Laboratory Management **Edited by: E. Wieland**

Hyojin Chae, Jaeeun Yoo, Joonhong Park, Kyoungho Cha, Jeong Joong Lee, Bongrae Cho, Myungshin Kim* and Yonggoo Kim*

Analysis of a 6-year pilot external quality assurance survey of free light chain using Sigma metrics

https://doi.org/10.1515/labmed-2018-0175
Received November 5, 2018; accepted August 19, 2019; previously published online September 14, 2019

Abstract

Background: A pilot external quality assurance (EQA) survey for the free light chain (FLC) assay was developed and implemented in Korea.
Methods: Survey data over 6 years (2010–2015) were collected retrospectively and Sigma metrics were calculated for method-specific peer groups.
Results: Nineteen to 29 laboratories participated in the EQA survey, and nephelometric (20%) and turbidimetric (80%) methods were used. Using a previously published clinically relevant reference change value (RCV) of 54.5% as the tolerance limit, the method-specific median Sigma metrics of kappa (κ) and lambda (λ) FLC achieved greater than Three-Sigma for 86–97% of all EQA distributions, and Five-Sigma for 48–72% of all distributions.
Conclusions: This EQA analysis of FLC assay applied clinically relevant quality specifications using Sigma metrics. During the 6-year EQA survey, we found that most of the results from participating laboratories meet clinically relevant quality specifications. In addition, method-specific differences were noted for λ FLC, at FLC concentrations above the initial measuring range that require a sample dilution.

Keywords: external quality assurance; free light chain; Sigma metric; Sigma Proficiency Assessment Chart.

Introduction

Free light chain (FLC) assays quantitatively measure light chains that are not bound to immunoglobulin heavy chains, and detect clonal restriction through a skewed ratio of free kappa (κ) to free lambda (λ) light chains [1]. The assay uses polyclonal antibodies that bind exclusively to the hidden epitopes of FLC molecules located at the junction of the immunoglobulin heavy and light chains in an intact immunoglobulin. These epitopes are only accessible when the light chain molecules are not bound to the heavy chain, which makes the assay highly specific. The use of latex enhancement further enhances the sensitivity of the assay to a few mg/L, and the assay can be performed by turbidimetry or nephelometry on a number of automated laboratory instruments.

The FLC assay has become standard practice and is recommended in various guidelines for diagnosis, prognostic assessment and evaluation of patient response to treatment in multiple myeloma and related plasma cell disorders [2–6]. However, technical concerns with the FLC assay include non-linearity and antigen excess, FLC polymerization, possibility of undetected “private” epitopes, lack of international standards, lot-to-lot variability of the reagent and inter-instrument variability.

In this regard, an external quality assurance (EQA) survey for the quantitative FLC assay was developed and implemented in collaboration with the Quality Assurance Committee of the Korean Society for Laboratory Medicine. We analyzed the 6-year EQA survey data and applied Sigma metrics to evaluate the observed analytical characteristics of the FLC assay and to assess the quality of the FLC assay that is being achieved.

***Corresponding authors: Myungshin Kim**, MD **and Yonggoo Kim**, MD**,** Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea; and Catholic Laboratory Development and Evaluation Center, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea, Phone: +82 2 2258 1645 (M. Kim); +82 2 2258 1642 (Y. Kim), E-Mail: microkim@catholic.ac.kr (M. Kim); yonggoo@catholic.ac.kr (Y. Kim)
Hyojin Chae, Jaeeun Yoo and Joonhong Park: Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea; and Catholic Laboratory Development and Evaluation Center, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
Kyoungho Cha, Jeong Joong Lee and Bongrae Cho: Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Materials and methods

Preparation of EQA materials and design of the EQA survey

The EQA material was derived from surplus serum from patients with monoclonal gammopathies and from healthy controls, and a serum pool was prepared as described in the Clinical and Laboratory Standards Institute document C37-A [7]. FLC concentrations were titrated to examine medical decision points, linearity, antigen excess and lower limit of quantitation. Serum material manufactured for quality assurance (QA) measures (Binding Site, Birmingham, UK) and commercial quality control (QC) materials purchased from Randox (Randox, Antrim, UK) were also used.

