

Joannes F.M. Jacobs^{a,*}, Katherine A. Turner^a, Maria Stella Graziani, Jody L. Frinack, Michael W. Ettore, Jillian R. Tate, Ronald A. Booth, Christopher R. McCudden, David F. Keren, Julio C. Delgado, Galina Zemtsovskaja, Robert O. Fullinfaw, Anna Caldini, Theo de Malmanche, Katina Katakouzinos, Matthew Burke, Giovanni Palladini, Sara Altinier, Martina Zaninotto, Gabriella Righetti, Marie Therese Melki, Stephen Bell and Maria Alice Vieira Willrich*

An international multi-center serum protein electrophoresis accuracy and M-protein isotyping study. Part II: limit of detection and follow-up of patients with small M-proteins

<https://doi.org/10.1515/cclm-2019-1105>

Received October 25, 2019; accepted December 17, 2019

Abstract

Background: Electrophoretic methods to detect, characterize and quantify M-proteins play an important role in the management of patients with monoclonal gammopathies (MGs). Significant uncertainty in the quantification and limit of detection (LOD) is documented when M-proteins are <10 g/L. Using spiked sera, we aimed to assess the variability in intact M-protein quantification and LOD across 16 laboratories.

Methods: Sera with normal, hypo- or hyper-gammaglobulinemia were spiked with daratumumab or elotuzumab, with concentrations from 0.125 to 10 g/L (n=62) along with a beta-migrating sample (n=9). Laboratories blindly analyzed samples according to their serum protein electrophoresis (SPEP)/isotyping standard operating

procedures. LOD and intra-laboratory percent coefficient of variation (%CV) were calculated and further specified with regard to the method (gel/capillary electrophoresis [CZE]), gating strategy (perpendicular drop [PD]/tangent skimming [TS]), isotyping (immunofixation/immunosubtraction [ISUB]) and manufacturer (Helena/Sebia).

Results: All M-proteins ≥ 1 g/L were detected by SPEP. With isotyping the LOD was moderately more sensitive than with SPEP. The intensity of polyclonal background had the biggest negative impact on LOD. Independent of the method used, the intra-laboratory imprecision of M-protein quantification was small (mean CV=5.0%). Low M-protein concentration and high polyclonal background had the strongest negative impact on intra-laboratory precision. All laboratories were able to follow trend of M-protein concentrations down to 1 g/L.

Conclusions: In this study, we describe a large variation in the reported LOD for both SPEP and isotyping; overall

*Joannes F.M. Jacobs and Katherine A. Turner contributed equally to this work.

***Corresponding authors:** Joannes F.M. Jacobs, PhD, MD, Laboratory Medical Immunology, Department of Laboratory Medicine, Radboud University Medical Center, Geert Grootenhuisplein 10, 6525 GA Nijmegen, The Netherlands, Phone: +31 24 36 15335, E-mail: H.Jacobs@Radboudumc.nl; and

Maria Alice Vieira Willrich, PhD, Division of Clinical Biochemistry and Immunology, Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA, Phone: +1-507-284-2511, E-mail: willrich.mariaalice@mayo.edu

Katherine A. Turner, Jody L. Frinack and Michael W. Ettore: Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

Maria Stella Graziani: Section of Clinical Biochemistry, University of Verona, Verona, Italy

Jillian R. Tate [posthumous] and Matthew Burke: Department of Chemical Pathology, Pathology Queensland, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia

Ronald A. Booth and Christopher R. McCudden: Department of Pathology and Laboratory Medicine, The Ottawa Hospital, Ottawa, ON, Canada

David F. Keren: Department of Pathology, The University of Michigan, Ann Arbor, MI, USA

Julio C. Delgado: ARUP Laboratories, Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, USA

Galina Zemtsovskaja: Clinical Chemistry Laboratory, North Estonia Medical Centre, Tallinn, Estonia

Robert O. Fullinfaw: Department of Chemical Pathology, The Royal Melbourne Hospital, Melbourne, Victoria, Australia

Anna Caldini: General Laboratory, Careggi University Hospital, Florence, Italy

Theo de Malmanche: NSW Health Pathology, Immunology Department, John Hunter Hospital, New Lambton Heights NSW, Australia

Katina Katakouzinos: Immunopathology Department, Royal Prince Alfred Hospital, Camperdown, NSW, Australia

Giovanni Palladini: Amyloidosis Research and Treatment Center, Foundation IRCCS Policlinico San Matteo, and Department of Molecular Medicine, University of Pavia, Pavia, Italy

Sara Altinier and Martina Zaninotto: Laboratory Medicine of the University Hospital of Padova, Padova, Italy

Gabriella Righetti: Clinical Chemistry Laboratory, University of Verona, Verona, Italy

Marie Therese Melki: Sebia Inc., Lisses, France

Stephen Bell: Helena Biosciences Europe, Sunderland, UK

LOD is most affected by the polyclonal immunoglobulin background. Satisfactory intra-laboratory precision was demonstrated. This indicates that the quantification of small M-proteins to monitor patients over time is appropriate, when subsequent testing is performed within the same laboratory.

Keywords: accuracy; immunofixation; immunosubtraction; limit of detection; monoclonal proteins; precision; protein electrophoresis.

Introduction

Monoclonal gammopathies (MGs) are defined by the clonal expansion of plasma cells, resulting in the characteristic excretion of a monoclonal immunoglobulin (M-protein). MGs encompass a broad spectrum of clinical disorders ranging from asymptomatic, MG of undetermined significance (MGUS) to life-threatening diseases, such as multiple myeloma (MM) and amyloid light chain (AL) amyloidosis [1, 2].

M-protein detection and quantification are integral parts of the diagnosis and monitoring of MG [3]. M-protein diagnostics is most commonly performed using serum electrophoretic methods, supplemented with additional assays for quantification and clonality testing [4]. Serum protein electrophoresis (SPEP), using either agarose gel (AGE) or capillary electrophoresis (CZE), is often the first test to screen for MG. SPEP abnormalities are confirmed and isotypes using the more sensitive methods of immunofixation electrophoresis (IFE) or immunosubtraction (ISUB). SPEP is further used for quantification, in which the M-protein is gated (M-spike) either by perpendicular drop (PD) or tangent skimming (TS) [3]. M-protein quantification is used for disease staging, monitoring of disease activity and to assess response to therapy [5].

Quantitation of the M-protein spike involves laboratory staff selecting the limits of the monoclonal protein. This judgement remains subjective; therefore, this is a potential source of variation, particularly if the M-protein is small and/or migration is aligned with proteins such as transferrin and C3 in the beta region [6, 7]. Recognition of the imprecision and inaccuracy of measurements of small M-proteins is reflected in the International Myeloma Working Group (IMWG) guidelines that define a ‘measurable’ M-protein as one that meets at least one of the following three criteria: serum M-protein ≥ 10 g/L, urine M-protein ≥ 200 mg/24 h or serum involved free light chains (FLCs) ≥ 100 mg/L provided that the FLC ratio is abnormal [5].

