Harmonization of laboratory results by data adjustment in multicenter clinical trials

Sang Gon Lee[1], Hee-Jung Chung[2], Jeong Bae Park[3], Hyosoon Park[4], and Eun Hee Lee[1]

[1]Department of Laboratory Medicine, Green Cross Laboratories, Yongin; [2]Department of Laboratory Medicine, National Cancer Center, Goyang; [3]Division of Cardiology, Department of Medicine, Cheil General Hospital & Women's Healthcare Center, Dankook University College of Medicine, Seoul; [4]Department of Laboratory Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Correspondence to
Eun Hee Lee, M.D.
Department of Laboratory Medicine, Green Cross Laboratories, 107 Ihyeon-ro 30beon-gil, Giheung-gu, Yongin 16924, Korea
Tel: +82-2-578-0649
Fax: +82-2-578-0649
E-mail: ehlee@gclabs.co.kr

Background/Aims: In multicenter clinical trials, laboratory tests are performed in the laboratory of each center, mostly using different measuring methodologies. The purpose of this study was to evaluate coefficients of variation (CVs) of laboratory results produced by various measuring methods and to determine whether mathematical data adjustment could achieve harmonization between the methods.

Methods: We chose 10 clinical laboratories, including Green Cross Laboratories (GC Labs), the central laboratory, for the measurement of total cholesterol, high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), serum triglycerides, creatinine, and glucose. The serum panels made with patient samples referred to GC Labs were sent to the other laboratories. Twenty serum samples for each analyte were prepared, sent frozen, and analyzed by each participating laboratory.

Results: All methods used by participating laboratories for the six analytes had traceability by reference materials and methods. When the results from the nine laboratories were compared with those from GC Labs, the mean CVs for total cholesterol, HDL-C, LDL-C, and glucose analyzed using the same method were 1.7%, 3.7%, 4.3%, and 1.7%, respectively; and those for triglycerides and creatinine analyzed using two different methods were 4.5% and 4.48%, respectively. After adjusting data using Deming regression, the mean CV were 0.7%, 1.4%, 1.8%, 1.4%, 1.6%, and 0.8% for total cholesterol, HDL-C, LDL-C, triglyceride, creatinine, and glucose, respectively.

Conclusions: We found that more comparable results can be produced by laboratory data harmonization using commutable samples. Therefore, harmonization efforts should be undertaken in multicenter trials for accurate data analysis (CRIS number; KCT0001235).

Keywords: Multicenter clinical trial; Harmonization; Comparable data; Traceability

INTRODUCTION

Lasting and continuous competence is imperative as the field of medicine is constantly evolving. The need to investigate multiple populations in medical research has led to multicenter and even multinational trials. In multicenter trials, many challenges arise that are not present in single-institution studies, particularly with respect to clinical laboratory results. To overcome challenges associated with compiling and comparing results

from different laboratories, standardization or harmonization of laboratory test results is needed. However, standardization and harmonization consume much time and resources, creating obstacles in multicenter trials [1].

In multicenter trials, clinical laboratory testing can be performed in the laboratory of each participating site or in a central laboratory. In general, routine tests, such as complete blood cell counts, general chemistries, and urinalysis, are usually performed at the respective laboratory of each participating site; and special tests, such as drug concentration and genetic testing, are performed in a central laboratory. Because researchers may assume that various methods for measuring or evaluating an analyte produce the same results, they may not be aware of variability in results between methods, which could result from a lack of traceability. As all laboratories do not use the same analytic methods, measurement principles, calibrators, and reagents, test results may vary based on the laboratory, which makes comparison of test results from different laboratories difficult. In multicenter trials carried out to establish diagnostic or therapeutic guidelines or to aid in drug development, variance of clinical laboratory test results caused by the use of different measurement methods should be considered in evaluating results for optimal guidelines and efficient drug development. If these variances are not considered, the accuracy of the analysis may suffer and result in negative clinical, technical, financial, and regulatory consequences [2].

In its 2015 survey, the College of American Pathologists (CAP), one of the largest external quality assessment organizations in the world, showed that the inter-assay coefficient of variation (CV) of total cholesterol was 3.3% across all methods, 3.9% for creatinine, and 3.2% for hemoglobin A1c [3]. Those CVs were much lower than those of other analytes.

As international guidelines were established in hyperlipidemia and diabetes, standardization in measuring glucose and lipids were developed by using reference materials and methods [1]. Standardization and harmonization are processes used to equalize results derived using different methods. Standardization can be accomplished by relating the result to a reference through a documented, unbroken chain of calibration. When such a reference is not available, harmonization is used to equalize results utilizing a consensus approach, such as application of an agreed-upon method mean [2]. However, accurate standardization rely on securing traceability from reagent manufacturers. Each laboratory in a trial must check the traceability of reagents in clinical laboratory tests before the trial begins. Although traceability may be confirmed, commutability of reference materials should also be considered. Without such standardization, harmonization, and traceability, accurate interpretation of trial results may be difficult. This study was conducted to evaluate the CV of laboratory results produced by various measuring methods and to determine whether mathematical data adjustment could achieve harmonization between the methods.

