

Katherine A. Turner, Jody L. Frinack, Michael W. Ettore, Jillian R. Tate, Maria Stella Graziani, Joannes F.M. Jacobs, Ronald A. Booth, Christopher R. McCudden, David F. Keren, Julio C. Delgado, Galina Zemtsovskaja, Robert O. Fullinfaw, Anna Caldini, Theo de Malmanche, Katina Katakouzinou, 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 I: factors impacting limit of quantitation of serum protein electrophoresis

<https://doi.org/10.1515/cclm-2019-1104>

Received October 25, 2019; accepted December 2, 2019; previously published online January 15, 2020

Abstract

Background: Serum protein electrophoresis (SPEP) is used to quantify the serum monoclonal component or M-protein, for diagnosis and monitoring of monoclonal gammopathies. Significant imprecision and inaccuracy pose challenges in reporting small M-proteins. Using therapeutic monoclonal antibody-spiked sera and a pooled beta-migrating M-protein, we aimed to assess SPEP limitations and variability across 16 laboratories in three continents.

Methods: Sera with normal, hypo- or hypergammaglobulinemia were spiked with daratumumab, Dara (cathodal migrating), or elotuzumab, Elo (central-gamma migrating), with concentrations from 0.125 to 10 g/L (n=62) along with a beta-migrating sample (n=9). Provided with total protein (reverse biuret, Siemens), laboratories blindly analyzed

samples according to their SPEP and immunofixation (IFE) or immunosubtraction (ISUB) standard operating procedures. Sixteen laboratories reported the perpendicular drop (PD) method of gating the M-protein, while 10 used tangent skimming (TS). A mean percent recovery range of 80%–120% was set as acceptable. The inter-laboratory %CV was calculated.

Results: Gamma globulin background, migration pattern and concentration all affect the precision and accuracy of quantifying M-proteins by SPEP. As the background increases, imprecision increases and accuracy decreases leading to overestimation of M-protein quantitation especially evident in hypergamma samples, and more prominent with PD. Cathodal migrating M-proteins were associated with less imprecision and higher accuracy compared to central-gamma migrating M-proteins, which is attributed to the increased gamma background contribution in M-proteins migrating in the middle of the gamma fraction. There is greater imprecision and loss of accuracy at lower M-protein concentrations.

*Corresponding author: 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, Tel.: +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

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

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

Joannes F.M. Jacobs: Department of Laboratory Medicine, Radboud University Medical Center, Laboratory Medical Immunology, Nijmegen, the Netherlands

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

David F. Keren: 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 Katakouzinou: Immunopathology Department, Royal Prince Alfred Hospital, Missenden Rd, 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, Gateshead, UK

Open Access. ©2020 Katherine A. Turner et al., published by De Gruyter.  This work is licensed under the Creative Commons Attribution 4.0 International License.

Conclusions: This study suggests that quantifying exceedingly low concentrations of M-proteins, although possible, may not yield adequate accuracy and precision between laboratories.

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

Introduction

Founded on the secretion of a monoclonal immunoglobulin (M-protein), plasma cell proliferative disorders are classified as monoclonal gammopathies and include multiple myeloma (MM), Waldenström's macroglobulinemia (WM), amyloidosis (AL), light chain deposition disease (LCDD), POEMS syndrome, and premalignant diseases such as monoclonal gammopathy of undetermined significance (MGUS) and smoldering multiple myeloma (SMM). Due to the wide range of disease presentations, identification of the M-protein by serum protein electrophoresis (SPEP) may be the first clue to the diagnosis of a monoclonal gammopathy followed by subsequent isotyping via immunofixation (IFE) or immunosubtraction (ISUB). SPEP and IFE/ISUB are indispensable clinical assays for the diagnosis, risk stratification and therapeutic monitoring of monoclonal gammopathies [1–3].

In normal serum, SPEP yields a broad gamma fraction with Gaussian distribution due to the millions of plasma cell clones that secrete immunoglobulins into the serum. In monoclonal gammopathies, M-proteins are visualized as a restricted area of migration in the electrophoretic pattern which can be gated and quantified [4]. M-protein migration is quite heterogeneous; in a cohort of 1027 MM patients, the M-protein migration was distributed in the following order: 54% of the M-proteins were in the gamma fraction, 12% in the beta-gamma bridge, 13% in the beta fraction and 1% in the alpha-2 fraction, adding up to 80%, while 2% of cases were biclonal, 8% hypogammaglobulinemic and 11% were normal when reviewing serum and urine protein electrophoresis [5]. A panel of assays combining SPEP with IFE/ISUB, serum free light chains (FLC), and urine electrophoresis and IFE/ISUB has the highest degree of clinical sensitivity for the diagnosis of these disorders [6].

After diagnosis of a monoclonal gammopathy, patients will be monitored for the rest of their lives, with either imaging studies, bone marrow biopsies and most likely repeat measurements of the serum, or urine test which identified the disease. Response to therapy over

time is associated with reduction of the M-protein size. The low analytical accuracy of SPEP is well known when measuring low concentrations of M-proteins (<10 g/L). Each laboratory decides whether such low-concentration M-proteins can be reported quantitatively or qualitatively. SPEP testing is then often followed by IFE/ISUB [2, 7–9]. Patients may receive therapy from different institutions over the course of their disease, and the patient's serum samples may be sent to different laboratories with different assay methods and gating practices. Inaccurate quantitation of small M-proteins may impact trending or the classification of response criteria.

SPEP can be performed using agarose gel (AGE) or capillary electrophoresis (CZE). Quantitation of the M-protein is achieved through integration of the M-protein peak in the electropherogram. Two methods are currently utilized: perpendicular drop (PD) and tangent skimming (TS) (Figure 1A). Quantitation by PD is mediated by the placement of vertical gates at the M-protein's anodal and cathodal limits, integrating the total area under the curve. Because the PD measurement includes the polyclonal base, it has been recommended that no measurement be made when the spike is <1/4 of the base [10]. Recognizing this limitation in the PD method, laboratories adopted strategies to limit the inaccuracy of the measurements by only performing measurements when the spike was at least one-quarter of the total gamma fraction background. TS draws a line connecting the M-protein's anodal and cathodal deflection points, where the M-protein peak meets the polyclonal background, and integrates only the triangular area above the line, excluding the underlying polyclonal immunoglobulins. Schild et al. reported accuracy by the percent recovery (measured by a 25% variance from the known absolute value) with PD as 15 g/L, and 1.5 g/L with TS using 50 cases and serial dilutions of sera [11]. A 2018 survey of clinical laboratories in 38 countries showed 521 out of 674 of them use PD (77%), while 61 out of 674 (9%) have adopted TS to routinely quantitate monoclonal proteins. Other respondents used other non-gel methods to quantitate the M-protein such as immunoglobulin quantitation or serum FLC [12].