Three levels of EQA specimens were distributed biannually to the participating laboratories. The specimens were shipped frozen with ice pack and were transported by registered mail. Instructions regarding the correct processing of the specimens upon receipt, care needed for reconstruction, storage, and deadlines for sample analysis and reporting were provided with each dispatch of EQA materials. For each sample, the participating laboratory reported the κ and λ FLC results and the κ/λ ratios. The participating laboratories were given a unique laboratory code known only to the EQA coordinator, staff and the laboratory, for anonymization and confidentiality.

Two types of reports were provided for each level of EQA material. One was a summary report with the mean, standard deviation (SD) and standard deviation index (SDI) of all participant results. SDI expresses the reported result as the number of SDs it is from the mean value (SDI = [result – mean]/SD).

The other report was specific to each method peer group. Due to the lack of an international standard material or a reference measurement procedure for FLC, the FLC EQA was designed as a consensus program with major performance goals (criteria) based on the laboratory consensus. For each type of EQA report, the mean of all participants' results and the mean of each method peer group results were regarded as the target values, respectively. Those results distributed more than 3 SD away from the mean were regarded as outliers and removed from statistical analyses [8].

Application of Sigma metrics and statistical analysis of FLC EQA survey data

Survey data from 32 EQA samples distributed in 11 EQA surveys over 6 years (from 2010 to 2015) were collected retrospectively. An Excel spreadsheet was used to tabulate the observed means, SDs and medians for the method subgroups to calculate observed bias, coefficient of variation (CV), total error (TE) and Sigma metrics, and to plot %bias and %CV for each subgroup for each specimen in each episode on the Sigma Proficiency Assessment Chart [9]. Sigma metrics were calculated as follows [9]:

$$\text{Sigma metric}_{\text{peer group}} = (\text{TEa} - |\%\text{bias}|)/\text{CV}_{\text{peer group}}$$

$$\text{TE} = \text{Bias} + 1.65\ \text{SD or CV}$$

Due to the small number of laboratories comprising specific instrument peer groups, and for a simplified analysis, only method-specific peer groups were defined. The method-specific bias was estimated as the difference from the weighted average. The weighted average was calculated to account for the number of laboratories in each method subgroup [9]. Due to the lack of a well-established consensus definition of quality requirements regarding the FLC assay, multiple surrogate definitions of quality requirements were assessed (Table 1). The Sigma Proficiency Assessment Chart was developed using a previously published reference change value (RCV) of 54.5% [12] derived from the estimate of monoclonal FLC variability in patients with plasma cell disorders as the tolerance [13] in calculating the Sigma metrics for κ FLC and λ FLC. Of note, the RCV value of 54.5% was chosen because it was comparable to the reported interlaboratory reproducibility values achieved

Table 1: Multiple surrogate definitions of quality requirements found in the literature.

Surrogate definitions of quality requirements							
	Desirable biological [10]	Minimal [11]	Desirable [11]	Optimal [11]	RCV [12]	RCV (modular P) [1]	RCV (BNII) [1]
κ FLC	8.0	16.1	10.7	5.4	54.5	77	69
λ FLC	8.6	19.3	12.9	6.4	54.5	121	106

FLC, free light chain; RCV, reference change value.

during the UK National External Quality Assessment Service (UKNEQAS) FLC EQA scheme (CV of FLC assays was >50%, and often over 100%) [14].

Continuous data are presented as mean (SD) for normally distributed data and median (range) for non-normally distributed data. Categorical data are presented as numbers (%). Normality was assessed using the D'Agostino and Pearson normality tests. Continuous variables were compared by the Mann-Whitney U-test and/or one-way analysis of variance for non-normally distributed data and categorical variables were compared by chi-squared (χ^2) tests. MedCalc version 14.12.0 (MedCalc Software, Mariakerke, Belgium) was used for statistical analyses. A p-value < 0.05 was considered statistically significant.

This study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki and was approved by the Institutional Review Board (IRB)/Ethics Committee of Seoul St. Mary's Hospital (IRB No. KC10SISI0037).

Results

Nineteen to 29 laboratories participated in the EQA survey. All participating laboratories used the FLC Freelite® kits manufactured by Binding Site. Nephelometric (20%) and turbidimetric (80%) methods were used for the FLC assay, and different instruments including Siemens BNII (Siemens Healthineers, Deerfield, IL, USA), Siemens BN-Prospec (Siemens Healthineers), Roche Modular P (Roche, Mannheim, Germany), Hitachi analyzers (Hitachi, Tokyo, Japan), Toshiba (Toshiba, Tokyo, Japan), Siemens Advia 1650 (Siemens Healthineers) and Binding Site SPAPLUS (Binding Site) were used (Table 2).