With the introduction of more effective treatments during the last two decades, a vast majority of patients achieve treatment responses with decreasing M-protein concentrations over time [8]. There are however no strict guidelines with regard to analytical methodologies, or guidance as to when an M-spike is too small to permit quantification. Formal studies that document the limit of quantitation (LOQ) and limit of detection (LOD) are not available even though variation in the quantitation and detection of small M-proteins has a direct impact on the classification of response criteria. The importance of reliable routine follow-up of disease activity in patients with small M-proteins was further shown by Murray et al. who demonstrated that small M-proteins persist over time and these individuals develop progressive disease at similar rates compared to individuals with larger M-proteins [9].

In Part I of this study, we concluded that the quantitation of concentrations of M-proteins below 1 g/L results in poor accuracy and high variation between laboratories [10]. In this second part of the study, we investigate the M-protein LOD and the intra-laboratory coefficient of variation (CV) of the various electrophoretic methods and describe the parameters that most strongly affect the LOD and intra-laboratory imprecision.

Materials and methods

Large-volume sample pools with M-proteins in varying gamma fraction backgrounds were prepared in enough quantity to be shared with 16 different institutions in three different continents. For a full description of the materials and methods, we refer to Part I of this study published in this same issue of Clinical Chemistry and Laboratory Medicine [10].

Results

Limit of detection for SPEP and IFE/ISUB

Both Daratumumab (Dara) and Elotuzumab (Elo) are of the IgG κ isotype and represent M-proteins in the gamma fraction. Both biologics have differing migration patterns, as Dara migrates at the cathodal end of the gamma fraction, whereas Elo migrates in the center of the gamma fraction on gel electrophoresis and at the anodal end of the gamma region by CZE. The IgA λ M-protein could be identified as a distinct peak migrating in the beta region (Figure 1).

As the concentration of the M-protein decreases, so does the number of laboratories that can detect it with

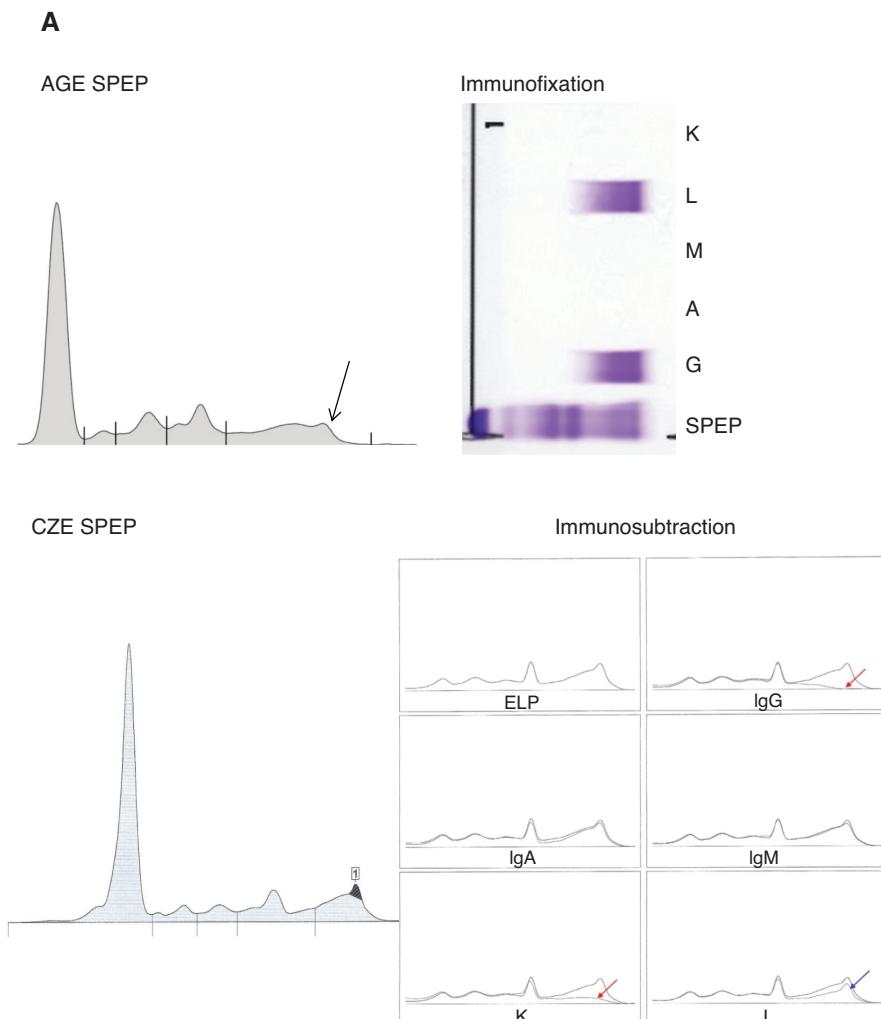


Figure 1: Representative immunofixation electrophoresis and immuno subtraction in a normal-gamma polyclonal background.

Images show the SPEP electropherogram and isotype of three M-proteins. Both Dara and Elo are IgG-k biologics, Dara migrates at the cathodal end of the gamma fraction and Elo migrates in the center of the gamma fraction. The IgAλ M-protein be identified as a distinct peak migrating in the beta region. (A) Daratumumab (1 g/L). AGE SPEP (Helena Laboratories), IFE (Sebia), CZE SPEP and ISUB (Sebia Capillarys) are represented. There is subtraction of the M-protein spike by anti-IgG and anti-kappa (red arrows), but with anti-lambda, although polyclonal lambda immunoglobulins are subtracted, the spike remains (blue arrow). (B) Elotuzumab (1 g/L). AGE SPEP (Helena Laboratories), IFE (Sebia), CZE SPEP and ISUB (Sebia Capillarys) are represented. There is subtraction of the M-protein spike by anti-IgG and anti-kappa (red arrows), but with anti-lambda, although polyclonal lambda immunoglobulins are subtracted, the spike remains (blue arrow). (C) Beta-migrating IgAλ M-protein. AGE SPEP (Helena Laboratories), IFE (Sebia), CZE SPEP and ISUB (Sebia Capillarys) are represented. There is subtraction of the M-protein spike by anti-IgA and anti-lambda (red arrows), but with anti-kappa, although polyclonal lambda immunoglobulins are subtracted, the spike remains (blue arrow).

either SPEP or IFE/ISUB. All M-proteins ≥ 1 g/L were detected by SPEP (Tables 1 and 2). The level of polyclonal gammaglobulin background had the biggest impact on the SPEP LOD, and M-proteins of 0.5 g/L were detected in all samples with low gammaglobulinemia. The concentration at which the detection of Dara and Elo becomes variable was slightly lower with IFE/ISUB (Figure 2 and Table 3) as compared to the SPEP in the different polyclonal backgrounds. All M-proteins ≥ 0.1 g/L in sera with low

gammaglobulinemia were detected by IFE. The increase of gamma fraction background led to less detection by both SPEP and IFE/ISUB. Less pronounced was the effect of the migrating properties of the M-protein on the LOD. Elo was detected more often by SPEP and IFE/ISUB as compared to the more cathodal-migrating Dara. However, in the hypo-gamma samples, detection by IFE/ISUB was not influenced by the migration pattern as both Dara- and Elo-spiked samples had the same detection rate. The

