total	Subgrouping	Total
Period 1	2020.03.09–2020.05.17	SD (u_{rw})	2020.03.09–2020.05.17	SD (u_{rw})
N	166	= 0.0355	164	=0.107
Mean	1.25		5.23	
u 1	0.0334		0.103	
Period 2	2020.05.18 – -2020.11.18		2020.05.18–2020.11.18	
N	569		575	
Mean	1.25		5.21	
u 2	0.0360		0.107	
Period 3	2020.11.19–2021.02.28		2020.11.19–2021.02.28	
N	289		292	
Mean	1.26		5.23	
u 3	0.0349		0.108	
u_{Rw}, U/L (pooled average)	0.0353		0.107	
Total mean	1.25		5.22	
Uncertainty estimation				
Standard (Rw)				
u_{Rw}	0.0353	0.0355	0.107	0.107
U_{Rw} (k=2)	0.0706	0.0710	0.214	0.214
U_c at IQC mean$=\sqrt{(U2cal+U2Rw)}$ (k=2)	0.0911	0.0915	0.221	0.222
Relative (rel)				
$u_{rel(Rw)}$	0.0282	0.0284	0.020	0.021
$U_{rel(Rw)}$ (k=2)	0.0564	0.0567	0.041	0.041
$U_{rel,c}=\sqrt{(U2rel(cal)+U2rel(Rw))}$ (k=2)	0.0581	0.0585	0.043	0.043
$U_{rel,c}$ (k=2)	5.81	5.85	4.33	4.34
Applied to patients' resultsdata (95% confidencet interval)	2.56±0.0745 g/dL or (2.88%)	2.56±0.0761 g/dL or (2.94%)	4.06±0.0969 g/dL or (2.48%)	4.06±0.100 g/dL or (2.56%)

Supplemental Data Table S4. Mean values and MU (SD and CV (%)) of IQC data of 14 analytes with two different QC lots used for two consecutive years (lot 1: March 2019 to February 2020, lot 2: March 2020 to February 2021)

Year Level	MU	2019		2020		%Diff. (2020 − 2019)	
		1	2	1	2	1	2
Albumin	mean	2.66	4.14	2.56	4.06	−3.92	−1.97
	SD	0.0373	0.0544	0.0360	0.0480	−3.68	−13.33
	CV(%)	1.40	1.31	1.41	1.18	0.236	−11.1
ALP	mean	92.1	382	94.7	365	2.76	−4.69
	SD	1.84	7.01	1.98	7.37	7.13	4.86
	CV(%)	1.99	1.84	2.09	2.02	4.49	9.12
ALT	mean	29.1	97.5	26.9	93.9	−8.23	−3.91
	SD	1.21	1.64	1.07	1.18	−12.5	−39.6
	CV(%)	4.15	1.68	3.99	1.25	−3.92	−34.4
AST	mean	39.4	196	40.2	200	2.06	2.34
	SD	1.13	2.76	1.08	2.83	−5.16	2.38
	CV(%)	2.88	1.41	2.68	1.41	−7.38	0.03
HDL	mean	22.3	59.9	21.1	58.9	−5.69	−1.67
	SD	0.664	2.12	0.45	1.48	−46.5	−43.0
	CV(%)	2.98	3.54	2.15	2.52	−38.6	−40.7
LDL	mean	61.3	135.3	61.6	136	0.563	0.703
	SD	1.58	3.41	1.31	2.56	−20.1	−33.3
	CV(%)	2.58	2.52	2.13	1.88	−20.8	−34.3
T-chol	mean	96.9	237	101	253	3.75	6.42
	SD	1.51	3.44	4.40	4.38	65.7	21.4
	CV(%)	1.56	1.45	4.37	1.73	64.3	16.0
Creatinine	mean	1.26	5.06	1.25	5.22	−0.78	3.04
	SD	0.04	0.10	0.04	0.11	−3.37	2.25
	CV(%)	2.91	2.07	2.83	2.05	−2.57	−0.82
Glucose	mean	83.7	284	82.5	275	−1.47	−3.40
	SD	1.18	3.40	1.04	2.72	−12.8	−25.1
	CV(%)	1.40	1.20	1.26	0.99	−11.1	−21.0
T-Protein	mean	4.31	6.64	4.19	6.66	−2.80	0.29
	SD	0.0550	0.0778	0.0472	0.0735	−16.5	−5.82
	CV(%)	1.28	1.17	1.13	1.10	−13.3	−6.13
K	mean	4.01	6.16	3.97	6.17	−1.07	0.16
	SD	0.0281	0.0355	0.0294	0.0376	4.39	5.60
	CV(%)	0.702	0.576	0.742	0.609	5.40	5.45
Na	mean	123	144	126	146	1.92	1.14
	SD	0.834	0.931	0.951	0.944	12.4	1.41
	CV(%)	0.676	0.647	0.756	0.649	10.7	0.27
TG	mean	98.9	181	103	187	3.80	3.50
	SD	1.26	2.03	1.18	2.10	−6.44	3.37
	CV(%)	1.27	1.12	1.15	1.12	−10.6	−0.136
BUN	mean	15.7	46.0	15.1	47.5	−4.21	3.15
	SD	0.326	0.867	0.297	0.694	−10.0	−24.9
	CV(%)	2.08	1.89	1.97	1.46	−5.60	−28.9
Total Diff.	mean					−0.951 (−8.23 to 3.80)	0.364 (−4.69 to 6.42)
	SD					−3.39 (−46.5 to 65.7)	−10.3 (−43.0 to 21.4)
	CV(%)					−2.06 (−38.6 to 64.3)	−10.5 (−40.7 to 16.0)

Abbreviations: MU, measurement uncertainty; IQC, internal quality control; QC, quality control material; ALP, alkaline phosphatase; T-chol, total cholesterol; K, potassium; Na, sodium; TG, triglyceride; BUN, blood urea nitrogen.