METHODS

Materials and methods
This study is part of the Cooperative Network Construction of a Nationwide Clinical Trial study [4] to evaluate the characteristics of and treatment strategies in patients with hypertension in 37 Korean centers. Of these 37 centers, the laboratories of nine centers were investigated in this study. This study was approved by the Institutional Review Board of Cheil General Hospital & Women's Healthcare Center (approval number: CGH-IRB-2013-33) and its associated centers, and written informed consent was obtained from all patients.

Basic data gathering
Nine laboratories (labeled A through I) as well as Green Cross Laboratories (GC Labs), the reference laboratory, were included in this study (Fig. 1). Instruments, analytic methods, reagents, lot information of reagents, and traceability of calibrators were surveyed for the six test items, serum total cholesterol, high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides, creatinine, and glucose, in all 10 laboratories. The traceability chain is shown in Fig. 2.

Manufacturing serum panel
A serum panel with 20 concentrations for each analyte was created for method comparison. Remaining patient samples at GC Labs were used to make the serum pan-
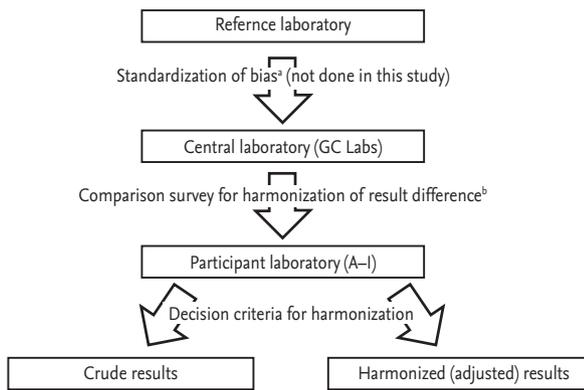
Lee SG, et al. Lab harmonization in clinical trial

KJIM

Figure 1. The hierarchy of participating laboratories. The calibration hierarchy in this study is as follows: a reference measurement procedure, a central laboratory (Green Cross Laboratories [GC Labs]), and nine participant laboratories. The uncertainty increases from top to bottom. [a]Estimation of bias between method of central laboratory and standard reference method can achieve comparable results by having calibration traceable to a reference measurement procedure, [b]Estimation of difference between method of participant laboratory and central laboratory can achieve comparable results among different measurement procedures.
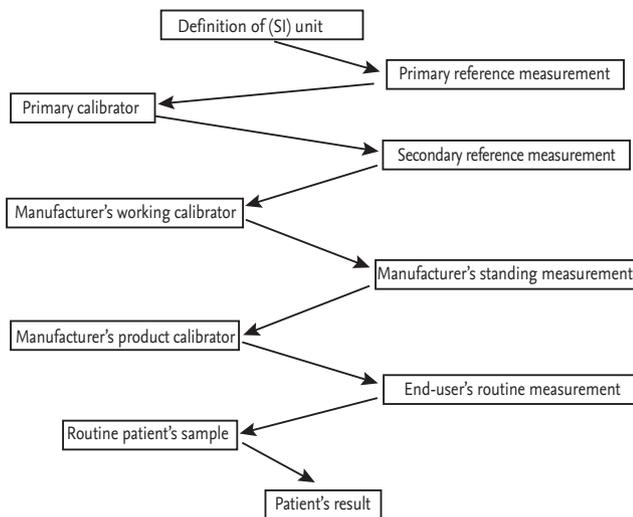

Figure 2. The traceability chain of a clinical test result. Traceability refers to the property of a measurement that has been related to references through an unbroken chain of comparisons. The procedure for equipment calibration should be designed and implemented so as to ensure that calibrations and measurements made by the laboratory are traceable to the International System (SI) of units.

els according to the Clinical & Laboratory Standards Institute EP9-A2 guideline. After 40 mL of the pooled serum samples were centrifuged at 3,000 rpm at 2°C to 8°C for 10 minutes, the supernatant was separated; and the supernatant was mixed for 10 hours under refrigerated conditions. Following filtration with a 0.22 μM filter (MF membrane, Sigma-Aldrich Co. LLC, St. Louis, MO, USA), 300 μL per tube were dispensed and kept in a deep freezer at –70°C before transportation to each participating laboratory under freezing conditions.

Measurement method

Each laboratory was asked to store samples in a deep freezer the day before measurement. All samples transported to each laboratory were analyzed in 14 days after being frozen. The analysis was done at the same time in both GC Labs and each laboratory to exclude bias due to pre-analytical conditions. Samples were thawed in a refrigerator for 1 hour and then were mixed for 30 minutes on a roller mixer. Each sample was analyzed in duplicate. The second measurement was done in reverse order compared to the first one. All measurements were done within 2 hours. Test results were reported in integers for serum total cholesterol, HDL-C, LDL-C, triglycerides, and glucose. Creatinine was reported to two decimal places.