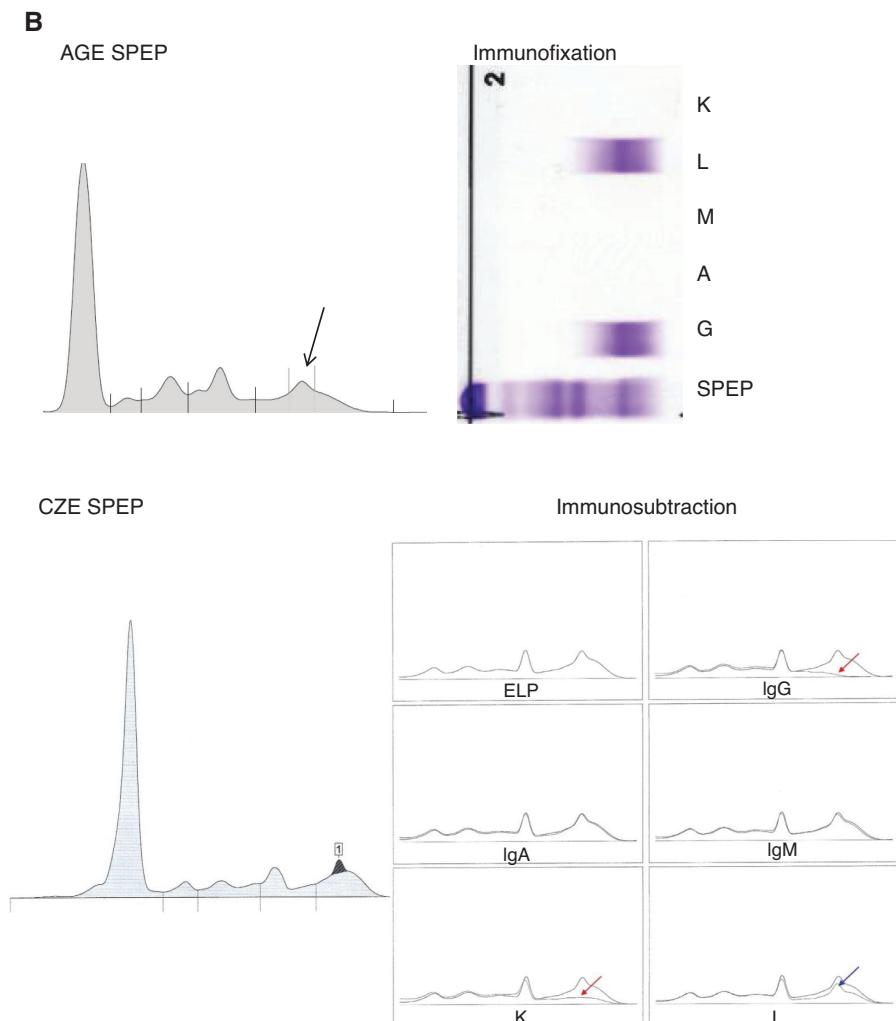


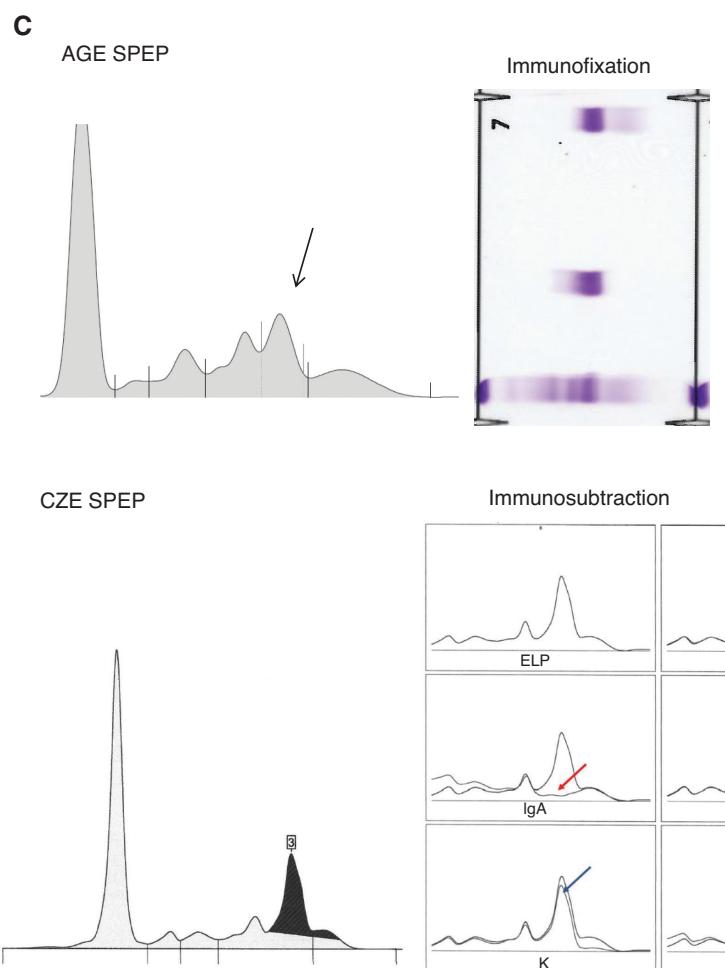
Figure 1 (continued)

LOD of the M-protein migrating in the beta region was 0.5 g/L, and abnormalities in all these samples were detected using both SPEP and IFE/ISUB.

A more detailed analysis of the individual methods reveals the strengths and weaknesses of each method. The SPEP LOD for each method shown in Figure 2 and Table 2 was analyzed based on the combined data of PD and TS as the gating technique should not have an effect on the analytical sensitivity of the method and was defined as the lowest concentration of the M-protein where all samples were detected. LODs between methods (AGE vs. CZE) and provider of the analytical system (Helena vs. Sebia) were comparable. However, once the detection became variable, the observed LOD values were lowest using CZE for the qualitative detection of low-concentration M-proteins in all sample types. No significant differences were observed between the SPEP LOD of Elo and the more cathodal-migrating Dara. As all concentrations were detected of the

beta-migrating M-protein, no further differentiation in the methods was possible.

The different isotyping methods showed minimal differences in terms of LOD (Table 3). Focusing on these small differences shows that the use of pentavalent antisera yielded the highest sensitivity regardless of the M-protein concentration for both Helena and Sebia systems. However, it must be stated that only two institutes reported results from pentavalent testing. No substantial differences in sensitivity were observed between IFE and ISUB. However, in hypo-gamma samples, IFE was more sensitive and detected all M-proteins down to 0.1 g/L and performed marginally better compared to ISUB that was able to detect eight of 18 (44%) of the M-proteins at a concentration of 0.1 g/L. In hyper-gamma samples with 1.0-g/L and 0.5-g/L M-proteins, the M-protein was not detected using monovalent antisera on the Sebia system in, respectively, two of 54 (4%) and 22 of 53 samples (4%).