Method-specific total error and Sigma metrics

The method-specific median TE values of κ FLC across 29 EQA specimens were 26% for nephelometry and 25% for turbidimetry. The method-specific median TE values of λ FLC across 29 EQA specimens were 24% and 19% for nephelometry and turbidimetry, respectively.

As the Sigma metric is inversely proportional to $CV_{peer\ group}$ and EQA distributions with low FLC concentrations (less than the lower limit of the reference interval) are associated with a disproportionately large $CV_{peer\ group}$, three κ FLC and three λ FLC EQA distributions with low FLC concentrations were not included in the Sigma metrics analysis. The method-specific median Sigma metrics of κ FLC across 29 EQA specimens were 4.9 for nephelometric methods and 4.9 for turbidimetric methods. The method-specific median Sigma metrics of λ FLC across 29 EQA specimens were 4.9 for nephelometric methods and 5.4 for turbidimetric methods. There were no statistically significant differences in TE and Sigma metrics between turbidimetric methods and nephelometric methods for κ FLC; however, for λ FLC, nephelometric methods showed larger TE and lower Sigma metrics compared with turbidimetric methods ($p < 0.05$) (Figure 1).

Method harmonization

We compared the median reported results of the nephelometric and turbidimetric methods for each EQA specimen. Due to the small number of laboratories reporting nephelometric results, the Bland-Altman plot was used. The Bland-Altman plot showed that the average of the difference between the nephelometric and turbidimetric

Table 2: Distribution of methods and instruments used for the FLC assay during the 6-year EQA survey.

	EQA event										
Method/manufacturer	**S10-02**	**S11-01**	**S11-02**	**S12-01**	**S12-02**	**S13-01**	**S13-02**	**S14-01**	**S14-02**	**S15-01**	**S15-02**
Nephelometry	8	7	6	5	5	4	3	5	4	3	3
Siemens BNII	7	6	5	4	4	4	3	5	4	3	3
Siemens BN-Prospec	1	1	1	1	1	0	0	0	0	0	0
Turbidimetric	12	12	19	19	21	20	21	22	23	24	25
Roche Modular/Hitachi	9	11	16	17	17	17	18	17	18	17	19
Toshiba	1	1	3	2	4	3	3	3	3	4	3
Siemens Advia	1	1	0	0	0	0	0	0	0	0	0
Binding Site SPAplus	0	0	0	0	0	0	1	2	2	3	3
All methods	20	19	25	24	26	24	24	27	27	27	28

EQA, external quality assurance; FLC, free light chain.