Analytic method and traceability

Instruments used to measure the analytes in this study were as follows: modular analytics (Roche Diagnostics, Manheim, Germany) for GC Labs; Cobas 8000 c702 (Roche Diagnostics) for laboratories A, E, and G; TBA-2000FR (Toshiba Medical System Corporation, Tochigi, Japan) for laboratories B, D, H, and I; ADVIA1800 (Siemens Healthcare Diagnostics, Marburg, Germany) for laboratory C; and TBA 200FR NEO (Toshiba Medical System Corporation) for laboratory F (Supplementary Table 1). Reagent and calibrator information for each analyte are described in Supplementary Table 2. Ranges of CVs for the six analytes at all concentrations are shown in Supplementary Table 3. The number of laboratories using the same calibrator and assigned values as the reference laboratory, GC Labs, was two for total cholesterol, one for LDL-C, two for triglycerides, two for glucose, and none for HDL-C and creatinine.

As seen in Table 1, the analytic method for total cho-

Table 1. Analytic methods and traceability information of the 10 laboratories

Laboratory	Assay principle	Traceability
Total cholesterol		
GC Labs	Enzymatic	IDMS/AK
A	Enzymatic	IDMS/AK
B	Enzymatic	NIST SRM911
C	Enzymatic	NCEP/CDC
D	Enzymatic	ReCCS JCCRM211
E	Enzymatic	IDMS/AK
F	Enzymatic	ReCCS JCCRM211
G	Enzymatic	IDMS/AK
H	Enzymatic	IDMS/AK
I	Enzymatic	NIST SRM911
HDL-C		
GC Labs	Direct method	CDC reference method
A	Direct method	CDC reference method
B	Direct method	ReCCS JCCRM224
C	Direct method	NCEP designated comparison method
D	Direct method	ReCCS JCCRM224
E	Direct method	CDC reference method
F	Direct method	ReCCS JCCRM224
G	Direct method	CDC reference method
H	Direct method	ReCCS JCCRM224
I	Direct method	ReCCS JCCRM224

GC Labs, Green Cross Laboratories; IDMS, isotope dilution mass spectrometry; AK, Abell-Kendall; NIST, National Institute of Standards and Technology; SRM, Standard Reference Material; NCEP, National Cholesterol Education Program; CDC, Centers for Disease Control and Prevention; ReCCS, Reference Material Institute for Clinical Chemistry Standards; JCCRM Japanese Serum Primary Reference Materials; HDL-C, high density lipoprotein cholesterol.

lesterol in all 10 laboratories was uniformly enzymatic. All methods were traceable with isotope dilution mass spectrometry or the Abell-Kendall method. For HDL-C and LDL-C measurements, direct methods were used (Table 1). Direct methods do not include pre-analytical processes of ultracentrifugation, precipitation, and calculation steps, which make direct methods suitable for auto-analyzers. Every laboratory used the direct method using cationic detergent, and all laboratories used reagents traceable to the Centers for Disease Control and Prevention reference method (Table 1) [5-7]. All laboratories used enzymatic methods for triglycerides. Five of the 10 laboratories used glycerol blank methods. For serum creatinine measurements, nine laboratories used the Jaffe method, one used an enzymatic method, and two did not use adjustment of pseudo-creatinine

chromogen [8,9]. All used the hexokinase method to measure glucose. Traceability was ensured with reference materials for glucose.

Statistics

For statistical analysis, EP evaluator release 11 (Data Innovations LLC, South Burlington, VT, USA) and Excel 2000 (Microsoft Corp., Redmond, WA, USA) were used. CV was calculated for central and each laboratory variation from each sample. The interassay CVs for each analyte was calculated by averaging the observed CVs over all 20 samples. We aimed to find the line of best-fit using by Deming regression equation, which was defined y-axis as a reference standard and x-axis as each laboratory result ($y = \beta_0 + \beta_1 \times X$). To evaluate the effect of harmonization using Deming regression analysis, in-

Lee SG, et al. Lab harmonization in clinical trial

KJIM

Table 2. Comparison data of the six analytes

Partic-ipant	Total cholesterol				HDL-C				LDL-C				Triglycerides				Creatinine				Glucose			
	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %	Correlation coefficient	Bias, %	CV, %	Adjusted CV, %
A	0.9979	0.3	0.8	0.8	0.9997	2.7			0.9970	7.7	5.2	1.9	0.9997	2.4	1.7	0.8	0.9997	5.0	3.5	1.9	0.9996	1.5	1.2	0.9
B	0.9976	−3.4	2.5	0.8	0.9996	−9.7			0.9980	−5.6	4.1	1.5	0.9995	−10.2	7.7	1.2	0.9995	−4.2	3.0	1.1	0.9997	−1.5	1.1	0.6
C	0.9975	−3.8	2.8	0.9	0.9911	0.2			0.9864	10.4	6.7	5.3	0.9994	4.6	3.2	0.8	0.9994	−2.0	3.2	1.4	0.9996	−4.7	3.4	1.1
D	0.9992	−0.7	0.8	0.6	0.9988	3.8			0.9996	3.9	2.7	0.7	0.9977	−12.1	9.3	1.9	0.9977	−10.9	8.2	1.0	0.9998	0.4	0.7	0.6
E	0.9990	0.7	0.7	0.6	0.9995	9.4			0.9982	6.6	4.4	1.9	0.9996	1.2	1.0	0.9	0.9996	−0.1	1.4	1.4	0.9993	1.7	1.4	0.9
F	0.9966	2.5	2.1	1.2	0.9966	4.1			0.9995	5.5	3.8	0.9	0.9975	−3.9	3.9	1.9	0.9975	73	4.9	1.3	0.9975	4.9	3.4	1.2
G	0.9995	0.5	0.5	0.4	0.9992	3.1			0.9978	9.4	6.3	2.0	0.9995	0.4	1.0	1.1	0.9995	0.2	1.8	1.9	0.9997	−0.8	1.0	0.7
H	0.9996	5.0	3.5	0.4	0.9983	0.7			0.9993	2.4	1.8	1.1	0.9978	−6.6	5.4	1.7	0.9978	15.6	9.9	1.6	0.9997	2.6	1.8	0.8
I	0.9994	2.9	2.0	0.5	0.9984	2.6			0.9993	4.8	3.3	1.0	0.9975	−9.3	7.1	2.0	0.9975	3.1	4.3	2.7	0.9998	1.7	1.3	0.8
Mean	-	-	1.7	0.7	-	-	3.7	0.7	-	-	4.3	1.8	-	-	4.5	1.4	-	-	4.5	1.6	-	-	1.7	0.8

HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; CV, coefficient of variation.

terassay CVs before and after harmonization were calculated. Differences in the proportion of diseases before and after adjustment for lab harmonization were analyzed by a proportion test. For box plot and strip charts, R general-public-license version (R Foundation, Vienna, Austria) was used.

RESULTS

Correlation among methods

To evaluate the correlation among methods at each location, the results from each of the nine laboratories were compared with those of GC Labs. The correlation coefficient was over 0.9 for all six analytes tested (Table 2).

Comparison according to analytic methods

The mean CVs for total cholesterol, HDL-C, LDL-C, and glucose analyzed by the same method as that used by GC Labs were 1.7%, 3.7%, 4.3%, and 1.7%, respectively; and those for triglycerides and creatinine, which were analyzed by two different methods in the 10 laboratories, were 4.5% and 4.5%, respectively (Table 2). After data adjustment with Deming regression, interassay CV decreased significantly compared with that before adjustment in total cholesterol, HDL-C, and LDL-C, and especially in triglycerides and creatinine (both, $p < 0.05$) but not in serum glucose ($p > 0.05$) (Table 2, Fig. 3). When comparing the results with reference criteria for CV, the number of samples falling within the criteria was increased for all analytes after adjustment. The proportion of samples within reference CV criteria for six analytes increased from 67.2% ± 11.4% before adjustment to 93.7% ± 5.1% after adjustment (Table 3). Those laboratories that used the same reagents, calibrator, and assigned values as GC Labs showed lower interassay CVs than other laboratories (Table 4).

Different methods demonstrated differences in results, especially for the triglycerides measurement: using or not using a glycerol blank resulted in a significant difference. Laboratories using glycerol blanks showed a mean percent difference of −4.4%, while those not using a glycerol blank showed a −1.3% difference. For serum creatinine analysis, methods that compensated for interference by the Jaffe-like chromogen showed smaller differences than those with no compensation.

Table 3. Comparison of CV (before and after adjustment) with reference criteria

Participant	Adjust-ment	No.[a] & CV, %	Reference criteria					
			Cholesterol total, ≤ 3.0%[b]	HDL-C, < 4.0%[b]	LDL-C, ≤ 4.0%[b]	Triglycerides, ≤ 5.0%[b]	Creatinine, ≤ 3.2%[c]	Glucose, ≤ 2.8%[c]
A	Before	Number[a]	20	20	5	20	14	19
		Mean CV, %	0.8	1.9	5.2	1.7	3.5	1.2
	After	Number[a]	20	20	18	20	18	20
		Mean CV, %	0.8	0.6	1.9	0.8	1.9	0.9
B	Before	Number[a]	13	2	10	6	13	20
		Mean CV, %	2.5	7.5	4.1	7.7	3.0	1.1
	After	Number[a]	20	20	19	20	19	20
		Mean CV, %	0.8	1.0	1.5	1.2	1.1	0.6
C	Before	Number[a]	10	9	12	19	13	6
		Mean CV, %	2.8	4.8	6.7	3.2	3.2	3.4
	After	Numbera	19	14	8	20	18	20
		Mean CV, %	0.9	3.1	5.3	0.8	1.4	1.1
D	Before	Numbera	20	17	17	2	0	20
		Mean CV, %	0.8	2.9	2.7	9.3	8.2	0.7
	After	Numbera	20	19	20	18	20	20
		Mean CV, %	0.6	1.4	0.7	1.9	1.0	0.6
E	Before	Numbera	20	0	11	20	18	19
		Mean CV, %	0.7	6.3	4.4	1	1.4	1.4
	After	Numbera	20	20	18	20	18	20
		Mean CV, %	0.6	0.8	1.9	0.9	1.4	0.9
F	Before	Number[a]	16	16	13	15	6	9
		Mean CV, %	2.1	2.9	3.8	3.9	4.9	3.4
	After	Number[a]	18	19	20	17	18	19
		Mean CV, %	1.2	1.5	0.9	1.9	1.3	1.2
G	Before	Numbera	20	18	4	20	14	20
		Mean CV, %	0.5	2.8	6.3	1	1.8	1.0
	After	Number[a]	20	20	16	20	14	20
		Mean CV, %	0.4	1.0	2.0	1.1	1.9	0.7
H	Before	Number[a]	4	18	18	11	6	16
		Mean CV, %	3.5	1.9	1.8	5.4	9.9	1.8
	After	Number[a]	20	19	19	19	18	20
		Mean CV, %	0.4	1.7	1.1	1.7	1.6	0.8
I	Before	Number[a]	18	18	15	8	10	18
		Mean CV, %	2.0	2.2	3.3	7.1	4.3	1.3
	After	Number[a]	20	19	20	18	13	20
		Mean CV, %	0.5	1.3	1.0	2.0	2.7	0.8
Total (A–I)	Before	No. (%)[d]	141 (78)	118 (66)	105 (58)	121 (67)	94 (52)	147 (82)
	After	No. (%)[d]	177 (98)	170 (94)	158 (88)	172 (96)	156 (87)	179 (99)