**Figure 1** (continued)**Table 1:** Limit of detection for SPEP and immunofixation/immunosubtraction.

	Dara			Elo Beta			<0.5
	Hypo	Normal	Hyper	Hypo	Normal	Hyper	
SPEP LOD							
Helena AGE	0.5	0.5	1.0	0.5	0.5	1.0	<0.5
Sebia AGE	0.5	0.5	1.0	0.5	0.5	1.0	<0.5
Sebia CZE	0.5	0.5	0.5	0.5	0.3	0.5	<0.5
IFE/ISUB LOD							
Helena							
Mono	<0.1	0.5	<0.5	<0.1	0.3	<0.5	<0.5
Penta	<0.1	<0.1	<0.5	<0.1	<0.1	<0.5	—
Sebia							
Mono	<0.1	0.5	1	<0.1	0.5	2	<0.5
Penta	<0.1	0.3	<0.5	<0.1	<0.1	<0.5	—
Sebia ISUB							
ISUB	0.3	0.5	<0.5	0.3	0.3	0.5	<0.5

SPEP LOD and IFE/ISUB LOD (g/L) were defined as the lowest M-protein concentration in which an M-protein was detected and qualitatively reported in all samples analyzed. Samples marked with (<) are cases where the LOD was below the tested concentrations.

For the beta-migrating sample, all dilutions were detected by all methods of IFE/ISUB and a LOD was never reached with the tested concentrations.

Within-laboratory variation of small M-protein quantification

In Part I we demonstrated a poor accuracy for the quantification of small M-proteins and significant variation in results between laboratories, even between laboratories that made use of identical methods [10]. However, as individual patients are often monitored within the same laboratory, we set out to investigate the intra-laboratory CV from two measurements performed blindly on two independent sample sets. In the total of 1203 duplicate measurements of M-proteins ranging from 1 to 10 g/L, the overall mean intra-laboratory CV was 5.0% (95% confidence interval [CI] 4.7%–5.4%). Parameters that strongly affect intra-laboratory variation are the M-protein concentration

Table 2: Qualitative M-protein detection by serum protein electrophoresis.

Expected recovery, g/L	10	8.0	6.0	5.0	4.0	3.0	2.0	1.0	0.5	0.3	0.1
Helena AGE											
Dara											
Hypo	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	4/16	0/16
Normal	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	4/16	0/16
Hyper	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	12/24	—	—
Elo											
Hypo	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	14/16	13/16
Normal	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	12/16	0/16
Hyper	24/24	24/24	24/24	24/24	24/24	24/24	24/24	24/24	20/24	—	—
Sebia AGE											
Dara											
Hypo	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	10/13	8/13	
Normal	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	5/13	0/13	
Hyper	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	12/13	—	—
Elo											
Hypo	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	11/13	9/13	
Normal	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	7/13	5/13	
Hyper	13/13	13/13	13/13	13/13	13/13	13/13	13/13	13/13	12/13	—	—
Sebia CZE											
Dara											
Hypo	23/23	23/23	23/23	23/23	23/23	23/23	23/23	23/23	21/21	15/21	
Normal	23/23	23/23	23/23	23/23	23/23	23/23	23/23	23/23	9/23	0/23	
Hyper	23/23	23/23	23/23	23/23	23/23	23/23	22/22	23/23	22/22	—	—
Elo											
Hypo	23/23	23/23	23/23	23/23	23/23	23/23	22/22	23/23	22/22	21/21	14/21
Normal	23/23	23/23	23/23	23/23	23/23	23/23	23/23	23/23	23/23	23/23	3/23
Hyper	23/23	22/22	22/22	22/22	22/22	21/21	22/22	22/22	22/22	—	—
Expected recovery, g/L	2.5	2.0	1.5	1.25	1.0	0.75	0.5	0.25	0.05		
Beta											
Helena AGE	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16	16/16		
Sebia AGE	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5		
Sebia CZE	6/6	6/6	6/6	6/6	6/6	6/6	6/6	6/6	6/6		

For each condition, the number of samples in which the M-protein is detected is followed by a slash and the total number of samples that was tested.

and the level of polyclonal background. Averaged over all methods, the intra-laboratory mean CV of Dara and Elo measured at 10 g/L in a hypo-gamma background was 2.6% (95% CI 1.7%–3.6%). The intra-laboratory variation gradually increased to a mean CV of 12.7% (95% CI 9.3%–16.1%) for M-proteins of 1 g/L measured in a hyper-gamma background (Figure 3A). A third parameter that made an impact on intra-laboratory CV was the method by which the laboratory performed the M-spike.

The mean CV of all duplicate measurements of M-proteins ranging from 1 to 10 g/L using the PD gating strategy was 4.1% (95% CI 3.8–4.5). The mean CV of 6.5% (95% CI 5.8–7.1) using TS was significantly higher ($p < 0.0001$) as shown in Figure 3B. The M-protein migration pattern had a small but significant effect on intra-laboratory CV. The mean CV for migrating Dara was 4.6% (95% CI 4.1%–5.0%), which was significantly lower than the mean

CV of 5.5% (95% CI 5.0%–6.0%) for Elo ($p = 0.005$). The intra-laboratory variation of M-proteins measured using the Sebia systems (mean CV 4.9%) was not significantly lower compared to the mean CV of 5.5% using the Helena systems ($p = 0.17$). Whether electrophoresis was performed using AGE or CZE did not affect intra-laboratory CV (Figure 3B). Below M-protein concentrations of 1 g/L, only a limited number of laboratories reported quantitative results, and imprecision within those laboratories increased to an average of 14% (95% CI 11.3–16.1) (data not shown).

Follow-up of individual M-proteins

As Dara, Elo and the beta-migrating monoclonal component each represent one M-protein, we were able to

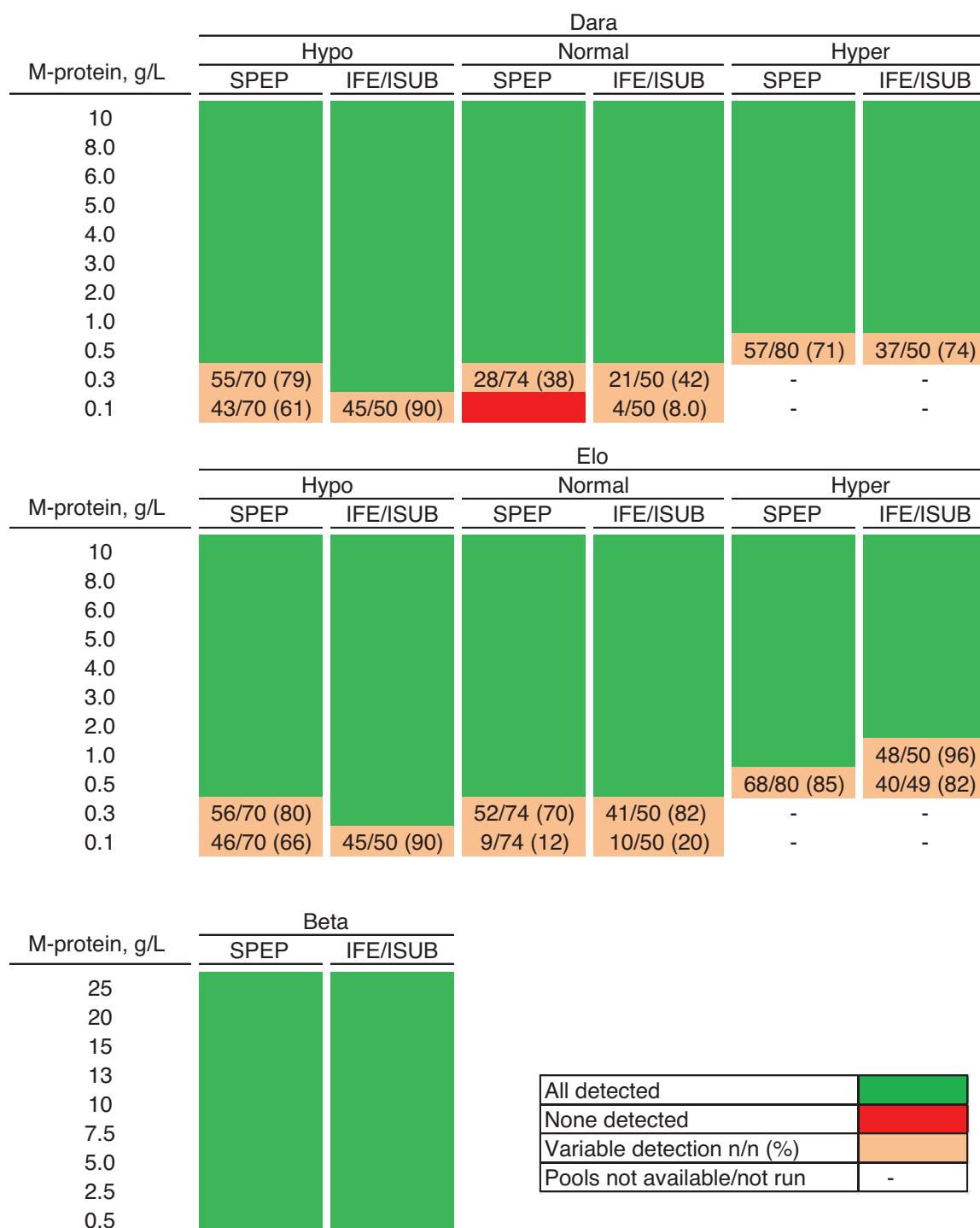


Figure 2: Limit of detection for SPEP and immunofixation/immunosubtraction.

Graphical illustration of LOD variation. The green cells represent conditions in which all samples were reported as abnormal by SPEP or IFE/ISUB. The peach-colored cells represent conditions in which a number of samples were reported as abnormal (ratio and percentage in brackets are shown in each cell). Cells in red would represent conditions in which the M-protein was not detected in all samples. A dash means pools were not prepared for the given concentration.

calculate how each participating laboratory performed in the hypothetical follow-up of each over various concentrations. The polyclonal background during this

hypothetical follow-up was kept constant as either hypo-gamma, normal or hyper-gamma. Because of the low intra-laboratory CV values reported earlier, M-protein

Table 3: Qualitative M-protein detection by immunofixation/immunosubtraction.

Expected recovery, g/L	10	8.0	6.0	5.0	4.0	3.0	2.0	1.0	0.5	0.3	0.1
Helena											
Mono											
Dara											
Hypo	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11
Normal	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	1/11	0/11
Hyper	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	—	—
Elo											
Hypo	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11
Normal	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	0/11
Hyper	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	11/11	—	—
Penta											
Dara											
Hypo	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Normal	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Hyper	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	—	—
Elo											
Hypo	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Normal	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Hyper	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	—	—
Sebia											
Mono											
Dara											
Hypo	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27
Normal	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	15/27	2/27
Hyper	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	14/27	—	—
Elo											
Hypo	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27
Normal	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	27/27	18/27	7/27
Hyper	27/27	27/27	27/27	27/27	27/27	27/27	27/27	25/27	17/26	—	—
Penta											
Dara											
Hypo	—	—	—	1/1	—	—	1/1	1/1	1/1	1/1	1/1
Normal	—	—	—	—	—	1/1	—	1/1	1/1	1/1	0/1
Hyper	—	—	—	—	—	1/1	—	1/1	1/1	—	—
Elo											
Hypo	—	—	—	—	1/1	—	1/1	—	1/1	1/1	1/1
Normal	—	—	—	—	—	1/1	1/1	1/1	1/1	1/1	1/1
Hyper	—	—	—	—	—	1/1	1/1	1/1	1/1	—	—
Sebia											
ISUB											
Dara											
Hypo	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	4/9
Normal	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	2/9	0/9
Hyper	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	—	—
Elo											
Hypo	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	4/9
Normal	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	0/9
Hyper	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	9/9	—	—
Expected recovery, g/L	25	20	15	13	10	7.5	5.0	2.5	0.5		
Beta											
Helena											
Mono	10/10	10/10	10/10	10/10	10/10	10/10	10/10	10/10	10/10		
Penta	—	—	—	—	—	—	—	—	—		
Sebia											
Mono	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15	15/15		
Penta	—	—	—	—	—	—	—	—	—		
Sebia											
SUB	6/6	6/6	6/6	6/6	6/6	6/6	6/6	6/6	6/6		

For each condition, the number of samples in which the M-protein is detected is followed by a slash and the total number of samples that was tested.