Although SPEP and IFE/ISUB are universally utilized clinical assays, there is a lack of standardization with considerable variability between laboratories in their reporting practices. In fact, there is limited guidance and recommendations for the reporting of monoclonal gammopathies by SPEP and IFE/ISUB, specifically, the limit of quantitation (LOQ) for these assays, when an M-protein is too small to be accurately quantified. Both SPEP and IFE/ISUB have been available in the clinical laboratories for decades. In more recent times, the regulatory environment

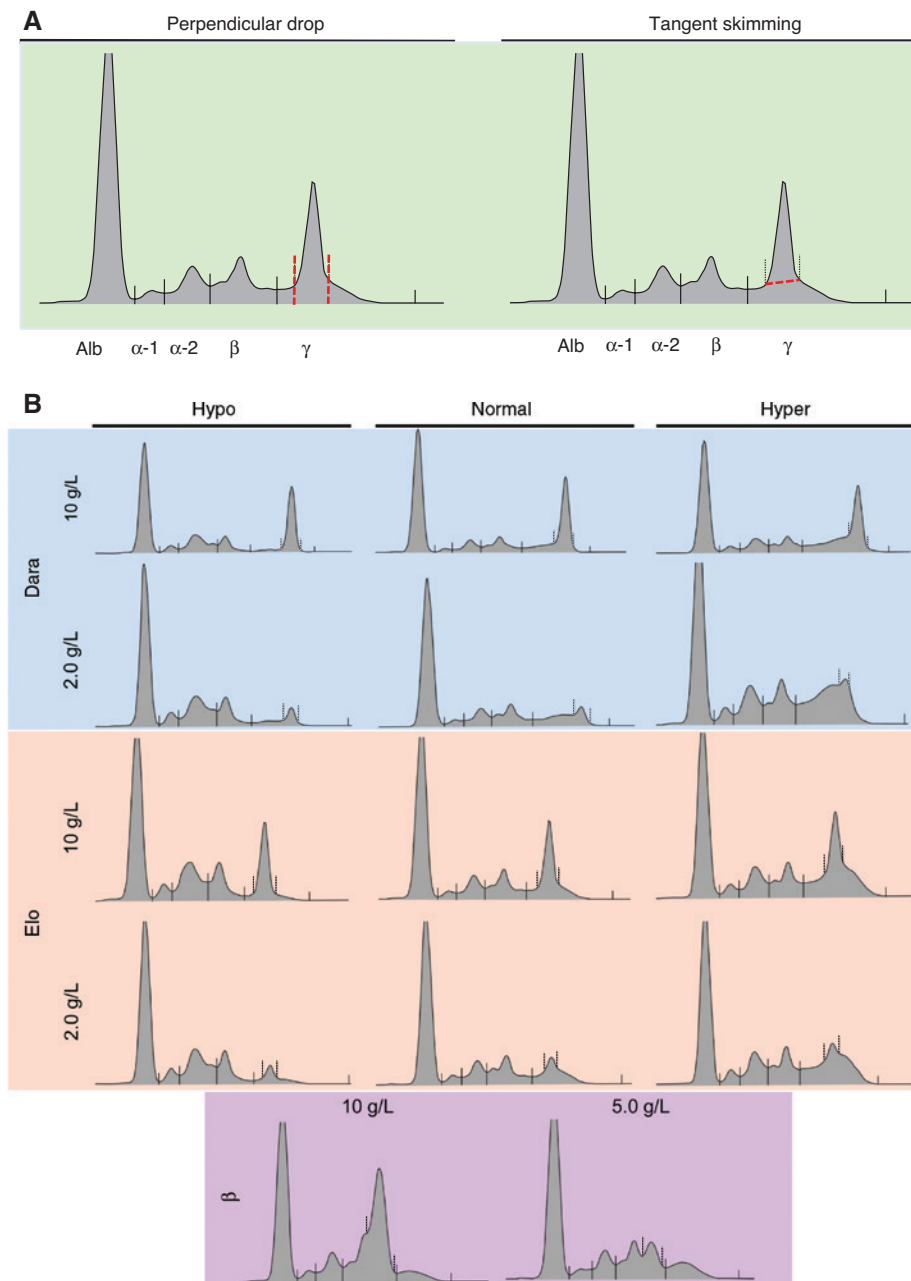


Figure 1: Representative electropherograms of serum protein electrophoresis.

(A) Perpendicular drop and tangent skimming methods to quantify a monoclonal protein migrating in the gamma fraction are shown.

(B) Migration pattern and different gamma fraction backgrounds may impact the ability to accurately quantitate an M-protein.

Concentrations of 10 g/L and 2.0 g/L are shown to illustrate the impact of the polyclonal background in quantitation. Dara (daratumumab) illustrates an M-protein which migrates at the end of the gamma fraction, closer to the cathode, while Elo (elotuzumab) illustrates an M-protein which migrates in the center of the gamma fraction. Approximately one-third of the monoclonal IgAs migrate in the beta fraction. Concentrations of 10 g/L and 5 g/L are shown to illustrate the impact of co-migrating proteins in the background impacting quantitation in the beta fraction. Electropherograms shown were obtained using AGE on a Helena system.

for test validations has become more stringent. Formal studies to document the LOQ and limit of detection (LOD) of these assays may have not been available. Additionally, while the International Myeloma Working Group (IMWG) has made specific recommendations on detection,

monitoring and follow-up of monoclonal proteins, it has not made recommendations with regard to analytical methodologies [13, 14].

The heterogeneity of the M-protein and the amount of background in the gamma fraction of SPEP are additional

contributing factors to the lack of a universal LOQ. With less background, such as the one observed in hypogammaglobulinemic samples, the M-protein will be more easily identifiable. In contrast, if there is high background as seen in hypergammaglobulinemia, the M-protein may be neither visible nor quantifiable even if present in relatively large concentrations (Figure 1B). Here, we present a study which was undertaken to more clearly define variability within the quantification of M-proteins. This study compares SPEP and qualitative isotyping – by each of IFE/ISUB – and also compares different laboratories, each using their respective methodologies. This includes M-proteins of differing concentrations, different gamma fraction backgrounds and different migration patterns. Additionally, a LOQ for M-protein analysis by SPEP and a LOD for both SPEP and IFE/ISUB for specific methodologies were characterized in an effort to streamline reporting guidelines between institutions and bring assay harmonization among laboratories, an initiative supported by the International Federation of Clinical Chemistry and driven by the sub-group for harmonization of reporting of protein electrophoresis and quantitation of small proteins.

Materials and methods

Materials and samples

Large-volume sample pools with M-proteins in varying gamma fraction backgrounds were prepared in enough quantity to be shared with 16 different institutions (Supplementary Table 1). Briefly, specimens containing gamma-migrating M-proteins were prepared by spiking known amounts of the therapeutic monoclonal antibodies (mAbs), daratumumab (Dara) (Darzalex, Janssen Pharmaceutical) or elotuzumab (Elo) (Empliciti, Bristol-Myers Squibb), into pooled sera. Pools were prepared by an experienced technologist following standard preparation procedures, and after a pilot study (data not shown), the spiked concentrations were considered the true value of the M-protein for the purpose of this study. Similar studies with Dara and Elo have been performed in serum matrix in the past, with acceptable recoveries [15]. Both of the IgGκ isotypes, Dara and Elo, represent M-proteins in the gamma fraction with differing migration patterns, as Dara migrates at the top of the gamma fraction (cathodal), whereas Elo migrates in the center of the gamma fraction on gel electrophoresis and at the anodal end of the gamma region by CZE. Importantly, the concentrations of Dara and Elo used in this study were chosen to represent small monoclonal proteins seen consistently in the clinical laboratory.

Dara and Elo pharmaceutical preparations were purchased from the Mayo Clinic pharmacy. The sterile, sealed vials of Dara were provided in a 20-mg/mL solution. The solution was diluted 1:2 with hypo, normal or hypergamma sera to yield a 10-mg/mL (10 g/L) Dara-spiked serum preparation. Elo was provided as a lyophilized powder in a sterile, sealed vial. The powder was reconstituted with 2 mL of

reagent-grade water for a final concentration of 150 mg/mL. The solution was diluted 1:15 with hypo-, normal or hypergamma sera to yield a 10-mg/mL (10 g/L) Elo-spiked serum preparation. Importantly, each concentration tested, ranging from 0.125 g/L to 10 g/L, was individually prepared from the 20-mg/mL Dara solution or 150-mg/mL Elo solution in each of the pools. No serial dilutions were performed throughout the preparation of the standard. The US pharmacopeia and the International Conference on Harmonization (ICH) guidelines followed throughout the world by pharmaceutical industries recommend thresholds for reporting degradation impurity (0.1%) and identifying it (1%) [16].