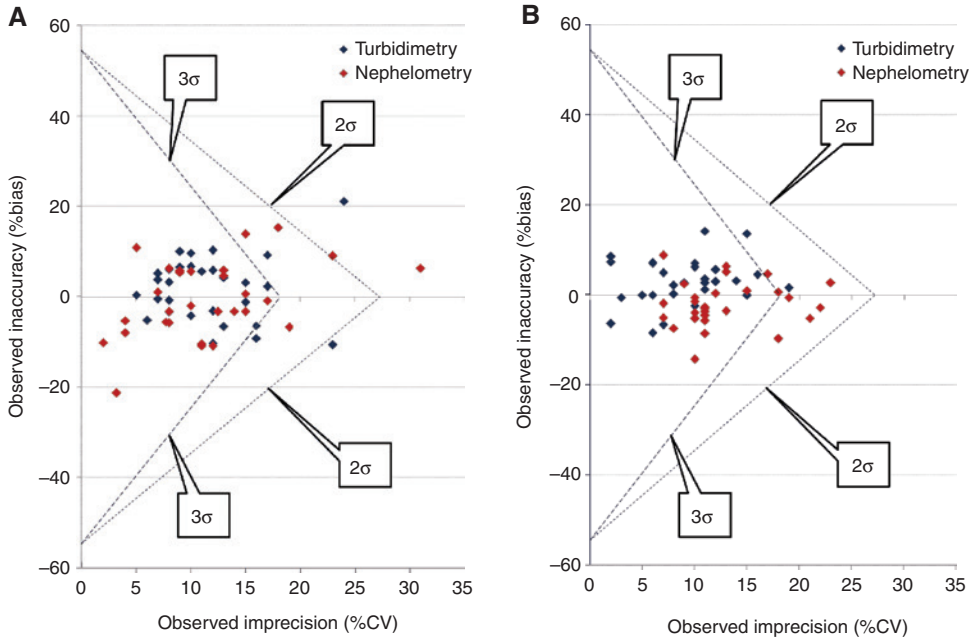

Figure 1: Sigma Proficiency Assessment Chart for κ FLC (A) and λ FLC (B) with TEa = 54.5%.
Diagonal lines from top to bottom represent 2σ and 3σ. The x-coordinate represents the interlaboratory precision observed for the method-specific peer group. The y-coordinate represents the bias compared to the laboratory consensus value used as the major performance goal of the EQA survey.

methods for κ FLC was 7.2 mg/L and for λ FLC was 42.1 mg/L (Figure 2). While the average difference between the two methods for κ FLC can be regarded as insignificant, for λ FLC there seems to be a significant difference between the two methods, and the difference seems to be due to higher λ FLC concentrations that would require a different sample dilution from that of the initial measuring range.

Interlaboratory reproducibility

The method-specific median interlaboratory reproducibility (CV) of κ FLC across 32 EQA specimens was 12% for both nephelometric and turbidimetric methods. The method-specific median interlaboratory reproducibility of λ FLC across 32 EQA specimens was 12% and 10% for the nephelometric and turbidimetric methods, respectively (Table 3).

The κ FLC and λ FLC values in healthy individuals occur at the lower end of the measuring range, within the manufacturer's reference intervals of 3–19 mg/L for κ FLC and 6–26 mg/L for λ FLC. FLC concentrations that border on the upper reference limit of the reference interval (medical decision point) could potentially be misclassified as normal or abnormal, and discordant results could occur by simply repeating the assay on those samples.

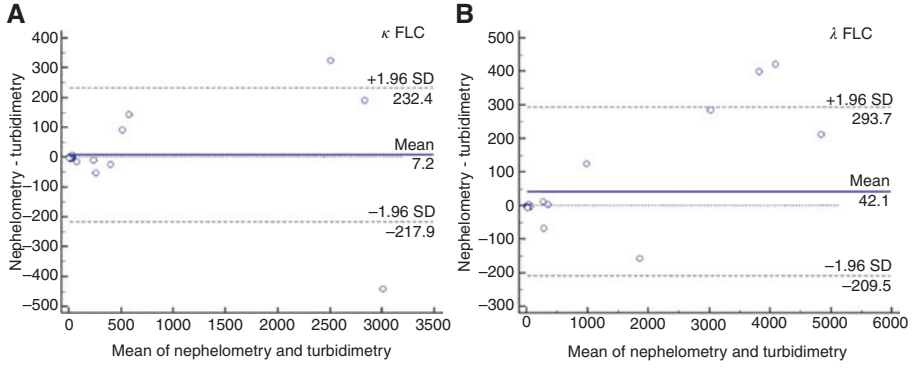

Figure 2: Bland-Altman plot of median reported results of nephelometry and turbidimetry for each EQA specimen.
Shown are absolute difference plots for κ FLC (A) and λ FLC (B).

Table 3: Statistical summary of the κ and λ FLC EQA results for all specimens distributed.