CV, coefficient of variation; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol.

[a]Number of samples within criteria among the 20 samples of each analyte.

[b]National Cholesterol Education Program test performance guideline [10].

[c]Desirable specifications for total error, imprecision, and bias, derived from intra- and inter-individual biologic variation [20], recently updated in the database located on the Westgard homepage (https://www.westgard.com/biodatabase1.htm).

[d]Number of samples within criteria among total 180 samples.

Comparison between enzymatic methods and the Jaffe method yielded a large percent difference, 9.60%.

Comparison of disease prevalence before and after lab harmonization

According to our proposed lab harmonization method, prevalence of diseases in 593 lab data from nine centers was compared before and after adjustment. The prevalence of dyslipidemia increased to 46.2% from 39.63% (p = 0.0012) and the prevalence of chronic kidney disease (defined as estimated glomerular filtration rate < 60 mL/min/1.73 m²) dropped to 8.26% from 20.57%.
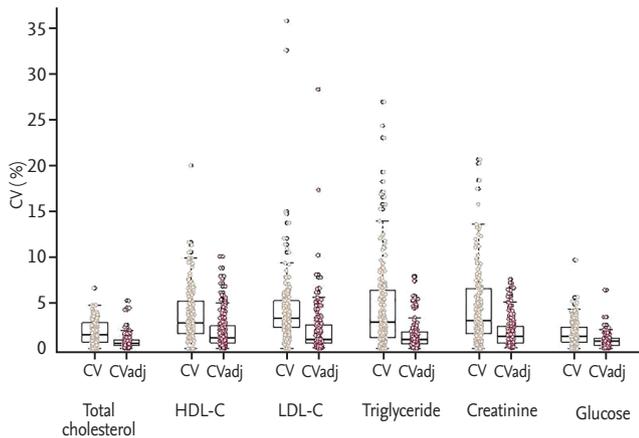

Figure 3. The comparison of inter-assay coefficient of variation (CV) before and after adjustment (CVadj). With the data adjusted based on Deming regression analyses between the central laboratory and each laboratory, mean CVs of all analytes decreased. HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol.

DISCUSSION

This study showed that result variation caused by different analytical methods can be reduced by harmonization. Harmonization may become a prerequisite in multicenter trials. The compatibility of data generated by multiple laboratories is not guaranteed due to different methods, reagents, calibrators, etc., used; and management of data from multiple sites is difficult and requires more effort for statistical analysis than what is needed in single-center studies [10].

When clinical laboratory tests may be performed by the laboratory of each participating site, more discussion and consideration in the planning stage of research should be given to whether a single, central clinical laboratory should be used instead of the laboratories at each center. If multiple laboratories will be analyzing samples, researchers should explore how to adjust or compare data prior to initiating the study.

This study was carried out to establish a method of postanalytical harmonization using data from various laboratories. The clinically acceptable total error, including precision and bias, is reported in international guidelines, the literature, and reports from external quality assessment organizations [11-13]. For example, in the CAP survey (external quality assessment), inter-assay CV among all methods for total cholesterol is 3.0% to 3.4%, 3.9% to 15.5% for creatinine, and 3.2% to 9.2% for glucose, which are rather good [3]. Accuracy of the analytic method used in clinical laboratories is established by a standardization process using the results from patient samples, the hierarchy structure, and traceability of the analytical measurement system (Fig.

Table 4. CV before and after adjustment by deming regression

Test item	Calibrator and assigned value			
	Same group		Different group	
	Mean CV, %	Mean CV after deming regression, %	Mean CV, %	Mean CV after deming regression, %
Total cholesterol	0.8	0.7	2.0	0.7
LDL-C	5.2	1.9	4.1	1.8
Triglycerides	1.3	0.9	5.4	1.5
Glucose	1.3	0.9	1.8	0.8

CV, coefficient of variation; LDL-C, low density lipoprotein cholesterol.

1) [14]. This association ensure traceability of analytical methods. Traceability applies a primary reference material with proper accuracy and precision, by cascades of reference material, to the calibrator by the manufacturer, and finally to the patient samples (Fig. 2). The interassay CV of test results from different laboratories is small for an analytical method with traceability that uses the same calibrator adjusted with the same primary reference material. This is standardization. All reagents used in this study have traceability. Thus, when compared with external quality assessment scheme results, all analytes had better or similar results except triglycerides (Supplementary Table 3).