Three different serum pools were prepared, with three gamma fraction intensities: normal, hypo- and hypergamma backgrounds. The normal gamma serum was purchased (EMD Millipore) and the hypo- and hypergamma sera were pools prepared from de-identified residual patient sera with similar background (Institutional Review Board approval IRB 17-001526). All residual sera used for the study had been previously tested with SPEP and IFE and were found to be negative for the presence of any endogenous M-proteins. The final preparation of the normal gamma pool had a total protein measurement of 68 g/L (albumin 36 g/L, alpha-1 3 g/L, alpha-2 8 g/L, beta 10 g/L and gamma fraction 11 g/L). Hypogammaglobulinemic residual sera with total gamma <5 g/L were pooled to yield a hypo pool with a total protein of 57 g/L (albumin 31 g/L, alpha-1 3 g/L, alpha-2 11 g/L, beta 8 g/L and gamma fraction 4 g/L). Finally, hypergammaglobulinemic residual sera with total gamma >17 g/L were pooled to yield a hyper pool with a total protein of 78 g/L (albumin 35 g/L, alpha-1 3 g/L, alpha-2 10 g/L, beta 10 g/L and gamma fraction 20 g/L). Samples containing the beta-migrating M-protein consisted of serum from a single patient with an M-protein of IgAλ isotype. The original concentration of the beta-migrating IgA M-protein measured by AGE using a Helena system with PD was 59 g/L with a neat total IgA concentration of 54 g/L as determined by the Mayo Clinic using nephelometry (Siemens BNII, Siemens Healthineers). Dilutions of the neat specimen from 25 g/L to 0.5 g/L were prepared with normal serum using the SPEP M-spike concentration only as a baseline concentration.

SPEP and IFE/ISUB testing

Sample aliquots were blinded and sent out to the testing laboratories frozen on dry ice, and laboratories were advised to keep samples frozen (–20 °C or colder) until testing. Laboratories were instructed to run the samples along-side their patient samples according to their institution's standard operating procedure for SPEP and IFE/ISUB. Laboratories were provided with total IgG (nephelometry, Siemens BNII) and total protein (reverse biuret, Siemens Advia 1200) on all samples to aid in determining the optimal dilutions for testing and to reduce the variability of this portion of the testing.

Academic institutions received two sets of samples and were asked to run at minimum SPEP on all the samples and a replicate measurement on a different day if possible. Additionally, academic laboratories with sufficient sample volume were asked to run samples on a secondary SPEP method (if available) and perform IFE/ISUB to isotype the M-proteins. Industry participants received five sets of samples and were asked to run all the samples at a minimum of once per day for 5 days or twice per day, if sufficient volume for both SPEP and IFE/ISUB. Sharepoint, a web-based, collaborative platform, was

set up for communication with all study participants, with access to an excel template to enter the results along with SPEP and IFE/ISUB methodologies and manufacturer utilized.

Data analysis

All results were collected and analyzed by the Mayo Clinic. Accuracy of quantitation was assessed by percent recovery, and a range of 80%–120% was empirically set as acceptable. Between-lab imprecision was defined as a target of a CV of 20% or less. LOQ was determined by the lowest M-protein concentration within the acceptable recovery range and CV $\leq 20\%$. LOD for SPEP and IFE/ISUB were defined conservatively as the lowest concentration of M-protein in which all samples analyzed had an M-protein reported qualitatively.

Results

Method distribution

With representatives in Australia, New Zealand, Italy, UK, Netherlands, Estonia, US and Canada, 16 laboratories reported their SPEP and IFE/ISUB methodology utilized to analyze the blinded sample set of over 3000 unique serum aliquots. Four different SPEP systems were represented in this study: an AGE platform from Helena Laboratories (Beaumont, TX, USA), a second AGE platform commercialized by Helena Biosciences Europe (UK), and an AGE and a CZE platform both from Sebia (Sebia Inc., France) (Table 1). Within each method, both PD and TS gating techniques were applied. For laboratories that submitted more than one SPEP method, their primary method of analysis, used in everyday clinical practice, was noted and the additional analysis was defined as a supplemental method. The exception to that rule was Sebia Inc., which has both methods listed as primary, adding to 17 primary methods reported.

The most represented method was CZE from Sebia using PD ($n=8$), followed by Sebia CZE with TS ($n=7$), Sebia AGE and Helena AGE with PD ($n=4$), Sebia AGE with TS, ($n=2$), and lastly Helena AGE with TS ($n=1$). Only two of the seven laboratories used Sebia CZE with TS as their primary method. In addition, no participant laboratory used Helena AGE with TS routinely. For isotyping the M-protein, the majority of laboratories used IFE and applied monovalent antisera on a Sebia platform ($n=9$), followed by ISUB by Sebia ($n=3$), monovalent antisera on a Helena platform ($n=2$), and pentavalent antisera applied to either Helena or Sebia platforms ($n=1$ each).

Table 1: Institutions' methodology distribution for serum protein electrophoresis and immunofixation/immunosubtraction.

| | North America | Europe | Australia |
|--------------|---------------|--------|-----------|
| SPEP | | | |
| Helena AGE | | | |
| PD | | | |
| Primary | 1 | 1 | 1 |
| Supplemental | – | – | 1 |
| TS | | | |
| Primary | – | – | – |
| Supplemental | – | – | 1 |
| Sebia AGE | | | |
| PD | | | |
| Primary | 1 | 1 | 1 |
| Supplemental | – | – | 1 |
| TS | | | |
| Primary | – | – | 1 |
| Supplemental | – | 1 | – |
| Sebia CZE | | | |
| PD | | | |
| Primary | 1 | 5 | 2 |
| Supplemental | – | – | – |
| TS | | | |
| Primary | 1 | 1 | – |
| Supplemental | – | 4 | 1 |
| IFE/ISUB | | | |
| Helena | | | |
| Mono | – | 1 | 1 |
| Penta | – | – | 1 |
| Sebia | | | |
| Mono | 2 | 5 | 2 |
| Penta | – | 1 | – |
| ISUB | 1 | 2 | – |

SPEP and IFE/ISUB methods have been used in 16 unique institutions across three continents. Primary method and alternative method results submitted are noted. AGE, agarose gel electrophoresis; CZE, capillary electrophoresis; IFE, immunofixation; ISUB, immunosubtraction; Mono, monovalent antisera; PD, perpendicular drop gating; Penta, pentavalent antisera; TS, tangent skimming gating. Three Australian and four European institutions used supplemental methodologies. Alternate AGE (Helena, 2 [1 PD and 1 TS gating] and Sebia, 1 [PD gating]) and Sebia CZE was used by Australian testing centers. In Europe, supplemental testing included one institution using Sebia AGE with TS gating and four using Sebia CZE with TS gating.

Influencing factors of method performance

Each method and their respective gating techniques showed remarkable differences in quantitation. There is an overall loss of accuracy and within-lab and inter-lab precision as the concentration of the M-protein decreases regardless of the method (Figure 2; for a detailed statistical summary of each method's performance, see Supplementary Tables 2–7). However, the extent of the deviation

from the true M-protein concentration, reproducibility and the factors that influence quantitation accuracy differ depending on the M-protein gating method, intensity of polyclonal background and migration pattern of the M-protein.

M-protein gating method

Gating low-concentration M-proteins by PD leads to over-estimation of the M-protein, whereas TS leads to underestimation. This is evidenced in Figure 2, by comparing the expected concentrations of Dara and Elo (x axis) with the measured values expressed as relative

recoveries on the y axis. To illustrate, samples spiked with 10 or 2 g/L Dara in a normal gamma fraction background and analyzed by Sebia CZE with PD showed a mean recovery (%CV) of 99% (3.2%) and 171% (7.6%), respectively, whereas analysis by Sebia CZE with TS had mean recovery (%CV) of 85% (4.0%) and 76% (7.4%), respectively (Figure 2A, center panels and Supplementary Tables 6 and 7). The event of over-estimation and under-estimation was also seen with Elo-spiked samples (Figure 2B) and in the beta-migrating sample (Figure 2C), where the recovery (%CV) of the 25-g/L beta-migrating IgA M-protein sample as measured by Helena AGE PD was 102% (1.7%), whereas the recovery (%CV) of the 5.0-g/L sample was 151% (11.6%).