	EQA event																											
Method			**E10-02**		**E11-01**			**E11-02**				**E12-01**		**E12-02**		**E13-01**			**E13-02**			**E14-01**		**E14-02**			**E15-01**			**E15-02**	
	S10-02-1	**S10-02-2**	**S10-02-3**	**S11-01-1**	**S11-01-2**	**S11-02-1**	**S11-02-2**	**S11-02-3**	**S12-01-1**	**S12-01-2**	**S12-01-3**	**S12-02-1**	**S12-02-2**	**S13-01-1**	**S13-01-2**	**S13-02-1**	**S13-02-2**	**S13-02-3**	**S14-01-1**	**S14-01-2**	**S14-01-3**	**S14-02-2**	**S14-02-3**	**S15-01-1**	**S15-01-2**	**S15-01-3**	**S15-02-1**	**S15-02-2**	**S15-02-3**
Nephelometry																												
Mean	12.2	13.6	1.8	13.0	33.1	18.6	64.1	24.0	27.0	15.0	2793.3	231.1	13.6	647.3	18.0	2673.3	15.3	12.7	12.0	2925.2	12.4	14.9	231.4	11.9	559.7	8.4	380.7	15.5	2.6
%bias	11	5	6	5	14	6	−11	−1	1	6	−8	−2	1	15	−3	6	−3	−5	−10	6	−11	−6	−10	−3	9	−7	−3	−5	−21
CV (%)	5	9	13	13	15	8	12	17	7	10	4	10	15	18	15	31	14	4	11	9	11	8	2	8	23	19	12.5	7.7	3.2
TEa	19	20	27	26	39	19	31	29	13	22	14	18	25	45	28	58	26	12	29	21	29	19	13	16	47	38	24	18	26
Sigma	10.9	6.0	4.2	4.2	3.6	6.8	4.5	3.2	7.8	5.4	13.6	5.4	3.6	3.0	3.6	1.8	3.9	13.6	4.9	6.0	4.9	6.8	27.2	6.8	2.4	2.9	4.4	7.1	17.0
Turbidimetric																												
Mean	10.3	12.5	1.8	12.0	26.0	16.6	78.6	24.3	26.6	13.6	3234.8	241.1	13.4	503.1	19.4	2347.8	16.4	14.1	14.7	2733.2	15.3	16.7	283.8	12.7	466.4	9.6	405.8	17.4	4.0
%bias	−6	−3	6	−3	−11	−5	9	0	0	−4	7	2	−1	−10	4	−7	4	5	10	−1	10	6	10	3	−9	7	3	6	21
CV (%)	16	12	12	8	23	6	17	5	7	10	10	17	8	12	13	13	7	7	10	15	9	11	12	8	16	9	15	8	24
TEa	33	23	26	17	48	15	37	9	12	21	23	30	14	30	26	28	15	17	26	26	25	24	30	16	36	22	28	19	61
Sigma	3.4	4.5	4.5	6.8	2.4	9.1	3.2	10.9	7.8	5.4	5.4	3.2	6.8	4.5	4.2	4.2	7.8	7.8	5.4	3.6	6.0	4.9	4.5	6.8	3.4	6.0	3.6	6.8	2.3
All methods																												
Weighted mean	11.0	12.9	1.7	12.4	29.1	17.5	71.9	24.2	26.7	14.2	3031.0	235.6	13.5	560.9	18.6	2512.8	15.8	13.4	13.4	2765.5	13.9	15.8	257.3	12.3	513.2	9.0	393.3	16.4	3.3
CV (%)	15	11	12	10	23	8	18	8	7	10	11	17	9	16	14	17	7	8	12	14	12	11	14	8	18	11	14	9	27
Number of outliers	0	0	2	0	0	0	1	0	0	0	1	0	2	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0

Method			**E10-02**		**E11-01**		**E11-02**		**E12-01**			**E12-02**			**E13-01**			**E13-02**			**E14-01**			**E14-02**			**E15-01**			**E15-02**	
	S10-02-1	**S10-02-2**	**S10-02-3**	**S11-01-1**	**S11-01-2**	**S11-02-1**	**S11-02-2**	**S12-01-1**	**S12-01-2**	**S12-02-1**	**S12-02-2**	**S12-02-3**	**S13-01-1**	**S13-01-2**	**S13-01-3**	**S13-02-1**	**S13-02-2**	**S13-02-3**	**S14-01-1**	**S14-01-3**	**S14-02-1**	**S14-02-2**	**S14-02-3**	**S15-01-1**	**S15-01-2**	**S15-01-3**	**S15-02-1**	**S15-02-2**	**S15-02-3**
Nephelometry																														
Mean	13.9	41.7	3160.4	13.7	39.9	17.1	14.7	13.7	55.4	14.3	14.0	1775.1	18.3	247.7	4940.4	1.9	15.5	357.1	13.3	269.0	4297.2	14.8	13.4	13.8	25.8	1048.9	11.7	12.6	4028.4
%bias	−3	1	3	−4	−3	−4	−5	−1	−1	−3	1	−5	−1	−10	5	−5	−2	0	−6	3	5	−5	−4	−7	9	6	−9	−14	5
CV (%)	22	15	23	13	11	10	10	10	10	11	18	21	19	18	17	7	7	12	11	9	13	11	11	8	7	13	11	10	13
TEa	39	26	41	25	21	20	22	18	17	21	30	40	32	39	33	17	13	20	24	17	27	23	22	21	20	28	27	31	27
Sigma	2.5	3.6	2.4	4.2	5.0	5.4	5.4	5.4	5.4	5.0	3.0	2.6	2.9	3.0	3.2	7.8	7.8	4.5	4.9	6.1	4.2	5.0	5.0	6.8	7.8	4.2	4.9	5.4	4.2
Turbidimetric																														
Mean	14.7	44.2	2874.8	14.9	42.6	18.9	16.6	14.2	55.7	15.1	13.9	1932.5	18.7	313.3	4726.7	2.0	16.0	353.7	14.9	256.0	3873.8	16.2	14.9	16.0	21.7	924.3	13.9	16.7	3626.7
%bias	3	7	−7	5	4	6	7	2	0	3	0	3	2	14	0	0	1	−1	6	−2	−5	5	7	7	−8	−6	9	14	−5
CV (%)	9	6	7	7	11	10	10	8	6	11	5	14	19	11	8	15	11	3	12	10	10	16	6	2	6	2	2	15	12
TEa	18	17	18	16	22	23	24	15	10	21	8	26	33	32	13	25	19	6	25	19	22	31	17	11	18	10	12	38	25
Sigma	6.1	9.1	7.8	7.8	5.0	5.4	5.4	6.8	9.1	5.0	10.9	3.9	2.9	4.9	6.8	3.6	5.0	18.2	4.5	5.4	5.4	3.4	9.1	27.2	9.1	27.2	27.2	3.6	4.5
All methods																														
Weighted mean	14.3	41.3	3077.0	14.2	41.1	17.8	15.5	13.9	55.7	14.7	13.9	1872.2	18.4	274.3	4718.6	2.0	15.8	355.7	14.1	262.4	4081.6	15.5	13.9	14.9	23.7	986.7	12.8	14.7	3831.2
CV (%)	19	12	20	12	11	10	11	10	9	11	16	20	11	19	17	9	7	11	12	9	13	12	11	9	9	13	12	15	13
Number of outliers	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	2	0	0	0	0	0	0	0	0	0	0	1	0	0

CV, coefficient of variation; EQA, external quality assurance; FLC, free light chain; TEa, total allowable error.

Regarding interlaboratory reproducibility at values closest to the medical decision point, four EQA specimens had weighted average values near the κ or λ medical decision points. The all-method interlaboratory reproducibility values (CV) were 8%, 14% and 9% at κ FLC concentrations of 17.5, 18.6 and 16.4 mg/L, respectively, and 9% at a λ FLC concentration of 23.7 mg/L.

Discussion

The FLC assay is extensively used for the diagnosis, prognostic assessment and management of patients with monoclonal gammopathy. The assay can be used on several types of automated analyzers and has been adopted for use in many laboratories nationwide in Korea as an indispensable resource. However, as with any clinical laboratory test, EQA (also known as proficiency testing) is considered a mandatory component of a laboratory quality system [15]. National EQA schemes for the FLC assay have been developed and implemented in France, United Kingdom and Czech [16], and a number of international schemes also exist. The lack of a national EQA scheme in Korea prompted the development and implementation of an EQA survey for the FLC assay in collaboration with the Quality Assurance Committee of the Korean Society for Laboratory Medicine. There have been more than 30 distributions over a 6-year period, and over 35 laboratories nationwide participated in the scheme.

In this retrospective analysis of the EQA scheme, we applied clinically relevant quality specifications using Sigma metrics. Sigma metrics combine bias, imprecision and the allowable TE and convert them into an overall assessment of the analytical quality of the test [15]. For that purpose, the previously described Sigma Proficiency Assessment Chart was used [9].