Our study results are also concerning. Although interassay CVs were within acceptable limits according to external quality assessment standards and comparison studies revealed good correlation coefficients, the results of some of the participating laboratories should be evaluated for adjustment against international guidelines like those from the National Cholesterol Education Program. International guidelines propose strict criteria, and not all our results could meet those requirements. In studies where the analysis of the crude data could influence the study's conclusion, failure to adjust the results could lead to challenges to the integrity of the study. Traceability is particularly important in multicenter trials where the different laboratories utilize various testing methods. However, even where the methods are traceable, differences between the various methods should not be ignored because commutability of reference material, application of the calibrated assigned value, and method compensation, such as the creatinine determination method, could be problematic [2]. Thus, traceability of methods coupled with studies comparing the results from different laboratories during the planning stages of multicenter trials would guide adjustments of results so that analysis of results from multiple sites would be accurate. As inter-assay CVs of the laboratories using the same calibrator and assigned value showed lower variance than the CVs of those that did not, traceability should be checked first when selecting analytical methods in multicenter trials.

Different methods impact CV differently. As the analytic method for triglycerides is affected by the use of a glycerol blank, the method use by each laboratory should be surveyed and considered in the interpretation of test results. Measuring triglyceride concentration without a glycerol blank has a high positive correlation with free glycerol concentration, which was also shown by regression analysis [15], so that methods with or without glycerol blank could be compared by regression analysis. In the current study, when results from the nine participating laboratories were adjusted based on the method used in the central laboratory, interassay CVs were then greatly improved. Compared with the 2015 CAP survey, interassay CVs of the participating laboratories were high. The reason may be that, in our group, 50% used a method with a glycerol blank, whereas, in the CAP group, only 10% used a method with a glycerol blank [3]. Data adjustment would greatly improve accuracy when analytic methods of participating laboratories demonstrate systemic bias.

In creatinine analysis, the Jaffe method and enzymatic method can be used. The differences between the analytic methods depend on avoiding interference of pseudo-creatinine chromogen, such as proteins, antibiotics, and ketones. Different measurements using the Jaffe method may yield significant analytic errors depending on whether compensation was made for those chromogens [8] because creatinine is a very low-concentration analyte. For such reasons, interassay CVs before and after adjustment were higher for creatinine than for other analytes. After adjustment, the interassay CV for creatinine decreased dramatically, and the correlation coefficient was good.

To measure glucose, all laboratories used the hexokinase method. As traceability was secured with reference methods and materials, the mean CV for glucose was excellent at 1.7% (0.7% to 3.4%).

Using a single, assigned analytic method in a central laboratory would be best in multicenter trials. However, because of the logistical challenges to sample transportation and the need for timely analysis, analyzing samples in the laboratories of the participating center where the sample was taken is how many multicenter trials currently manage the analysis. In this situation, data harmonization is an option. Laboratory data harmonization is especially needed when reference methods or materials are absent or when reference material is non-commutable. As described above, calibrator values and method characteristics can influence results and, thus, analysis of data. For detecting these possible ob-

Lee SG, et al. Lab harmonization in clinical trial

KJIM

stacles to accurate analysis and deciding on harmonization, pre-study surveys of the participating laboratories should be considered. In multicenter trials, direct comparison without data adjustment is best practice provided no variance in results existed; but, in older studies, traceability was seldom addressed, complicating comparisons between data sets. This study confirms the need to compare assays and methods used in obtaining older data and to adjust data if needed. In the future, if a central laboratory could compare its results with those of a reference laboratory, such as Cholesterol Reference Method Laboratory Network, and calculate bias, the central and participating laboratories could coordinate using a hierarchical structure permitting the central laboratory to communicate the reference material obtained to the participating laboratories and creating consistency in measurement of analytes across the participating laboratories.

The major limitation of our study is no external validation of this model in clinic. Although the prevalence of diseases such as dyslipidemia and chronic kidney disease changed significantly after application of our proposed lab harmonization method, it is still strongly required to test whether this regression reduces the interlaboratory gap. Further studies are needed.

Perspective

This study will encourage researchers to focus on traceability and improve the quality of study results through harmonization. For analytes with available reference materials and methods, standardization would be the best approach; and harmonization should be applied to the analytes for which reference materials and reference methods are not available or cannot be developed. Additional studies on harmonization of reference ranges may broaden the scope of the interpretation of clinical laboratory results [16].

KEY MESSAGE

1. By comparing the mean coefficients of variation of laboratory results between participating laboratories and central laboratory, more comparable results were produced.
2. Laboratory data harmonization can be a worthy alternative to compare all participating laboratories and so strongly recommended for accurate data analysis in multicenter clinical trials.

REFERENCES

1. Miller WG, Tate JR, Barth JH, Jones GR. Harmonization: the sample, the measurement, and the report. Ann Lab Med 2014;34:187-197.

2. Greg Miller W, Myers GL, Lou Gantzer M, et al. Roadmap for harmonization of clinical laboratory measurement procedures. Clin Chem 2011;57:1108-1117.