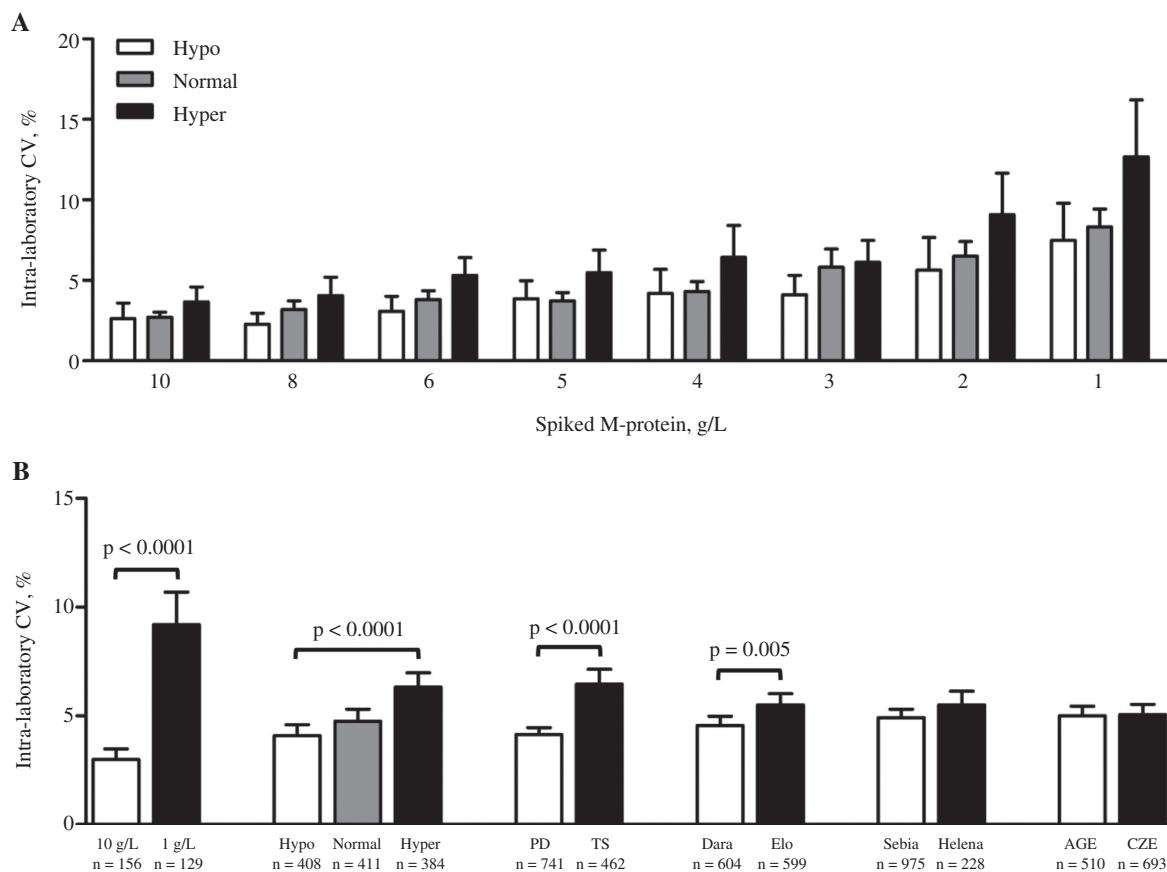


Figure 3: Intra-laboratory coefficient of variation (CV) of M-protein quantification.

(A) Intra-laboratory CV (%) of M-proteins ranging from 10 to 1 g/L, subdivided by the level of polyclonal immunoglobulin background. Plotted are mean values with 95% confidence interval. (B) Intra-laboratory CV (%) of different parameters that affect intra-laboratory precision. The number of duplicate measurements performed within each condition are shown under the bars. Plotted are mean values with 95% confidence intervals. Significance is indicated by p-values using two-tailed unpaired t-tests.

quantification within the same laboratory could successfully be applied to follow trend of M-protein levels (Figure 4). The different panels in Figure 4 demonstrate that the M-protein quantifications using the various methods are often inaccurate (i.e. more than 20% away from the dotted line that indicates perfect accuracy). As already pointed out in Part I, each method (AGE/CZE/PD/TS/Helena/Sebia) has its strengths and weaknesses in quantifying small M-proteins in a hypo-gamma (red), normal (green) and hyper-gamma (black) background. However, without exception to any of the methods and the level of polyclonal background, each individual laboratory correctly indicated in >99% of the samples when there was a small increase or decrease of the M-protein. Small differences in M-protein concentration of 0.25 g/L or even 0.125 g/L were correctly identified by all participating laboratories. Laboratories were equally effective to follow trend in both the gamma-migrating (Figure 4A) and beta-migrating (Figure 4B) M-proteins.

Discussion

Even though large parts of the M-protein analyses are fully automated, the actual detection and M-spike quantification remain subjective procedures [6]. As there are no strict guidelines with regard to M-protein analytical methodologies or reporting standards [11], many laboratory specialists struggle with questions as to when an M-spike is too small to permit quantification, detection and/or reporting. We performed an international, multi-center SPEP accuracy and M-protein isotyping study. Part I of this study focused on the LOQ of SPEP. In this second part of the study, we utilized the shared sample set in duplicate to objectively determine the limitations in detecting small M-proteins by defining method-specific LODs. Three institutions received five blinded sample sets of 71 samples each, and the remaining 13 institutes received two blinded sample sets of 62 samples each. Based on these replicate measurements, we assessed the

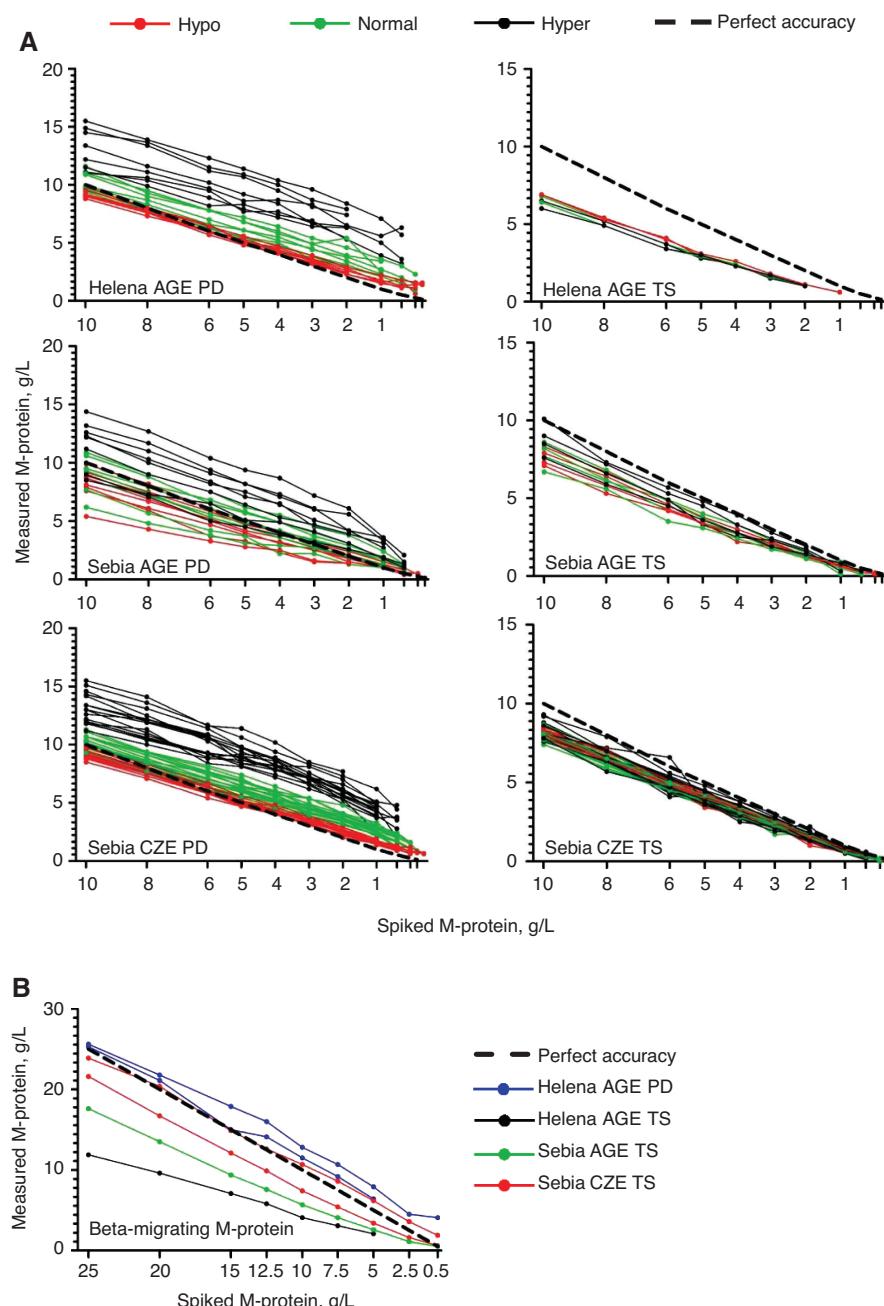


Figure 4: Follow-up of individual M-proteins.

(A) Each line represents the Dara/Elo concentrations reported by each individual laboratory. Red lines represent hypo-gamma samples, green lines represent normal-gamma samples and black lines represent hyper-gamma samples. The line of perfect accuracy is indicated by the dashed line. The method used is indicated in each panel. Data of the beta-migrating M-protein are plotted separately (B).

intra-laboratory variability of M-protein quantification in an effort to harmonize the field.

In this study, we confirm that IFE/ISUB is moderately more sensitive to detect intact M-proteins compared to SPEP. LOD differences between IFE and ISUB are small; however, IFE seems slightly more sensitive. We identified the intensity of the gamma-fraction background as

the most important factor that influences the LOD. We noted some variability between laboratories in the SPEP and IFE/ISUB LODs, even between laboratories using the same methods. The clinical consequence is that a given patient may or may not meet the response criteria of complete remission (requiring IFE negativity) depending on the laboratory in which the M-protein diagnostics was

performed [5]. LOD is important, as a negative IFE/ISUB result in patients with MM could lead to high sensitivity minimal residual disease (MRD) testing which includes a bone marrow biopsy. If one system was more likely to be negative due to a poorer LOD, it could lead to MRD overutilization. In this study, we found that all M-proteins ≥ 1 g/L are detected using SPEP and most M-proteins ≥ 0.3 g/L are detected using IFE/ISUB. This is in line with the observation that therapeutic Dara concentrations (reaching serum concentration up to 1 g/L) can be detected using electrophoretic isotyping in patients treated with this therapeutic monoclonal antibody [12–14].

Several groups previously reported the total variation (biological variation as well as analytical variation) of M-protein quantification in clinically stable patients [7, 15, 16]. In this study, we assessed the analytical variation as biologics were spiked into sera as surrogate M-proteins.

In Part I of this study, we concluded that the quantification of small M-proteins results in poor accuracy and high variation between laboratories [10]. For following patients in clinical practice, however, it may still provide relevant information to measure the M-protein as it changes during disease progression or in response to therapy. The present study demonstrated low intra-laboratory quantification variation of small M-proteins. The overall mean intra-laboratory CV was 5.0% (95% CI 4.7–5.4) when M-proteins from 1 to 10 g/L are quantified within the same laboratory using the same methodology. This is in line with the previously reported analytical variation of M-protein quantification which ranged from 2.1 to 6.2% [7, 15, 16]. Parameters that had a significant negative impact on intra-laboratory variation were: (1) low M-protein concentration, (2) hyper-gamma background, (3) TS used for M-spike and (4) central-gamma migration of the M-protein. But even a mean intra-laboratory CV of only 12.7% (95% CI 9.3%–16.1%) was found for M-proteins of 1 g/L measured in sera with a hyper-gamma background.

We show that because of these low intra-laboratory CV values, M-protein quantification within the same laboratory could successfully be applied to monitor small changes in M-protein levels in samples with a constant level of polyclonal background. Without exception to any of the methods or the nature of the sample, all participating laboratories correctly indicated in >99% of the samples when there was an increase or decrease of the M-protein. Small changes in M-protein concentration between samples, even down to 0.25 g/L, were correctly identified by those laboratories that quantified these small M-proteins. It is important to stress that the reported results only indicate a correct trend. In other words, the reported outcome provides exclusively information

on whether the M-protein concentration decreases or increases from one sample to another. Depending on the polyclonal background and the methodology used, the exact quantification of the changes in M-protein concentration is often inaccurate. It is also important to note that in case the polyclonal background changes in individual patients during follow-up, it becomes more challenging to monitor small changes in M-protein concentration because the same M-protein is quantified differently in a different polyclonal background as was previously shown by Schild et al. [17]. As a consequence, a falsely increased M-protein concentration might therefore be caused, for example, by an acute infection that induces a transient polyclonal hypergammaglobulinemia or in patients receiving IVIG therapy. Patients that start treatment also often experience changes in their polyclonal immunoglobulin background. However, most patients who undergo active treatment are monitored frequently. Therefore, often a gradual decrease of polyclonal background is observed over a couple of monitoring samples which limits the negative impact on the M-protein trend. To help interpret follow-up of M-protein quantification, three independent groups calculated M-protein reference change values as a tool to detect significant changes in a serial laboratory result to identify a true change in the pathological state of a patient with an MG. The reference change value is based on a combination of analytical variation and biological variation. The M-protein reference change values reported by these groups range from 25% to 38% [7, 15, 16].

Overall, our data are in line with the current recommendation to monitor MG patients in the same laboratory using the same method to improve the reproducibility and comparability of serial measurements [11, 18, 19]. Our data suggest that quantification performed within a single laboratory can give reliable follow-up data of M-proteins down to concentrations of 1 g/L, with the acceptance that the quantification of M-proteins at this level is analytically inaccurate. Our data stress the suitability of the IMWG response criteria in which the response is calculated as % reduction of the M-protein [8]. Because even when the measured M-protein concentration may be inaccurate, the calculated % reduction is consistent within laboratories provided there are no large fluctuations in the polyclonal background. Our studies indicate that comparing M-protein measurements between laboratories is not recommended. Even among laboratories using the same method to measure the M-protein, there was considerable variation.

It is important to note that it is challenging to conclude which method has the lowest LOD. Cultural

practice (human factors) may be a source of result variation between laboratories. Two examples include (a) the assay reader judges the extent of the M-protein and (b) laboratory specialists may differ in their practices of reporting small abnormal results or not, with some being more conservative than others. No conclusions could be drawn on the LOD of the beta-migrating M-protein, as the M-protein was not diluted far enough and LOD was never reached with the tested concentrations. It is also important to note that the beta-migrating M-protein in this study was visible as a distinct peak; it is generally appreciated that the LOD for SPEP of such an M-protein is lower compared to that of an M-protein that migrates in the exact center of the beta region. With regard to specificity, the participants were not asked to confirm a positive pentavalent signal using monovalent IFE antisera to verify that indeed the reported band coincides with the isotype of the M-protein that was spiked into the sample. An additional limitation of this study is that there may have been some bias in the sense that likely the experiments were run in a short period of time, possibly with a single operator, which may result in better performances compared to the ones observed in routine laboratory practice. Finally, intra-laboratory variation was calculated on duplicate samples. This, however, was compensated by the large number of duplicate analyses performed in this multi-center study.

Overall, our data suggest that SPEP quantification of small M-proteins can provide results with poor accuracy; however, it provides suitable information to monitor individual patients in case follow-up is performed within the same laboratory.

Acknowledgments: The authors would like to dedicate this work to Jillian R. Tate who died on December 4, 2018. Ms. Tate was a scientist at the Department of Chemical Pathology, Pathology Queensland, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia, and chair of the working group on standardization of reporting of small monoclonal proteins by electrophoresis with the IFCC at the time of study design and execution. Although she did not get to see the final manuscript of this work, the study would not have been accomplished without her input on study design, leadership with all the participants and invaluable mentorship. The authors would like to thank the RCPA Quality Assurance Program (St. Leonards, NSW, Australia), Dr. Louise Wienholt for shipping expenses of materials to all laboratories in Australia and Dr. Sara Altinier from UOC Medical Laboratory at the University of Padova (Padova, Italy) for coordinating the distribution and shipping of samples in Italy.

In addition, the authors would like to thank Dr. David Keren, MD, Professor of Pathology from the University of Michigan Ann Arbor, MI, USA, for providing images of capillary electrophoresis and immunosubtraction experiments, and Dr. David L. Murray, MD, PhD (Mayo Clinic, Rochester, MN, USA) and Dr. Michael Linden, MD, PhD (University of Minnesota, Minneapolis, MN, USA) for the critical reading of the manuscript and insightful comments. Technologists who performed SPEP, IFE and ISUB runs and interpreted gels are thoroughly thanked by the authors here: Corrie de Kat Angelino (Radboud University Medical Center), Yoke Leong (Department of Chemical Pathology, The Royal Melbourne Hospital), Anfernee Tseng (Pathology Queensland, Royal Brisbane and Women's Hospital), Megan Rae (Royal Prince Alfred Hospital), Karla Lemmert and Christine Burns (NSW Health Pathology), Margherita Berardi (General Laboratory, Careggi University Hospital, Florence), Maddalena Marini and Lavinia Nicolini (Clinical Chemistry Laboratory, University of Verona), Riccardo Albertini and Tiziana Bosoni (Servizio Analisi Chimico Cliniche, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy), Tatjana Tverskaja and Galina Trofimova (Clinical Chemistry Laboratory, North Estonia Medical Centre), Stacey Avondet and Gerald Ockey (ARUP Laboratories), Donald Giacherio and Theresa Nurmi (Clinical Immunology Laboratory, The University of Michigan) and Amy Marriott (Helena Biosciences Europe).

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission. Marie Therese Melki and Stephen Bell did not participate in the study design or data analysis.

Research funding: This project was funded by Sebia, Inc. and the Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA (funder Id: <http://dx.doi.org/10.13039/100000871>). Maria Willrich has received a \$6000 grant in research support from Sebia, Inc. to perform this study.

Employment or leadership: Marie Therese Melki is an employee of Sebia Inc., manufacturer of the products evaluated in this study. Stephen Bell is an employee of Helena Biosciences Europe, manufacturer of the products evaluated in this study.

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.

Conflict of interest: None of the other authors have conflicts of interest to declare.

References

1. Glavey SV, Leung N. Monoclonal gammopathy: the good, the bad and the ugly. *Blood Rev* 2016;30:223–31.
2. Rollig C, Knop S, Bornhauser M. Multiple myeloma. *Lancet* 2015;385:2197–208.
3. Willrich MA, Murray DL, Kyle RA. Laboratory testing for monoclonal gammopathies: Focus on monoclonal gammopathy of undetermined significance and smoldering multiple myeloma. *Clin Biochem* 2018;51:38–47.
4. Dimopoulos M, Kyle R, Fermand JP, Rajkumar SV, San Miguel J, Chanan-Khan A, et al. Consensus recommendations for standard investigative workup: report of the International Myeloma Workshop Consensus Panel 3. *Blood* 2011;117:4701–5.
5. Rajkumar SV, Harousseau JL, Durie B, Anderson KC, Dimopoulos M, Kyle R, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. *Blood* 2011;117:4691–5.
6. Keren DF, Schroeder L. Challenges of measuring monoclonal proteins in serum. *Clin Chem Lab Med* 2016;54:947–61.
7. 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.
8. Kumar S, Paiva B, Anderson KC, Durie B, Landgren O, Moreau P, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016;17:e328–46.
9. Murray DL, Seningen JL, Dispenzieri A, Snyder MR, Kyle RA, Rajkumar SV, et al. Laboratory persistence and clinical progression of small monoclonal abnormalities. *Am J Clin Pathol* 2012;138:609–13.
10. Turner KA, Frinack JL, Ettore MW, Tate JR, Graziani MS, Jacobs JF, et al. An international multi-center serum protein electrophoresis accuracy and M-protein isototyping study. Part I: factors impacting limit of quantitation of serum protein electrophoresis. *Clin Chem Lab Med* 2019.
11. Tate J, Caldwell G, Daly J, Gillis D, Jenkins M, Jovanovich S, et al. Recommendations for standardized reporting of protein electrophoresis in Australia and New Zealand. *Ann Clin Biochem* 2012;49:242–56.
12. van de Donk NW, Otten HG, El Haddad O, Axel A, Sasser AK, Croockewit S, et al. Interference of daratumumab in monitoring multiple myeloma patients using serum immunofixation electrophoresis can be abrogated using the daratumumab IFE reflex assay (DIRA). *Clin Chem Lab Med* 2016;54:1105–9.
13. McCudden C, Axel AE, Slaets D, Dejoe T, Clemens PL, Frans S, et al. Monitoring multiple myeloma patients treated with daratumumab: teasing out monoclonal antibody interference. *Clin Chem Lab Med* 2016;54:1095–104.
14. Xu XS, Yan X, Puchalski T, Lonial S, Lokhorst HM, Voorhees PM, et al. Clinical implications of complex pharmacokinetics for daratumumab dose regimen in patients with relapsed/refractory multiple myeloma. *Clin Pharmacol Ther* 2017;101:721–4.
15. Evliyaoglu O, van Helden J, Jaruszewski S, Imoh M, Weiskirchen R. Reference change values of M-protein, free light chain and immunoglobulins in monoclonal gammopathy. *Clin Biochem* 2019;74:42–6.
16. Salamatmanesh M, McCudden CR, McCurdy A, Booth RA. Monoclonal protein reference change value as determined by gel-based serum protein electrophoresis. *Clin Biochem* 2018;51:61–5.
17. Schild C, Wermuth B, Trapp-Chiappini D, Egger F, Nuoffer JM. Reliability of M protein quantification: comparison of two peak integration methods on Capillaries 2. *Clin Chem Lab Med* 2008;46:876–7.
18. Tate J, Panteghini M. Standardisation – the theory and the practice. *Clin Biochem Rev* 2007;28:127–30.
19. Tate JR, Keren DF, Mollee P. A global call to arms for clinical laboratories – harmonised quantification and reporting of monoclonal proteins. *Clin Biochem* 2018;51:4–9.