However, due to the lack of a well-established consensus definition of quality requirement regarding the FLC assay, multiple surrogate definitions of quality requirements, including those based on biological variations [12, 17, 18] and/or measurement uncertainties [1], were assessed. We chose a reported RCV value of 54.5% derived from the estimate of monoclonal FLC variability in patients with plasma cell disorders [12] and applied as tolerance limits in calculating the Sigma metrics for κ FLC and λ FLC. We acknowledge the limitations of using an RCV value as a quality specification as it does not have component for bias and therefore may not be an appropriate substitute for an allowable TE. Moreover, an RCV value of 54.5%, which is derived from the estimate of monoclonal FLC variability in patients with plasma cell disorders, is much larger than reported biological variation of FLC in healthy individuals. However, FLC is clinically a tumor marker and typically in myeloma the monoclonal FLC levels are variably increased above the upper limit of the reference interval. Therefore, a biological variation-based quality requirement may not be a clinically relevant quality specification for the FLC assay. In this regard, the RCV value of 54.5% was chosen because it was comparable to the reported interlaboratory reproducibility values achieved during the UKNEQAS FLC EQA scheme [14] and was similar to a clinical cut-point of 50% FLC reduction defining a clinically relevant FLC response reported in a previous clinical recommendation [10].

The method-specific Sigma metrics of κ and λ FLC achieved greater than Three-Sigma for 86–97% of all EQA distributions, and the higher quality specifications consistent with Five-Sigma, beyond which little improvement to quality can be achieved [11], was achieved in 48% of all κ FLC EQA distributions, for the nephelometric and turbidimetric methods, respectively, and 48% and 72% of all λ FLC EQA distributions for the nephelometric and turbidimetric methods, respectively. When only EQA distributions with consensus κ and λ FLC values in the vicinity of κ or λ medical decision points were considered, the method-specific Sigma metrics of κ and λ FLC achieved greater than Three-Sigma for all EQA distributions for both methods, and Five-Sigma was achieved in 67% and 100% of all κ and λ FLC EQA distributions, for the nephelometric and turbidimetric methods, respectively.

We also assessed the harmonization of results between the nephelometric and turbidimetric methods. While the average difference between the two methods is not significant for κ FLC, for λ FLC, the Bland-Altman plot showed a significant difference between the two methods at higher λ FLC concentrations that require a different sample dilution from that of the initial measuring range. This is not surprising as significant method-specific differences in FLC concentrations have been previously reported [14]. Hence, despite the use of the same source of commercial reagent, values were not completely harmonized between assay systems that use different method principles. Possible causes of discordant results between methods include different reaction times, differences in the antigen (sample)/antibody (reagent) ratio in the reaction mixture, different concentrations of polyethylene glycol, different sample dilution protocol and different linearity. The positive difference in nephelometry at higher λ FLC concentrations demonstrated in this study should be investigated further in a larger number of samples.

The method-specific median interlaboratory reproducibility (CV) of κ FLC and λ FLC across 32 EQA specimens ranged from 10 to 12% for both the nephelometric and turbidimetric methods. Interlaboratory reproducibility values reported in this survey were considered acceptable for clinical practice and were better than those achieved during the UK NEQAS external control cycles. Moreover, EQA specimens that had consensus values in the vicinity of κ or λ medical decision points achieved an all-method interlaboratory reproducibility value (CV) of 8–14%, which may be larger than, but not significantly different from, the manufacturer's predetermined acceptance criteria of total precision (%CV < 8.5%) and between-instrument precision (%CV < 8.5%), for a relatively narrow FLC concentration range without the need for serum dilution [19].

EQA provides an essential tool for medical laboratories for evaluating assay performances and establishing quality standards [20]. During the 6-year EQA survey, we found that most of the results from participating laboratories meet clinically relevant quality specifications. We could also provide an estimate of the interlaboratory variability in the FLC assay that participants and physicians should be aware of. In addition, we have highlighted the method-specific differences for λ FLC and also how they tend to increase with increasing FLC concentrations. This may have the potential to mislead clinical assessment of the FLC assay results when serial FLC results involving different methods are used.

There are acknowledged limitations of this study. First, the small number of participating laboratories comprising the nephelometric method subgroup may have introduced bias in the estimation of method-specific performance, despite the use of a weighted average. Second, due to the lack of a well-established consensus definition of quality requirement regarding the FLC assay, a reported RCV value of 54.5% derived from the estimate of monoclonal FLC variability in patients with plasma cell disorders was applied as the tolerance limit in calculating the Sigma metrics. Lastly, as a reference measurement procedure or certified reference material is not available for FLC, we could not assess the reported results of the EQA scheme with reference to trueness.

Despite these limitations, the data collected by the first national EQA scheme for the FLC assay in Korea represent a unique resource for clinical laboratories. We used Sigma metrics and a clinically relevant tolerance limit in analyzing the FLC EQA scheme. The study also shows that between-method differences still exit, and further standardization and harmonization efforts are necessary.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Wang L, Chan PC. Measurement uncertainty for serum free light chain assays: estimation and implication on result interpretation. Clin Biochem 2013;46:381–84.
2. Dispenzieri A, Kyle R, Merlini G, Miguel JS, Ludwig H, Hajek R, et al. International myeloma working group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. Leukemia 2009;23:215–24.
3. Moreau P, San Miguel J, Ludwig H, Schouten H, Mohty M, Dimopoulos M, et al. Multiple myeloma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013;24(Suppl 6):vi133–7.
4. Tan D, Chng WJ, Chou T, Nawarawong W, Hwang SY, Chim CS, et al. Management of multiple myeloma in Asia: resource-stratified guidelines. Lancet Oncol 2013;14:e571–81.
5. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International myeloma working group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15:e538–48.
6. Pratt G, Jenner M, Owen R, Snowden JA, Ashcroft J, Yong K, et al. Updates to the guidelines for the diagnosis and management of multiple myeloma. Br J Haematol 2014;167:131–3.
7. Clinical and Laboratory Standards Institute. Preparation and validation of commutable frozen human serum pools as secondary reference materials for cholesterol measurement procedures. Approved Guideline. C37-A. Wayne, PA: Clinical and Laboratory Standards Institute, 1999.
8. Healy MJ. Outliers in clinical chemistry quality-control schemes. Clin Chem 1979;25:675–7.
9. Westgard JO, Westgard SA. Assessing quality on the Sigma scale from proficiency testing and external quality assessment surveys. Clin Chem Lab Med 2015;53:1531–5.
10. Dispenzieri A, Zhang L, Katzmann JA, Snyder M, Blood E, Degoey R, et al. Appraisal of immunoglobulin free light chain as a marker of response. Blood 2008;111:4908–15.
11. Jassam N, Yundt-Pacheco J, Jansen R, Thomas A, Barth JH. Can current analytical quality performance of UK clinical laboratories support evidence-based guidelines for diabetes and ischaemic heart disease? – a pilot study and a proposal. Clin Chem Lab Med 2013;51:1579–84.
12. Katzmann JA, Snyder MR, Rajkumar SV, Kyle RA, Therneau TM, Benson JT, et al. Long-term biological variation of serum protein electrophoresis M-spike, urine M-spike, and monoclonal serum

free light chain quantification: implications for monitoring monoclonal gammopathies. Clin Chem 2011;57:1687–92.
13. Westgard JO, Westgard SA. The quality of laboratory testing today: an assessment of Sigma metrics for analytic quality using performance data from proficiency testing surveys and the CLIA criteria for acceptable performance. Am J Clin Pathol 2006;125:343–54.
14. Vercammen M, Meirlaen P, Broodtaerts L, Vande Broek I, Bossuyt X. Effect of sample dilution on serum free light chain concentration by immunonephelometric assay. Clin Chim Acta 2011;412:1798–804.
15. Westgard SA. Utilizing global data to estimate analytical performance on the Sigma scale: a global comparative analysis of methods, instruments, and manufacturers through external quality assurance and proficiency testing programs. Clin Biochem 2016;49:699–707.
16. Vávrová J, Maisnar V, Tichý M, Friedecký B, Čermáková Z, Dastych M, et al. Interlaboratory study of free monoclonal immunoglobulin light chain quantification. Clin Chem Lab Med 2011;49:89–92.
17. Ricos C, Alvarez V, Cava F, García-Lario JV, Hernández A, Jiménez CV, et al. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59:491–500.
18. Braga F, Infusino I, Dolci A, Panteghini M. Biological variation of free light chains in serum. Clin Chim Acta 2013;415:10–1.
19. U. S. Food & Drug Administration. 510(k) Substantial Equivalence in Determination Decision Memorandum. https://www.accessdata.fda.gov/cdrh_docs/reviews/k150658.pdf (updated on September 2015).
20. Vlašić Tanasković J, Coucke W, Leniček Krleža J, Vuković Rodriguez J. Peer groups splitting in Croatian EQA scheme: a trade-off between homogeneity and sample size number. Clin Chem Lab Med 2017;55:539–45.