3. Horowitz GL, Alter DN, Baskin LB, et al. Hemoglobin A1c (5 Challenge) Participant Summary. Northfield (IL): College of American Pathologists, 2015.

4. Kim SA, Kim JY, Park JB. Significant interarm blood pressure difference predicts cardiovascular risk in hypertensive patients: CoCoNet study. Medicine (Baltimore) 2016;95:e3888.

5. Rifai N, Warnick GR, McNamara JR, Belcher JD, Grinstead GF, Frantz ID Jr. Measurement of low-density-lipoprotein cholesterol in serum: a status report. Clin Chem 1992;38:150-160.

6. Sugiuchi H, Uji Y, Okabe H, et al. Direct measurement of high-density lipoprotein cholesterol in serum with polyethylene glycol-modified enzymes and sulfated alpha-cyclodextrin. Clin Chem 1995;41:717-723.

7. Matsuzaki Y, Kawaguchi E, Morita Y, et al. Evaluation of two kinds of reagents for direct determination of HDL-cholesterol. Seibutsu Shiryo Bunseki 1996;19:419-427.

8. Chung HJ, Chun S, Min WK. Creatinine determination with minimized interference. J Lab Med Qual Assur 2008;30:229-231.

9. Zawta B, Delanghe J, Van Den Noortgate N, Taes Y, Lam-

eire N, Engel W. Arithmetic compensation for pseu-do-creatinine Jaffé method and its effect on creatinine clearance results. Clin Chem 2001;47:A148-A149.

10. Chahal AP. A knowledge-based process for offshoring clinical trials. J Clin Res Best Pract 2006;2. http://www.firstclinical.com/journal/2006/0601_Offshoring.pdf.

11. Sacks DB, Arnold M, Bakris GL, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Diabetes Care 2011;34:e61-e99.

12. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002;106:3143-3421.

13. Ricos C, Alvarez V, Cava F, et al. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59:491-500.

14. Myers GL, Kimberly MM, Waymack PP, Smith SJ, Cooper GR, Sampson EJ. A reference method laboratory network for cholesterol: a model for standardization and improvement of clinical laboratory measurements. Clin Chem 2000;46:1762-1772.

15. Choi JS, Shon BR, Kim YK. A clinical study on the need for glycerol-blanking in triglycerides measurement. Korean J Clin Pathol 1995;15:204-211.

16. Ichihara K. Statistical considerations for harmonization of the global multicenter study on reference values. Clin Chim Acta 2014;432:108-118.

Supplementary Table 1. Instruments in each participating laboratory

Laboratory	Instrument	Manufacturer
GC Labs	Modular analytics	Roche Diagnostics, Manheim, Germany
A	Cobas 8000, c702	Roche Diagnostics, Manheim, Germany
B	TBA-2000FR	Toshiba Medical Systems Corporation, Tochigi, Japan
C	ADVIA1800	Siemens Healthcare Diagnostics, Marburg, Germany
D	TBA-2000FR	Toshiba Medical Systems Corporation, Tochigi, Japan
E	Cobas 8000, c702	Roche Diagnostics, Manheim, Germany
F	TBA 200FR NEO	Toshiba Medical Systems Corporation, Tochigi, Japan
G	Cobas 8000, c702	Roche Diagnostics, Manheim, Germany
H	TBA-2000FR	Toshiba Medical Systems Corporation, Tochigi, Japan
I	TBA-2000FR	Toshiba Medical Systems Corporation, Tochigi, Japan

GC Labs, Green Cross Laboratories.

Supplementary Table 2. Reagents and calibrator information for the 10 laboratories

Participant laboratory	Reagent	Manufacturer	Calibrator	Calibratorlot no.	Calibrator assigned value, mg/dL
Total cholesterol					
GC Labs	CHOL	Roche[a]	C.f.a.s	175190	158
A	CHOL2	Roche	C.f.a.s	175190	158
B	Determiner-C-TC	Kyowa[b]	HDL.LDL-C	235ADG	202
C	Total cholesterol	Siemens[c]	Chemistry CAL	459400	183
D	Determiner-C-TC	Kyowa	Liquid cal.	153081	221
E	CHOL2	Roche	C.f.a.s	175190	158
F	Determiner-C-TC	Kyowa	Liquid cal.	164051	221
G	CHOL2	Roche	C.f.a.s	176123	160
H	Cholesterol N	Denka Seiken[d]	C.f.a.s	177953	175
I	Determiner-C-TC	Kyowa	HDL.LDL-C	235ADG	202
HDL-C					
GC Labs	HDLC3	Roche	C.f.a.s Lipids	172156	76.5
A	HDLC3	Roche	C.f.a.s Lipids	174327	75
B	Determiner-L	Kyowa	HDL.LDL-C	235ADG	74.9
HDL-C					
C	HDL-C	Siemens	HDL/LDL	308123	68
D	Determiner-L HDL-C	Kyowa	HDL.LDL-C	235ADG	74.9
E	HDLC3	Roche	C.f.a.s Lipids	178923	79
F	Determiner-L HDL-C	Kyowa	HDL.LDL-C	234ADD	74.9
G	HDLC3	Roche	C.f.a.s Lipids	175839	75
H	Determiner-L HDL-C	Kyowa	HDL.LDL-C	235ADG	74.9
I	Determiner-L HDL-C	Kyowa	HDL.LDL-C	235ADG	74.9
LDL-C					
GC Labs	LDL-C	Roche	C.f.a.s Lipids	172156	146
A	LDL-C	Roche	C.f.a.s Lipids	174327	146
B	Determiner-L LDL-C	Kyowa	HDL.LDL-C	235ADG	110
C	LDL-C	Siemens	HDL/LDL Cholesterol CAL	308123	128
D	Determiner-L LDL-C	Kyowa	HDL.LDL-C	235ADG	110
E	LDL-C	Roche	C.f.a.s Lipids	178923	166
F	Determiner-L LDL-C	Kyowa	HDL.LDL-C	235ADD	110
G	LDL-C	Roche	C.f.a.s Lipids	175839	159
H	Determiner-L LDL-C	Kyowa	HDL.LDL-C	235ADG	110
I	Determiner-L LDL-C	Kyowa	HDL.LDL-C	235ADG	110