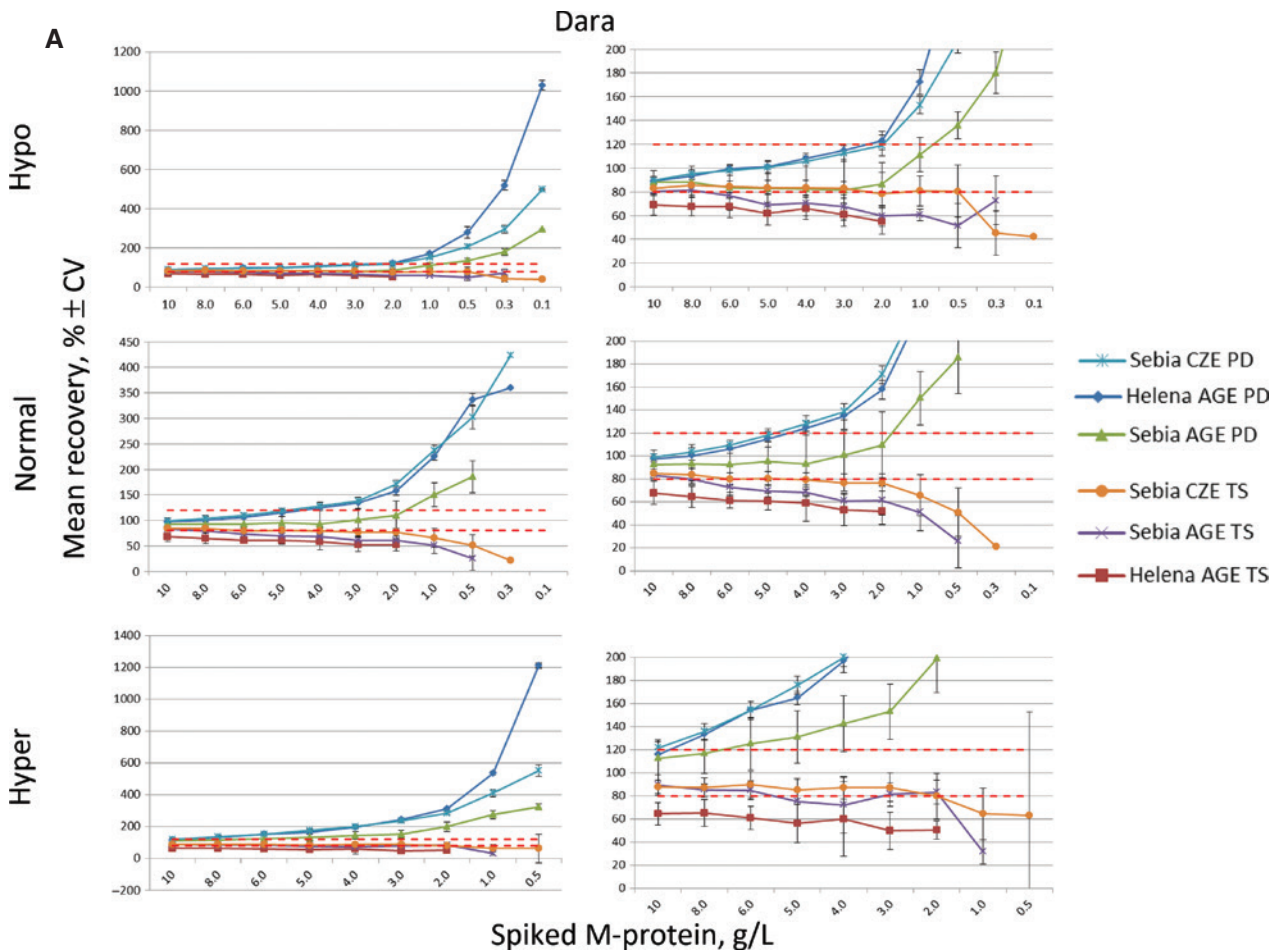


Figure 2: Spiked M-protein recovery by the serum protein electrophoresis method.

Each line represents a different SPEP system and gating technique. M-protein expected concentration is on the x axis, plotted against the experimentally obtained value expressed as the mean percent recovery (error bars represent %CV of each replicate measurement). Red dashed lines represent the acceptable recovery range of 80%–120%. (A) Dara-spiked samples in three gamma fraction backgrounds on the left panes. On the right panes, zoomed in on recoveries between 0% and 200%. (B) Elo-spiked samples in three gamma fraction backgrounds on the left pane. On the right panes, zoomed in on recoveries between 0% and 200%. (C) Beta-migrating M-protein sample on the left pane. On the right pane, zoomed in on recoveries between 0% and 200%. PD, perpendicular drop; TS, tangent skimming; AGE, agarose gel electrophoresis; CZE, capillary zone electrophoresis.

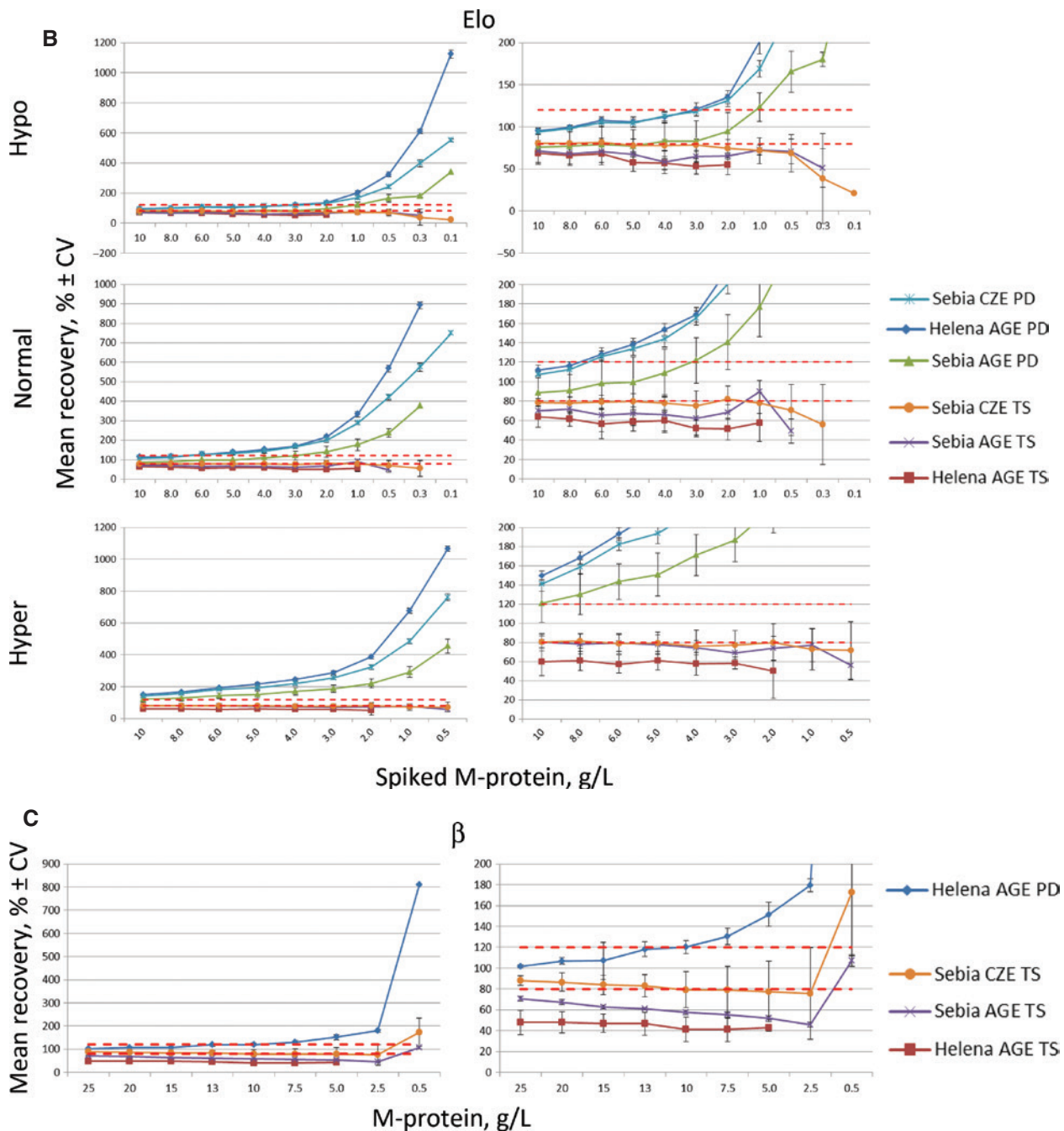


Figure 2: (Continued)

The recoveries (%CV) of the same sample analyzed by Sebia CZE with TS were 88% (4.5%) and 78% (29%), respectively (Figure 2C). In contrast to the pharmaceutical preparations obtained for Dara and Elo which had a manufacturer-provided concentration, the beta-migrating M-protein concentration was estimated using the M-spike value obtained by the leading laboratory, which used AGE PD. For this specific sample, if the IgA concentration derived from the nephelometric method

is utilized to determine the accuracy of quantitation instead of the M-spike value, Helena AGE with PD now only remains accurate until 14 g/L and Sebia CZE with TS has increased accuracy down to 2.3 g/L (Supplementary Figure 1).

Although the different gating techniques lead to opposite deviations from the acceptable range, as the concentration of the M-protein decreases, the extent of the loss of accuracy is influenced by other factors that

compound the loss of accuracy related to low protein concentration.

The intensity of the polyclonal background in the gamma fraction

This is especially relevant in the use of PD. As the concentration of the gamma fraction increases, there is a substantial decrease in accuracy and precision. This is evident by the difference in mean recoveries (%CV) between the hypo-, normal and hypergamma samples (Figure 2A,B). In samples spiked with 10 g/L of Dara and analyzed by Helena AGE with PD, the recoveries (%CV) were as follows: hypogamma 89% (2.6%), normal 97% (2.5%) and hypergamma 116% (5.1%) (Supplementary Table 2). Similar recoveries were observed when Sebia CZE with PD was analyzed at the same concentration:

90% (2.8%), 99% (3.2%) and 122% (6.4%), respectively (Supplementary Table 6). Both accuracy and precision worsen as the relative proportion of the gamma background increases for all methods. The effect is further compounded by samples with M-proteins of lower concentrations. In samples spiked with 2.0 g/L of Dara and analyzed by Helena AGE with PD, recoveries (%CV) were: hypogamma 123% (7.6%), normal 157% (8.2%) and hypergamma 311% (10%). The same was seen in Elo-spiked samples; for 2.0 g/L of Elo, recoveries (%CV) were: hypogamma 135% (7.9%), normal 216% (14%) and hypergamma 388% (12%), and similarly for Sebia CZE with PD (Supplementary Table 6). The bias in PD quantitation proportionally increases with the gamma background and is inversely proportional to the M-protein concentration.

The TS gating technique is less susceptible to the influence of the intensity of the gamma fraction (Figure 3A,B). In samples spiked with 10 g/L of Dara and

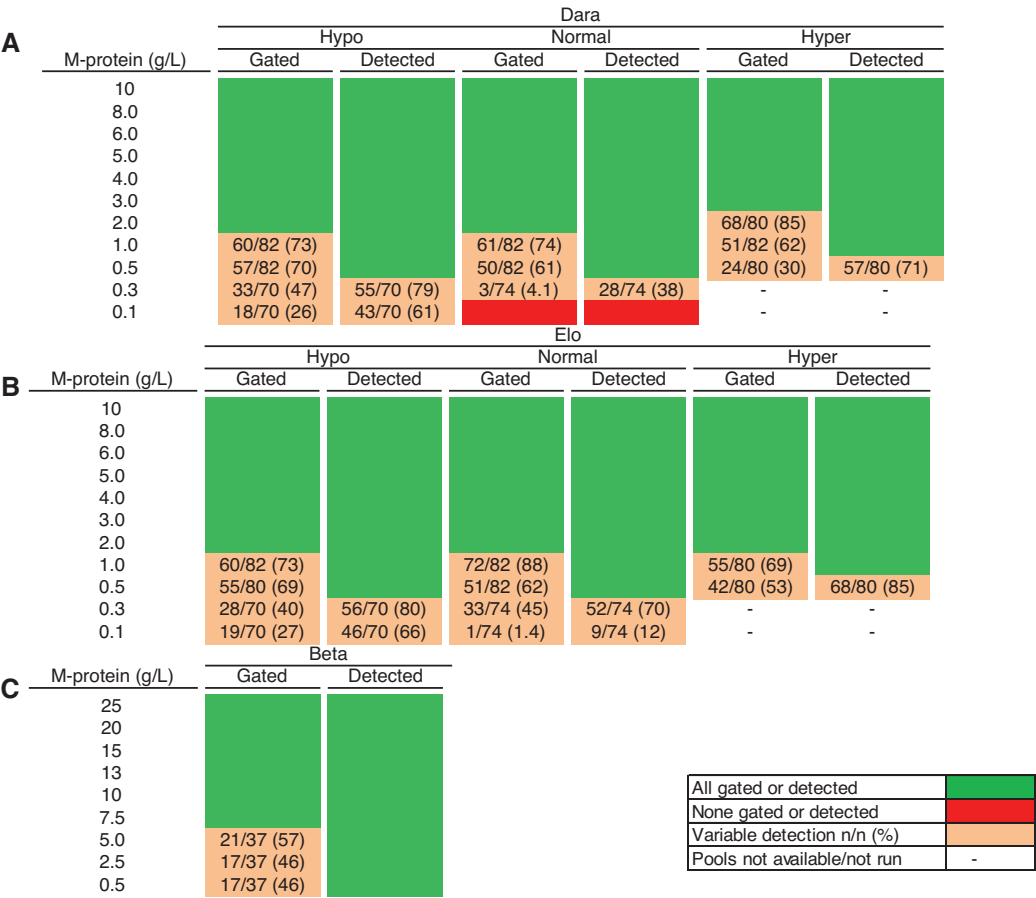


Figure 3: Serum protein electrophoresis limits of quantitation and limits of detection for all methods. The green cells represent all samples where an abnormality was gated and quantified by SPEP and/or qualitatively detected by SPEP. Number of samples, depicted as ratio (percentage) that were gated and quantified by SPEP and/or qualitatively detected by SPEP in peach-colored cells. Red cells represent samples where no abnormality was gated or detected. (A) Dara-spiked samples. (B) Elo-spiked samples. (C) Beta-migrating samples.

analyzed by Sebia CZE with TS (Supplementary Table 7), the recoveries (%CV) were closer to the target than PD, with 83% (3.3%) in hypo-, 84% (4.0%) in normal and 88% (5.9%) in hypergamma samples. Even with lower concentrations, under-recoveries were similar and did not trend, as exemplified by the analysis of 2 g/L Dara by Sebia CZE with TS with recoveries (%CV) as follows: hypogamma 78% (18%), normal 77% (7.4%) and hypergamma 80% (20%). The same was seen in Elo-spiked samples. In the samples spiked with 10 g/L of Elo and analyzed by Sebia CZE with TS, the recoveries (%CV) were as follows: hypogamma 80% (3.3%), normal 79% (3.8%) and hypergamma 81% (6.0%). In samples spiked with 2.0 g/L of Elo, recoveries (%CV) were: hypogamma 74% (10%), normal 82% (14%) and hypergamma 80% (19%). This suggests a constant negative bias for TS rather than a proportional positive bias as observed with PD.

The M-protein migration pattern

The position of the M-protein, even within the gamma fraction, affects the accuracy and precision of quantitation. This is shown by comparing Dara-spiked samples (cathodal-migrating) to Elo-spiked samples (central-migrating on AGE and anodal in CZE). Overall, M-proteins that migrate closer to the cathode and less in the central portion of the gamma fraction have higher accuracy and precision, as they suffer less influence from the polyclonal gamma background. In the samples spiked with 10 g/L of Dara or Elo and analyzed by Helena AGE with PD, the recoveries (%CV) in various gamma fraction backgrounds were as follows: hypogamma 89% (2.6%) and 95% (3.1%), normal 97% (2.5%) and 112% (5.4%), and hypergamma 116% (5.1%) and 150% (4.7%), respectively (Figure 2A,B). The difference in accuracy becomes more apparent as the M-protein concentration decreases. Given the significantly lower impact of the gamma fraction background on the TS technique, TS is also less influenced by the location of the M-protein within the gamma fraction. Dara and Elo had similar recoveries (%CV) ranging from 79% to 88% (3.3%–6.0%) when measured by Sebia CZE with TS in the three tested gamma backgrounds when spiked with 10 g/L of Dara or Elo.

Limit of quantitation

For each sample type, a LOQ was calculated and defined as the lowest M-protein concentration within the acceptable recovery range of 80%–120% with a CV

Table 2: Calculated limit of quantitation for serum protein electrophoresis.

| Dara | | | | Elo | | | Beta |
|------------|--------|-------|------|--------|-------|-----|------|
| Hypo | Normal | Hyper | Hypo | Normal | Hyper | | |
| SPEP LOQ | | | | | | | |
| Helena AGE | | | | | | | |
| PD | 3.0 | 5.0 | 10 | 4.0 | 8.0 | >10 | 13 |
| TS | >10 | >10 | >10 | >10 | >10 | >10 | >25 |
| Sebia AGE | | | | | | | |
| PD | 1.0 | 5.0 | 8.0 | 2.0 | >10 | >10 | – |
| TS | 8.0 | 8.0 | 2.0 | >10 | >10 | 10 | >25 |
| Sebia CZE | | | | | | | |
| PD | 2.0 | 5.0 | >10 | 3.0 | 8.0 | >10 | – |
| TS | 1.0 | 5.0 | 2.0 | 6.0 | >10 | 8.0 | 13 |

SPEP LOQ (g/L) for each sample pool by each method and gating technique. The LOQ was defined as the lowest M-protein concentration within the acceptable recovery range (80%–120%) and with a CV ≤20%. Samples marked with (>) are cases where a LOQ was not obtained at the tested concentrations. AGE, agarose gel electrophoresis; CZE, capillary electrophoresis; LOQ, limit of quantitation; PD, perpendicular drop gating; SPEP, serum protein electrophoresis; TS, tangent skimming gating.

≤20% (Table 2). Within each method (platform plus gating technique), there are variations in the LOQ which are dependent on the intensity of the gamma fraction background and the migration pattern of the M-protein. When the LOQ was estimated for all platforms (Helena AGE, Sebia AGE and Sebia CZE) with PD gating, it was influenced by the gamma background. Dara-spiked samples in hypo, normal and hypergamma backgrounds as measured on the Helena AGE using PD were 3 g/L, 5 g/L and 10 g/L, respectively. The LOQ was also affected by the migration pattern of the M-protein, with Dara-spiked samples having a lower LOQ as compared to Elo-spiked. In Dara- and Elo-spiked samples in normal gamma fraction background as measured by Helena AGE using PD, the LOQ was 5 g/L and 8 g/L, respectively. With PD, in some cases such as with Sebia AGE measurement of Elo-spiked samples in the normal and hypergamma fraction background, an acceptable LOQ was not obtained within the range of concentrations tested and noted as >10 g/L. This means that none of the samples had both an acceptable recovery (80%–120%) and acceptable %CV (≤20%).

For platforms where TS was applied, the LOQ increased between the hypo- and normal gamma fractions in the same fashion as seen with PD; however, the LOQ was improved in the setting of the hypergamma background. The LOQ for Dara-spiked samples in hypo-, normal and hypergamma backgrounds as measured on the

Sebia CZE using TS was 1 g/L, 5 g/L and 2 g/L, respectively. Dara-spiked samples showed a lower LOQ when compared to Elo-spiked samples using TS: 5 g/L and >10 g/L, respectively. LOQ was not obtained for Dara or Elo samples with any of the gamma backgrounds for Helena AGE using TS gating. The LOQ for the beta-migrating sample was 13 g/L for both Helena AGE with PD and Sebia CZE with TS, but was not able to be determined with Helena AGE with TS or Sebia AGE with TS methods.

The decision to gate the M-protein in SPEP

As each laboratory standard operating procedure was used and a LOQ criterion was not pre-determined, it was common to observe that the M-proteins were quantitated (or gated) below the LOQ found by the study results. The decision of the laboratory to gate an M-protein is greatly dependent on its concentration. Down to the concentrations of 2 g/L for the gamma-spiked samples and 7.5 g/L for beta samples, all M-proteins were gated and reported quantitatively (Figure 3). Below these concentrations, the number of laboratories choosing to gate progressively decreased. The heterogeneity in the laboratory's gating practices can be appreciated in Figure 4. Four laboratories performed SPEP by Helena AGE with

PD with three of the four being the laboratory's primary method. For each laboratory, the concentration of the M-protein at which they decided to stop gating differed. The laboratory that performed this method as a supplemental method gated the lowest concentration and had the highest over-recovery (360%) as compared to other laboratories that stopped gating at higher concentrations, such as laboratory 4, which stopped gating at 2 g/L and had a recovery of 148%.

As the gamma fraction background increased, the number of laboratories gating decreased. The number of gated samples at 0.5 g/L Dara for hypo-, normal and hyperwere 57/82 (70%), 50/82 (61%) and 24/82 (30%), respectively (Figure 3A). The influence of the gamma fraction intensity on the decision to gate was only seen in samples with an M-protein <2 g/L. Above this concentration, the influence of the background is not a great contributor on the decision to gate. Moreover, the migration pattern plays a role in the decision to gate low-concentration M-proteins even within the gamma fraction. Laboratories were more likely to gate low-concentration M-proteins if they were migrating more centrally within the gamma fraction. This is evident from the differences between the number of gated M-proteins between Dara, cathodal-migrating and Elo, central migrating, spiked samples. At 0.3 g/L in a normal gamma background, only 3/74 (4%) of the Dara

| | | | | | | | | |
|----------------------------|--------|--------|-------|--------|-------|--------|--------|-------|
| Helena AGE PD | Lab 1* | Lab 2 | Lab 3 | Lab 4 | | | | |
| Last gated M-protein (g/L) | 0.3 | 0.5 | 1.0 | 2.0 | | | | |
| Recovery (%) | 360 | 345 | 208 | 148 | | | | |
| Helena AGE TS | Lab 1* | | | | | | | |
| Last gated M-protein (g/L) | 2.0 | | | | | | | |
| Recovery (%) | 52 | | | | | | | |
| Sebia AGE PD | Lab 1 | Lab 2* | Lab 3 | Lab 4 | | | | |
| Last gated M-protein (g/L) | 0.5 | 0.5 | 1.0 | 1.0 | | | | |
| Recovery (%) | 143 | 228 | 100 | 151 | | | | |
| Sebia AGE TS | Lab 1 | Lab 2* | | | | | | |
| Last gated M-protein (g/L) | 1.0 | 1.0 | | | | | | |
| Recovery (%) | 25 | 60 | | | | | | |
| Sebia CZE PD | Lab 1 | Lab 2 | Lab 3 | Lab 4 | Lab 5 | Lab 6 | Lab 7 | Lab 8 |
| Last gated M-protein (g/L) | 0.3 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.0 |
| Recovery (%) | 424 | 217 | 223 | 260 | 334 | 345 | 392 | 255 |
| Sebia CZE TS | Lab 1* | Lab 2* | Lab 3 | Lab 4* | Lab 5 | Lab 6* | Lab 7* | |
| Last gated M-protein (g/L) | 0.3 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.0 | |
| Recovery (%) | 21 | 27 | 53 | 53 | 55 | 58 | 65 | |

Figure 4: Inter-laboratory gating practices.

The lowest concentration of M-protein that was gated in the Dara-spiked sample in a normal gamma fraction background for each method with its respective recovery. Labs noted with (*) are results from labs in which the method was not their primary method. PD, perpendicular drop; TS, tangent skimming; AGE, agarose gel electrophoresis; CZE, capillary zone electrophoresis.

samples were gated, while 33/74 (45%) of the Elo samples were gated (Figure 3A,B). The influence of the migration pattern was absent from hypogamma samples, as the difference between the number of gated samples between Dara and Elo was not substantial.

Discussion

While a number of clinical guidelines exist in relation to the diagnosis, monitoring and treatment of monoclonal gammopathies, these guidelines only give passing attention to the laboratory aspects of protein electrophoresis resulting in a lack of reporting standards [17]. Here we performed an international, multicenter study utilizing a single shared sample set to objectively determine the limitations in quantifying small M-proteins by defining method-specific LOQs and the variability in reporting practices among laboratories in an effort of harmonization and standardization within the field.

Currently, for the monitoring of patients with monoclonal gammopathies, it is recommended to use the same method in the same laboratory to improve the reproducibility and comparability of serial measurements [17–19]. Between-laboratory variation will greatly impact patient care if the patient gets results from different labs. Thus, standardization of quantitative results is necessary for trending of results between laboratories. Standardizing SPEP quantitation is challenging given the great heterogeneity in methodologies and gating technique as depicted in this study. Making a universal recommendation for the use of a single method and gating technique is impractical considering the diversity of different healthcare systems and hematology practices, the available technology in each country, and the costs/resources necessary to adapt and change existing instrumentation. In addition, this could result in the need to re-baseline patients with the new method, which would not only be time-consuming for the laboratory, but may also cause confusion for the clinical service and management. Nonetheless, knowledge about the limitations of your institution's current method or a method under consideration for implementation is useful and valuable. In future publications on the topic, detailing the method description of the test system in use for the conclusions should be taken into consideration by researchers in the field.

This study quantified the effects of four major factors that influence the accuracy and precision of quantitation of M-proteins: M-protein concentration, intensity of the gamma-fraction background, M-protein migration

and gating method. As a result, in methods applying PD, the gating of small M-proteins leads to over-estimation with a bias proportional to the gamma background and inversely proportional to the M-protein concentration. Over-estimation is attributed to the contribution of the polyclonal background inside the gate. As the M-protein concentration decreases, there is a much more significant inclusion of the background into the gate and therefore the reported M-protein value. The same is true for beta-migrating samples, M-proteins that migrate in the center of the gamma fraction and M-proteins present in a hypergammaglobulinemic background. In a comparison of all the methods that applied PD (i.e. Helena AGE [n=4], Sebia AGE [n=4], Sebia CZE [n=8]), Sebia AGE showed superior accuracy and LOQ as compared to when TS methods were applied. This was true for all samples tested in this study.

TS gating of low-concentration M-proteins led to underestimation with a constant negative bias for TS rather than a proportional bias as observed with PD. However, the extent of deviation from the acceptable range and target was not as substantial as seen with PD. Similar findings were also seen in other studies [11]. Higher-concentration M-proteins gated by TS can be accurately separated out from the polyclonal background. With lower-concentration M-proteins, it appears that TS is not only removing the background but is also removing a part of the M-protein. TS appeared to be not as affected by the intensity of the gamma background or the migration within the gamma fraction. The LOQ increased consistently as the gamma fraction increased when PD gating was utilized, whereas the LOQ when gating with TS increased between the hypo- and normal gamma samples and then decreased in the setting of the hypergamma background. This shows that TS has better performance than PD in the setting of hypergammaglobulinemia. However, the concomitant finding of hypergammaglobulinemia and a monoclonal protein has low prevalence, as hypergammaglobulinemia is present in autoimmune, inflammatory or infectious processes, whereas a monoclonal protein is present in pre-malignant or malignant conditions. Laboratories should comment upon the degree of gamma globulins in the background, both to aid in the appreciation of uncertainty and to indicate possible associations with such clinical conditions. In a comparison of all the TS methods (i.e. Helena AGE [n=1], Sebia AGE [n=2] and Sebia CZE [n=7]), Sebia CZE performed the best with a superior LOQ for each sample type.

We believe that these challenges with measurement reflect the fact that M-proteins have an architecture resembling a triangle with a peak at the top and continuously broaden toward the base. As the demarcation points for

both PD and TS are placed where the M-protein has a perceptible deviation, they would underestimate M-proteins unless there are no polyclonal immunoglobulins at all. In our study, under conditions of hypogammaglobulinemia, both methods underestimate the M-spike because the M-spike broadens below this point as it extends to the baseline. As shown by our data, PD has the advantage that the perpendicular drop to the baseline includes underlying polyclonal material that at least partially makes up the difference in hypogammaglobulinemic serum. As the M-protein decreases and polyclonal gamma globulins increase, the relative amount of the underlying polyclonal material becomes the major contributor yielding recoveries of 150% or higher. TS, by measuring the area above demarcations originating at the more petite end of the M-spike, consistently underestimates the quantity of the M-protein as shown in our study. Despite consistent low recovery, when the polyclonal background becomes the major factor (<4 g/L in normal or hypergammaglobulinemic samples), TS provides consistent recovery relative to the initial measurement. Both methods have advantages and disadvantages. However, one should not switch from one method when following a patient.

It is generally stated that the minimum concentration for quantification is 1 g/L, with the acceptance that quantification of M-proteins at this level is inaccurate and imprecise [17]. M-proteins <1 g/L visible on SPEP should not be quantified as the measurement is unreliable and should therefore be reported qualitatively. Those guidelines were developed when the only option for M-protein measurement by electrophoresis was to use PD. While this method had the virtue of simplicity and worked well with large M-proteins and when normal immunoglobulin production was suppressed, with small M-proteins and polyclonal backgrounds much of the measurement included was background [2]. Here we show that many laboratories gate and quantitatively report M-proteins with concentrations much lower than 1 g/L. The resulting value is inaccurate and may not truly reflect the disease status of the patient. Laboratories still gate below what is analytically acceptable as evident by the discrepancy between gating practices and calculated LOQs. Although the LOQ is held as the quantitative analytical sensitivity of SPEP, an abnormality within the gamma or beta fraction was visually down to a median 16-fold lower (range: 2–100-fold lower) than the reported LOQ, characterizing the LOD of the SPEP systems (Table 2).

There are two situations where the measurement of a low-concentration M-protein would clinically occur: at initial diagnosis, with the finding of a small non-quantifiable M-protein and in the monitoring of a therapeutic

response. In these situations, the laboratory is faced with a decision between quantifying and reporting out an analytically inaccurate result or reporting qualitatively as a small abnormality observed on SPEP and reflex to IFE/ISUB. In an initial diagnostic sample found to have any M-protein abnormality, long-term monitoring for disease progression is recommended. Most patients with a monoclonal gammopathy have a relatively small monoclonal peak, have no clinical symptoms and have the diagnosis of MGUS. This situation can be common as monoclonal gammopathies have been detected in 1% of the population older than 50 years and in 3% of those 70 years or older with incidental findings. It may be important to identify and monitor these patients as in 20%–25% serious disease such as MM, WM or primary systemic AL may develop during long-term follow-up. As a result of the long-term monitoring, there is no harm in not giving a numerical value at initial diagnosis. Only when the M-protein is large enough to be quantified accurately should a quantitative result be given.

Moreover, in therapeutic monitoring, when a patient is responsive to therapy and there is a substantial decrease in the M-protein concentration, a small non-quantifiable M-protein can be encountered. In fact, international guidelines for the classification of myeloma response recommend the use of dFLC (difference between the involved and uninvolved FLC) in place of the M-protein concentration determined by densitometric analysis if the serum M-protein is <10 g/L. If the FLC ratio becomes normal, then IFE/ISUB is required to further monitor the presence or absence of the M-protein [20]. Stopping the quantitation when the M-protein is too small to be accurately reported is acceptable because SPEP and IFE/ISUB reports need to contain adequate information to enable assessment of very good partial response (VGPR), complete response (CR) and stringent complete response (sCR). The differentiation between VGPR and CR response criteria are dependent on the performance of the IFE/ISUB to demonstrate the absence of M-protein when one was previously detected. Therefore, the practice of not gating M-proteins that are expressed at low levels is sufficient as the M-protein can be reported qualitatively. From here, clinical service can follow these qualitative results instead of monitoring quantitatively, which may be inaccurate. Furthermore, in the era of therapeutic mAbs, it has become common to see small monoclonal IgG kappa bands in SPEP and IFE/ISUB, that do not necessarily represent the original disease clone, and gating the band as an M-protein may change response criteria for patients undergoing treatment if the therapeutic mAbs cannot be accurately differentiated from the original disease clone

with subsequent testing. Reporting the therapeutic mAb qualitatively without recognizing the difference from the original clone will have a similar effect, and could mean absence of CR for the patient.

Quantitating M-proteins in the beta region can often be problematic as the M-protein may co-migrate and be obscured by other factors such as C3 complement or transferrin within the fraction making their quantification challenging. As the beta-migrating sample was pooled from a single patient and not generated from spiked mAbs, the initial undiluted concentration was determined in house (Mayo Clinic, Rochester, MN, USA) using Helena AGE with PD and total IgA. Therefore, all resulting values generated by other methods should be compared relative to Helena AGE with PD and trended within the dilutions of each method to assess loss of accuracy, as no method is truth. Both Helena AGE with PD and Sebia CZE with TS showed good accuracy down to 10 g/L. This is most likely attributed to the specific migration of this M-protein which had its peak separate from the rest of the beta fraction. One way to improve the accuracy of low-concentration M-proteins is to quantify isotype-specific immunoglobulin classes [21]. This can be especially useful with beta-migrating IgAs overlapping with other proteins.

A conclusion about which methodology is most sensitive is challenging. Firstly, the sensitivity in detection may be dependent on the reader and not on the methodologies with some labs being more conservative in what they call an M-protein, and not all methods were run and read in the same laboratory and by the same reader. Additionally, there may have been some bias in the reading as no negative samples were included in the blinded sample set. The reader may have expected that all sent samples contained an M-protein, as the initial letter with communication to the participants, if shared with the technologists, stated details of the study design. The true value of the spiked Dara and Elo in the pools relied upon the concentration stated in the vials of the pharmaceutical preparations without further verification by the lead laboratory besides measurement of total protein and total IgG before and after pools were prepared. A small error margin in the standards preparation can be considered. Another limitation for the study is that likely the experiments were run in a short period of time, with a single operator, and dedicated attention to gating and detection of M-proteins and results could be superior to the ones observed in routine practice.

In conclusion, a large, multicenter SPEP and IFE/ISUB study was performed utilizing a single shared sample set. The study displays the lack of accuracy and precision of quantifying low-concentration M-proteins along with other factors that influence the performance of

the assay. With an understanding of the limitations, laboratories may become more conservative in what they are quantitating and reporting.

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 the 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 Prof. Mario Plebani, from the Department of Laboratory Medicine, University-Hospital of Padova (Padova, Italy) for coordinating the distribution of samples in Italy. The authors would like to thank 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: Marie Therese Melki and Stephen Bell did not participate in the study design or data analysis. All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

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/10.13039/1000000871>. 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 interests' disclosures: None of the other authors have conflicts of interest to declare.

References

- Katzmann JA. Screening panels for monoclonal gammopathies: time to change. *Clin Biochem Rev* 2009;30:105–11.
- Keren DF, Schroeder L. Challenges of measuring monoclonal proteins in serum. *Clin Chem Lab Med* 2016;54:947–61.
- Willrich MA, Katzmann JA. Laboratory testing requirements for diagnosis and follow-up of multiple myeloma and related plasma cell dyscrasias. *Clin Chem Lab Med* 2016;54:907–19.
- Keren DF. Protein electrophoresis in clinical diagnosis. London, New York: Arnold; Oxford University Press, 2003, 402 p.
- Kyle RA, Gertz MA, Witzig TE, Lust JA, Lacy MQ, Dispenzieri A, et al. Review of 1027 patients with newly diagnosed multiple myeloma. *Mayo Clin Proc* 2003;78:21–33.
- Katzmann JA, Kyle RA, Benson J, Larson DR, Snyder MR, Lust JA, et al. Screening panels for detection of monoclonal gammopathies. *Clin Chem* 2009;55:1517–22.
- 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.
- Mills JR, Kohlhagen MC, Dasari S, Vanderboom PM, Kyle RA, Katzmann JA, et al. Comprehensive assessment of M-proteins using nanobody enrichment coupled to MALDI-TOF mass spectrometry. *Clin Chem* 2016;62:1334–44.
- Murray DL, Ryu E, Snyder MR, Katzmann JA. Quantitation of serum monoclonal proteins: relationship between agarose gel electrophoresis and immunonephelometry. *Clin Chem* 2009;55:1523–9.
- Katzmann J, Keren D. Strategy for detecting and following monoclonal gammopathies. In: Detrick B, Schmitz JL, Hamilton RG. *Manual of molecular and clinical laboratory immunology*. Washington, DC: ASM Press, 2016:112–24.
- Schild C, Wermuth B, Trapp-Chiappini D, Egger F, Nuoffer JM. Reliability of M protein quantification: comparison of two peak integration methods on Capillarys 2. *Clin Chem Lab Med* 2008;46:876–7.
- Genzen JR, Murray DL, Abel G, Meng QH, Baltaro RJ, Rhoads DD, et al. Screening and diagnosis of monoclonal gammopathies: an international survey of laboratory practice. *Arch Pathol Lab Med* 2018;142:507–15.
- 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.
- 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.
- Mills JR, Kohlhagen MC, Willrich MA, Kourelis T, Dispenzieri A, Murray DL. A universal solution for eliminating false positives in myeloma due to therapeutic monoclonal antibody interference. *Blood* 2018;132:670–2.
- The United States Pharmacopeia. National formulary. Chapter 476 & Chapter 1086. Rockville, MD: United States Pharmacopeial Convention, Inc.; 1979.
- 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.
- Tate J, Panteghini M. Standardisation – the theory and the practice. *Clin Biochem Rev* 2007;28:127–30.
- 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.
- 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.
- Tate JR, Graziani MS, Mollee P, Merlini G. Protein electrophoresis and serum free light chains in the diagnosis and monitoring of plasma cell disorders: laboratory testing and current controversies. *Clin Chem Lab Med* 2016;54:899–905.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/cclm-2019-1104>).