Lee SG, et al. Lab harmonization in clinical trial

KJIM

Supplementary Table 2. Continued

Participant laboratory	Reagent	Manufacturer	Calibrator	Calibratorlot no.	Calibrator assigned value, mg/dL
Triglycerides					
GC Labs	TG	Roche	C.f.a.s	175190	127
A	TRIGL	Roche	C.f.a.s	175190	127
B	Determiner-C-TG	Kyowa	Liquid Calibrator	235ADG	77
C	TG2	Siemens	Chemistry CAL	459400	186
D	Determiner-C-TG	Kyowa	Liquid Calibrator	153081	151
E	TRIGL	Roche	C.f.a.s	175190	127
F	Determiner-C-TG	Kyowa	Liquid Calibrator	164051	153
G	TRIGL	Roche	C.f.a.s	176123	134
H	Determiner-C-TG	Kyowa	C.f.a.s	17795301	146
I	Determiner-C-TG	Kyowa	Liquid Calibrator	235ADG	77
Creatinine					
GC Labs	CREA	Roche	C.f.a.s	175190	4.29
A	CREJ2	Roche	C.f.a.s	175190	4.2
B	CREA	Roche	Liquid Calibrator	153081	4.3
C	Creatinine	Siemens	Chemistry CAL	459400	9.12
D	Creatinine FS	DiaSys[e]	Trucal U	19846	3.65
E	CREJ2	Roche	C.f.a.s	175190	4.21
F	CREA	Roche	Liquid calibrator	164051	4.15
G	CREJ2	Roche	C.f.a.s	176123	4.06
H	Creatinine FS	DiaSys	C.f.a.s	177953	3.9
I	L-CRE-L	Shinyang[f]	C.f.a.s	175190	4.04
Glucose					
GC Labs	GLU	Roche	C.f.a.s	175190	200
A	GLUC3	Roche	C.f.a.s	175190	200
B	Glucose	Denka Seiken	Liquid Calibrator	153081	203
C	Glucose	Siemens	Chemistry CAL	459400	244
D	Glucose	Denka Seiken	Liquid Calibrator	153081	203
E	GLUC3	Roche	C.f.a.s	175190	200
F	Glucose	Denka Seiken	Liquid Calibrator	164051	188
G	GLUC3	Roche	C.f.a.s	176123	195
H	L-type glucose 2	Wako[g]	C.f.a.s	17795301	205
I	Glucose Hexokinase FS	DiaSys	C.f.a.s	175 190-01	196

GC Labs, Green Cross Laboratories; C.f.a.s, calibrator for automated systems; TC, total cholesterol; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; CAL, calibrator; TG, triglyceride; TRIGL, triglyceride; CREA, creatinine; CREJ2, creatinine Jaffe generation 2; FS, fluid stable; GLU, glucose; GLUC3, glucose hexokinase generation 3.

[a]Roche Diagnostics, Manheim, Germany.
[b]Kyowa, Tokyo, Japan.
[c]Siemens Healthcare Diagnostics, Marburg, Germany.
[d]Denka Seiken, Tokyo, Japan.
[e]DiaSys, Waterbury, CT, USA.
[f]Shinyang, Seoul, Korea.
[g]Wako, Richmond, VA, USA.

Supplementary Table 3. Ranges of coefficients of variance for the six analytes at all concentrations

Test item	CAP 2015 results		Current investigation	
	CV, %	No. of participants	CV, %	No. of participants
Total cholesterol	3.0–3.4	4,524–4,536	0.9–3.7	10
HDL-C	11.4–29.2	4,430–4,445	1.5–8.1	10
LDL-C	15.1–29.9	2,525–2,542	1.4–15.2	10
Triglycerides	4.4–5.1	4,510–4,545	1.7–16.5	10
Creatinine	3.9–15.5	5,427–5,449	1.3–14.5	10
Glucose	3.2–9.2	5,425–5,448	0.4–4.2	10

Values are presented as range.

CAP, College of American Pathologists; CV, coefficient of variation; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol.