

Review

Standardizing plasma protein measurements worldwide: a challenging enterprise

Giampaolo Merlini^{1,*}, Søren Blirup-Jensen²,
A. Myron Johnson³, Joanna Sheldon⁴ and Ingrid
Zegers⁵ on behalf of the IFCC Committee on Plasma
Proteins (C-PP)

¹ Biotechnology Research Laboratories, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo, Department of Biochemistry, University of Pavia, Pavia, Italy

² Department of Clinical Chemistry, University Hospital, Lund, Sweden

³ Department of Obstetrics and Gynecology, University of North Carolina School of Medicine, Chapel Hill, NC, USA

⁴ Protein Reference Unit, St. George's Hospital, Blackshaw Road, London, UK

⁵ Institute for Reference Materials and Measurements (IRMM), Joint Research Centre, European Commission, Geel, Belgium

Abstract

The need for harmonizing laboratory results is particularly intense in the field of quantitative protein assays in consideration of the clinical impact of specific protein measurements and their relevance in monitoring disease. We report the efforts made by the Committee on Plasma Proteins of the IFCC Scientific Division to achieve worldwide comparability in plasma protein results. We focus on the production of reference materials and the methods applied throughout their production process. Particularly, the recent characterization of ERM-DA470k/IFCC and ERM-DA472/IFCC has demonstrated that it is possible to reproduce the earlier established procedures and thereby maintain standardization. Plasma protein reference materials have had a substantial impact in improving the harmonization of patient protein results that should translate into better patient care.

Clin Chem Lab Med 2010;48:1567–75.

Keywords: harmonization; plasma proteins; reference material; traceability.

*Corresponding author: Giampaolo Merlini, Biotechnology Research Laboratories, Fondazione IRCCS Policlinico San Matteo, Department of Biochemistry, University of Pavia, P.le Golgi, 19, 27100 Pavia, Italy
Phone: +39-0382-502995, Fax: +39-0382-502990,
E-mail: gmerlini@smatteo.pv.it
Previously published online October 20, 2010

Introduction

It has been stressed, particularly by the European laboratory medicine community and by the Joint Committee on Traceability in Laboratory Medicine (JCTLM), that it is extremely important to harmonize medical laboratory results, especially with the current levels of population migration and mobility. Among other things, standardization approaches must include traceability of values of reference materials, calibrators, and controls, and demonstration of commutability of reference materials among assay methods (1).

The quantitative determination of human proteins in biological fluids, such as plasma, serum, cerebrospinal fluid, and urine serves as an important tool in diagnosing diseases and in monitoring the course of a disease and the effect of treatment. In order to improve, the clinical value of these measurements, standardized assays must be selected, and a yardstick in the form of a reference material has to be identified. However, standardization of human serum proteins presents several challenges due to a wide array of possible post-translational modifications, genetic variability, presence of ligands bound to proteins, their sensitivity to proteolysis, and the wide range of concentrations in pathophysiological conditions. The task of standardizing protein measurements has been undertaken by the scientific community using different approaches, with variable rates of success. The main step forward toward achieving the standardization of plasma protein measurement has been the production of the stable and reliable certified reference material (CRM) 470 (now ERM-DA470) and its more recent successor ERM-DA470k/IFCC. We describe the challenges, pitfalls and success of such enterprise, highlighting the essential role of a working reference measurement system in order to validate the final outcome represented by results in biological samples.

The quest for harmonizing protein quantification assays

The increasing and widespread use of serum protein quantification in the clinical setting led to the necessity to standardize the employed assays. Several attempts have been made to reach this goal. In the late 1960s, Rowe et al. produced a reference preparation for human serum immunoglobulin measurements (2, 3) in an attempt to increase the comparability of immunoglobulin results among laboratories. It was thought that the large between laboratory variability was due to differing standard preparations, antisera with dif-

ferent affinities and specificities, and use of a variety of manual methods. These authors reported that when a radial immunodiffusion method and a common calibrator were used, the reproducibility of the results improved considerably (2). This material was the first major step in harmonization of proteins assays. The batch was processed in two sites and became the International Reference Preparations WHO 67/86 and WHO 67/95, 67/97 and 67/99. Batch 67/86 was assigned 100 U per ampoule for IgG, IgA and IgM, and values for the other three preparations were derived by comparison to 67/86 using a radial immunodiffusion method. Some years later, mass concentration values were ascribed to these preparations despite wide variations among the participating "expert" laboratories (4).

In 1973, the WHO and the IUIS (International Union of Immunological Societies) embarked on the production of a new material. The chosen preparation became the US National Reference Preparation (USNRP) and the WHO Reference Preparation for six human serum proteins (WHO6HSP). International units were arbitrarily assigned for albumin (ALB), α_1 -antitrypsin (AAT), ceruloplasmin (CER), α_2 -macroglobulin (A2M), transferrin (TRF) and C3c, while immunoglobulin concentrations were assigned with respect to 67/86 (4, 5). Mass concentrations were assigned to USNRP and, although this generated some differences from the values for WHO 67/86, USNRP became the primary reference material for serum proteins. During this same period, other preparations were available. For example, Reference Preparation for Serum Proteins (RPSP) lots 1, 2 and 3 were produced by the College of American Pathologists (CAP). As the value assignment was done by consensus comparison to the preceding material, over time there was a noticeable drift from the original values assigned with respect to USNRP.

The 1970s also saw a major change in how serum proteins were measured, and the introduction of turbidimetric and nephelometric assays generated an additional requirement for a non-turbid reference material. The first significant intervention by the IFCC in standardizing plasma protein assays was the preparation of IFCC 74/1, a protein reference material that was sterile, non-turbid liquid suitable for all techniques (gel-based and turbidimetric/nephelometric). This material was collected according to strict protocols to minimize contact with tissue and leukocyte proteases, interference from fibrinogen and denaturation by freeze drying. Evaluation of this material considered method-specific variation in the calibration of the new standard, so the three available methods (Laurell electroimmuno assay, or "rocket immunoelectrophoresis", radial immunodiffusion and automated immunoprecipitation) were used with antiserum from one defined source. There were certainly confounding issues due to the high turbidity of WHO 67/86, non-linearity, and different behaviors of the IFCC and WHO materials, but the outcome was a new material that "undoubtedly improved the quality of immunoglobulin measurements" (6).

The 1980s were a time of unprecedented change in protein laboratories. Procedures moved from gel-based quantification of proteins (requiring 24–48 h) to fixed-time nephelo-

metry (requiring 3–5 h) to rate nephelometry or automated turbidimetry with results generated within minutes. The workload expanded dramatically and the major diagnostic companies introduced automated protein analyzers. Quality assessment (QA) schemes for specific proteins were introduced and there was the expectation that results were reliable and comparable. The investigation of monoclonal proteins, B cell malignancies and immune deficiency became integral parts of laboratory repertoires; rapid analysis of C-reactive protein (CRP) was introduced and many labs started to quantify a number of serum proteins, such as AAT, TRF, complement factors C3 and C4 and haptoglobin (HPT). Instead of protein analyses being confined to specialized laboratories, these tests became part of the routine activity in clinical chemistry laboratories, and certain protein assays began to be used as emergency tests.

The protein community should have been in a situation where theoretically the major serum protein measurements in automated systems were well standardized. However, QA data from the US and Western Europe showed that some protein values varied by as much as 100%. This could have been related to confusion about the use of purified reference materials (lacking commutability), uncritical use of different secondary reference materials, drifts in the value assignment to these materials, and/or the availability of materials that were not suitable for the modern optical measurement systems. The severity of the problem and the possible solution were clearly shown in a study by the UK National External Quality Assurance Scheme (UK NEQAS) for specific proteins (7). In this experience, results from pairs of dilution related samples were collected over five UK NEQAS distributions. The next two sample distributions were sent with a new reference preparation (designated SPS-01), with instructions to return results that were calibrated against it. The effect of using SPS-01 recalibration in inter-laboratory agreement in the UK NEQAS for immunoglobulins, C3, C4 and AAT concentrations was significant. There was a statistically significant ($p < 0.001$) improvement in the overall between-laboratory precision for all six analytes, when SPS-01 was used as a common calibration material. Interestingly, the within-laboratory imprecision also improved for the majority of analytes when this new calibrant material was used (7). Another study increasing the number of investigated proteins and the number of participating countries showed similar results, with marked improvement in inter-laboratory agreement when a common calibration material was used (8).

From these results, it was clear that a single internationally agreed reference material to reduce between-laboratory variability was needed, and in 1989 the Committee on Plasma Proteins (C-PP) of the IFCC Scientific Division began the process of preparing, characterizing and calibrating a new secondary (matrix) reference preparation for 14 serum proteins. The material was also spiked with and certified for CRP. CRP concentrations in serum can increase more than 100-fold in cases of infection or inflammation, and the concentration in serum from healthy people is too low for a reference material. Therefore, purified protein was spiked

into the processed serum. This material was produced under yet more stringent conditions than previous reference materials for serum proteins. For example, donors were required to have fasted overnight before blood collection, and lipemic, turbid, hemolyzed or jaundiced donations were discarded and any residual lipids were absorbed over microparticulate silica. The material was also subjected to considerably more analysis; donor samples had both AAT and HPT phenotypes recorded and samples showing positive rheumatoid factor or monoclonal proteins by electrophoresis were discarded. The donations were screened for hepatitis B and C virus and also for HIV. Four pure protein preparations were produced and a new value assignment procedure was introduced. This material was certified by the European Community Bureau of Reference (BCR) as certified reference material (CRM) 470 and was released by BCR and CAP in a joint initiative; it became available in Europe in 1993 and in the US in 1994 (9). It was later renamed ERM-DA470.

Production of a reference material to replace ERM-DA470

In 2004, supplies of ERM-DA470 (originally released as CRM 470) were close to being exhausted. Therefore, the C-PP and the European Institute for Reference Materials and Measurements (IRMM) took up the task of producing a new replacement material. In order to maintain harmonization among assays, the new material had to be produced in such a way that protein properties in the material, and values assigned to them, were in line with those of ERM-DA470, as this material had clearly been very successful (10). This was a demanding task, as small changes in processing procedures could have a major impact on the properties of the material as a whole or on properties of individual analytes. However, the development and production of a new batch of reference material was also an opportunity to optimize or change certain aspects. Particularly, it was decided to include β_2 -microglobulin (B2M) in the panel of certified proteins and not to certify α_1 -antichymotrypsin (ACT), according to their clinical use. B2M is a powerful prognostic indicator in multiple myeloma and is commonly measured in patients with B cell malignancy (11). However, the lack of designated reference material for B2M results in marked differences in the concentrations measured by various commercial assays, thus preventing comparability among different measurement centers. Also, ACT is now seldom measured in clinical laboratories.

Pilot studies

Pilot studies were performed in order to verify the consequences of planned changes, such as the addition of recombinant B2M, changes in sodium azide concentrations and the replacement of dialysis with diafiltration (12). The pilot batch was spiked with both CRP and B2M; the latter was shown to be stable in the material that demonstrated commutability for at least three major assay systems. Furthermore, spiking the material with B2M did not affect the

stability of the other proteins. Freeze-drying and subsequent reconstitution of the material led to incomplete CRP recovery when this protein was measured with two commercial methods. This lack of recovery could be alleviated to a considerable extent by the addition of calcium prior to freeze-drying (12). However, it was decided to not do this for the final ERM-DA470k/IFCC batch, as the effect of calcium on the stability and commutability of other proteins was unknown. Consequently, a separate liquid frozen material was produced in order to provide optimal conditions for CRP (ERM-DA472/IFCC).

Suitability of the candidate reference material: homogeneity, stability and commutability

ERM-DA470 was produced in the early 1990s after extensive feasibility studies and was characterized carefully in terms of the physicochemical properties of the proteins and the stability of the material (10). The stability of the certified proteins has since been monitored on a yearly basis. The properties of the new reference materials ERM-DA470k/IFCC and ERM-DA472/IFCC were also evaluated in detail. The between-vial heterogeneity and stability over short (4 weeks) and long periods of time (1 year) were determined for all proteins to be certified (13, 14). The possible uncertainties due to long-term stability was included in the uncertainty of the certified values, as required by ISO Guide 34 (15).

The commutability, or potential for harmonization of results from different methods calibrated with the material, had been demonstrated in practice for ERM-DA470 for all certified proteins, except for CER. As the new materials were produced along similar lines, they were expected to have similar properties in terms of commutability. As mentioned above, for the newly certified protein B2M, commutability studies were performed on the pilot batch (12).

Characterization of ERM-DA470k/IFCC

It was possible to establish certified values for A2M, α_1 -acid glycoprotein (AAG), AAT, ALB, C3, C4, HPT, IgA, IgG, IgM, TRF, and transthyretin (TTR) in ERM-DA470k/IFCC (16) and for CRP in ERM-DA472/IFCC (14). For B2M, the assignment of a certified value to ERM-DA470k/IFCC is planned in the near future. The material concentrations are comparable to those in ERM-DA470, although somewhat lower for most proteins. The establishment of the certified values was performed using the value transfer procedure developed originally for ERM-DA470, with ERM-DA470 as calibrator (17). The characterization measurements were performed using a wide range of immunoassays, based on different analytical principles (nephelometry, turbidimetry) and different antibody specificities (13). The principles of these procedures are currently listed by the Joint Committee for the Traceability of Laboratory Medicine (JCTLM) as reference methods/procedures. The transfer procedures have been further optimized for use with these immunoassays. The "closed" value transfer procedure (described in more detail below), previously used with only one type of instrument, has now also been optimized for all the platforms partici-

pating in the study. It was shown to give results equivalent to those obtained with the “open” value transfer procedure.

The between-method variation was primarily around 2%, and in all cases below 4%, for the 12 certified proteins. This is quite low, particularly taking into consideration the variety of assay designs and antibodies used, as well as the complexity of the measurands, with often a wide mixture of isoforms being measured.

The certification strategy was designed to include all relevant uncertainty components in the final uncertainties, but also to keep the uncertainties as low as possible. The uncertainties of the certified values were estimated by combining the uncertainties from the calibrant ERM-DA470 and the characterization measurements with terms reflecting the possible inhomogeneity and uncertainty due to long-term storage. As such, they were necessarily increased compared to the uncertainties of ERM-DA470 (16). In particular, some proteins (e.g., AAG, AAT) had uncertainties well below 2% (e.g., 0.8%) in ERM-DA470. In these cases, the addition of all relevant uncertainty components in ERM-DA470k/IFCC led to a significant increase, bringing the uncertainties of the values for these proteins to between 2% and 3%, in line with most of the other proteins. The relative increase in uncertainty is comparatively smaller for the other proteins (e.g., from 2.0% to 2.8% for IgA, and from 2.0% to 3.2% for albumin), as for these proteins a larger estimate of the uncertainty was made from the beginning in the ERM-DA470.

There are no generally accepted target values for the uncertainties of serum proteins, and discussions on this topic are ongoing (18). Still, the final uncertainties are acceptable compared to some indicators. For example, Target Errors compiled in the Westgard biological variations database (13.5% for IgA and 3.9% for albumin, <http://www.westgard.com/biodata->

base1.htm) and the maximum allowable deviations for ring trials as recommended by the German Bundesärztekammer (20% for both IgA and ALB, values of November 2007). The final uncertainties of the reference material must be considered in the estimation of uncertainties of patient results. As the aim is to keep these as low as possible, all steps in the traceability chain should be examined so as to minimize the added uncertainties.

The traceability chain

In order to measure a specific protein present in biological fluids, a method with well-defined performance must be selected, and a yardstick in the form of reference or calibrator materials needs to be identified. Particularly, the manufacturer's product calibrator must be value assigned using higher-level reference materials. In other words, a chain of validated steps should be established from a stated reference to the final reportable result. In metrological terms, traceability is defined as the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty (19). The uppermost level of the calibration chain should be a stable reference. The SI (Système International d'Unités), and the high-level physical standards held in Paris, represent such a reference that is well guarded (by the BIPM), exactly defined, and stable over long periods.

In the standardization work with international reference preparations for human serum proteins (ERM-DA470, ERM-DA470k/IFCC, and ERM-DA472/IFCC), protocols for the production, characterization and value assignment have been developed to secure optimal transfer of measurement trueness in all the steps of the traceability chain (Figure 1).

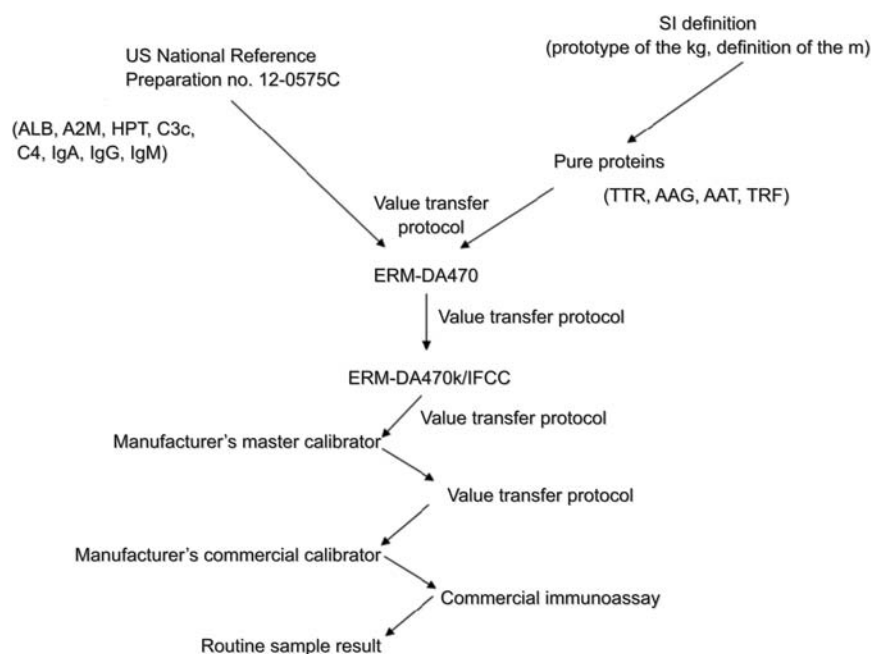


Figure 1 Traceability chain for serum protein measurements showing the calibration hierarchy going from the physical form of the kg and meter down to the patient results, via pure protein and matrix materials.

Pure protein preparation

In the first step of the traceability chain for plasma proteins, values are transferred from the definition of the measurand to the pure protein preparation. This requires the production of purified proteins, either as native from, e.g., human sources or as recombinant proteins. Procedures for the purification of TTR, AAG and TRF have been described (20). Similarly, a purified reference preparation based on recombinant cystatin C has been obtained (21). Such primary reference preparations, intended to be used for assignment of values to serum-based secondary reference materials, must meet several criteria:

1. The pure protein should be representative of the same protein found in a serum or plasma pool, both in terms of macro-heterogeneity (e.g., genetic variants) and micro-heterogeneity (e.g., different carbohydrate isoforms).
2. The protein should be as native as possible, with both biological and immunological identity preserved. The only way to achieve this is to use mild and highly-reproducible purification methods with high recovery rates to ensure that no significant sub-fractionation has occurred. For proteins like TRF, AAG and TTR, recovery of 50% could be obtained.
3. The purity of the final preparation should be as high as possible (preferably >98%), as judged by both immunochemical and non-immunochemical methods. However, in a balance between recovery and purity, recovery must be given preference.
4. The pure protein must remain stable at least until the calibration of the secondary reference material has been performed.

Dry mass determination

Once the protein has been successfully purified, a reliable method for the exact determination of the concentration must be established. A procedure has been developed for the determination of the dry mass of a pure protein preparation dissolved in aqueous solution (0.1 mol/L KCl) (22). The procedure – based on repeated dryings and weighings – not only renders the concentration (in g/L), but additionally gives the partial specific volume of the protein (in mL/g) – the latter being an important parameter for the characterization of a specific protein. This procedure has been used with excellent results for TTR, AAG, AAT and cystatin C (21, 22).

Method optimization

Quantitative protein determinations in clinical laboratories are today most often performed using automated systems based on turbidimetry or nephelometry. However, slight variations in the assay principle, in the programming of the instrument or in the reagents may lead to different results. This has led to the prerequisite of method optimization and standardization (23). An optimized test system should preferably offer a wide measuring range, combined with a wide security range, which for the user means fewer repeat runs and maximum security against antigen excess.

Value transfer procedures

The concept for assignment of values to serum proteins in a target material using a reference preparation was developed in the late 1980s for the transfer of values to 14 serum proteins in the international reference preparation for human serum proteins CRM 470 (now ERM-DA470) (9, 10, 17). Two models were developed: the direct value transfer between serum matrices, and the indirect value transfer from a pure protein preparation to a serum protein material. The practical procedure (the transfer protocol) is based on six dilutions of the reference preparation assayed together with six dilutions of the target material. Using this method imprecision is reduced and the proportionality of the two materials (i.e., the presence or absence of matrix effects) can be assessed directly by evaluating a single regression plot. If no matrix effects are found the regression line will pass through zero, with a slope equal to the ratio of the concentrations of the two materials. The transfer protocol is based on a multiple-point value assignment obtained by several measurements made each day, repeated over several (usually four) days. An important prerequisite is that all reconstitutions and dilutions are controlled by weighing. The more theoretical description of the procedure and the necessary mathematical equations are discussed in (17), whereas a practical protocol with examples is given in (24).

The direct procedure for transferring values between matrix materials can be used in the traceability chain for assigning values to new reference preparations and to manufacturer's working calibrators. In open measuring systems, dilutions of the reference preparation with known protein concentrations are used to calibrate the instrument platform, whereas dilutions of the target material are measured as samples. In closed measuring systems it is not possible to use the reference preparation as the calibrator, and the manufacturer's dedicated calibrator must be used as the calibrator, and both the reference preparation and the target material be assayed as unknown samples. In this case, the data will therefore yield two regression lines, and the ratio of these two slopes will give the ratio of the concentrations of the analyte in the two materials. In both cases, several measurement series with independent calibrations are performed on each of 4 days.

This value assignment procedure has been shown to be very robust and thereby minimizes the risk of value drift. From previous experience some recommendations must be considered:

1. In order to obtain good homogeneity of results in ascribing values by immunochemical methods, it is necessary to optimize the method – as described in the protocols – and to use defined reagents optimized for the method.
2. Before launching a value assignment campaign, it is important to conduct a trial value assignment, in which the developed transfer protocols are tested and the participating laboratories qualified.

Final value assignment results

In the certification campaign for CRM 470, 36 laboratories in Europe, Japan and the USA were preliminarily selected for

value assignment. Of these, 30 qualified for the final certification campaign. The laboratories obtained excellent results with CVs well below 5% for all serum proteins, except for CER (CV 8.48%). For pure protein preparations, the CVs were below 3% (10). For the characterization of the new serum protein reference material ERM-DA470k/IFCC, 20 laboratories participated in the pilot trial, and 18 continued in the final value assignment. Again, all laboratories obtained excellent results, with CVs well below 4% for all serum proteins certified in the material (16). In the certification of the new CRP reference material the same value assignment principles were used. Eight laboratories participated and the overall CV was 4.3% (14). The low variation makes it possible to certify values with uncertainties that are sufficiently low compared to, e.g., the biological variation of CRP (Westgard Target Error 56%) and allowable relative deviations specified by the *bundesärztekammer* for ringtrials (20%), to make the material fit for purpose.

The value assignment procedure has been used not only in the official certification campaigns, but also when certified reference materials have been used to assign values to commercial calibrators and controls.

The effect of availability of reference materials ERM-DA470 on harmonization of serum protein assays

As noted above, the rather large between-laboratory and between-manufacturer variance in values obtained for serum protein assays before 1990 was the major driving force for the production of ERM-DA470. The C-PP intended to evaluate the effect of this material on assay harmonization by sending two high-quality samples to several thousand laboratories prior to and a few years after release of the material, with data compared for the same laboratories and assay methods to reduce extraneous influences. However, this comparison could not be performed as planned: first, over half of the participating laboratories did not report results for the first pair of samples to the C-PP. Thus, the second sample was sent to a different set of laboratories. Second, the assay methods and manufacturers of kits had changed in the interim between the two sets of assays. For the first set, most assays were performed using immunonephelometry, primarily with Beckman (now Beckman Coulter) and Behringwerke/Dade Behring (now Siemens) nephelometers. In contrast, for the second set, assays were performed using a large number of manufacturers' systems, most utilizing immunoturbidimetry. The available data were analyzed by the C-PP and presented at conferences, but were not published because of the discrepancies.

However, several analyses utilizing the results of national quality control programs have been published, including data from the UK (UK NEQAS), Sweden (EQAS), Germany (DKGC-RfS), the Czech Republic, and the US (CAP) (25–29). The data could not be pooled because of differences in the evaluation of data by the QA programs: various methods of outlier removal were used (with no trimming for the UK NEQAS), and some programs submitted individual laboratory

results, while others (e.g., CAP) included only means, medians, and variances for individual assay groups (instruments and manufacturers) and for the overall data sets. In spite of this, the results were remarkably similar among countries, with substantial harmonization of between-laboratory results for AAT, HPT, and TRF, and moderate harmonization for C3, C4, and immunoglobulins. The between-assay CVs, as reported for the CAP data, were as expected moderately lower than the overall between-laboratory results from other countries, which included both within- and between-manufacturer variances.

Results varied for AAG and TTR, while CER variability was essentially unchanged, as has been recently reported (30, 31). Data for CRP have not been included in most previous publications. However, CRP assays, both traditional and "high-sensitivity", showed higher variability between laboratories and manufacturers in 2000 than before the release of CRM 470, possibly due to the rush by manufacturers to introduce high-sensitivity assays.

Recent QA programs have shown essentially no further decrease in total between-laboratory variability for most proteins, with increases in some cases. Previously published results for the US and the UK (26), modified to include 2009 data, are shown in Figure 2. The graph for the UK QA shows between-laboratory CVs, while that for the US shows between-assay, or "manufacturer," CVs (some manufacturers have more than one assay system). Results from individual laboratories data were not available for the latter. For the previous report by C-PP, as well as for this report, results for proteins with very low concentrations were excluded (e.g., samples with PiZZ AAT phenotype or immunoglobulin concentrations below 0.050 g/L), as were results for immunoglobulins in samples containing known monoclonal immunoglobulins. As can be seen, there is still substantial between-manufacturer variability in the CAP surveys for several proteins, as has also been shown in German QA (*Deutsche Vereinte Gesellschaft für Klinische Chemie und Laboratoriumsmedizin-Referenzinstitut für Bioanalytik* or *DGKL-RfB*) (32). Several factors may be involved, including suboptimal value transfers from ERM-DA470 to the manufacturer's master calibrators and subsequently to product calibrants and controls. Poor calibration curve fitting and the use of unstable calibration and control materials may also contribute.

A graph of the overall data from the German program was also included in the previous publication (29), but it is not included here because of changes in their presentation of the data. The previous analysis was based on the total data sets, while means, standard deviations, and CVs are now computed on only the subset of results that meet the acceptable limits of deviation from target values (32). As a result, the current CVs are much lower.

Problems persist for most common assay methods if protein concentrations are very low, if the protein analyte is labile, if there are significant differences in epitope recognition by the antibodies used, or combinations of these. For example, CER is a very labile protein, and different antisera may react with different fragments if the protein is not intact. As stated above, harmonization of assays for this protein has been less than optimal in essentially all reports. Complement component C4,

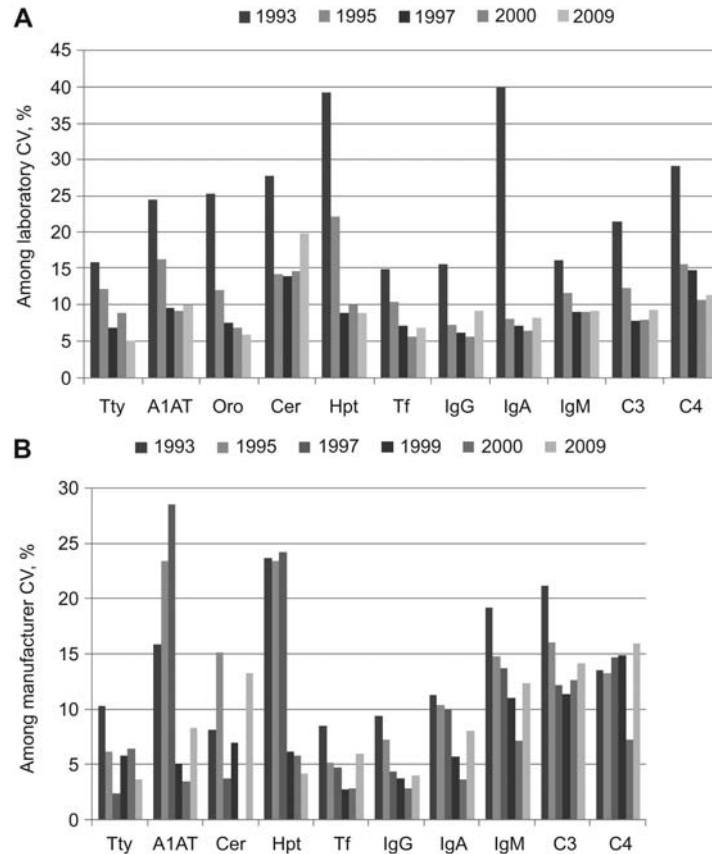


Figure 2 Mean coefficient of variations (CVs) for external quality assessment schemes for protein assays in (A) the UK (UK NEQAS) (between-laboratory CV) and (B) the US (College of American Pathologists) (between-manufacturer CV). TTY, transthyretin or prealbumin; A1AT, α_1 -antitrypsin; Oro, orosomucoid or α_1 -acid glycoprotein; Cer, ceruloplasmin; Hpt, haptoglobin; Tf, transferrin.

like CER and TTR, is present in relatively low concentrations in healthy individuals, challenging the sensitivity and precision of routine (non-enhanced) immunoassays. Physiologically, CRP concentrations are below 2 mg/L in most individuals, and at this concentration assays require enhanced-sensitivity (examples include particle-enhanced nephelometry or turbidimetry and enzyme or fluorescence immunoassays). Even using these assays, harmonization of CRP has been poor. Figure 3 is a Youden plot of manufacturer medians for two 2009 CAP specimens with the median of medians reported concentrations of 1.2 and 2.9 mg/L. As can be seen, the values varied over an eight-fold range. The Youden plot for two samples in the 2009 DKGC-RfS IG4/09 survey is shown in Figure 4; median concentrations in this case are approximately 15-fold higher than for the CAP samples, but marked variation is noted even at these higher concentrations (32). Figure 5 is a Youden plot of all CER results for two samples in the 2009 UK NEQAS surveys. As can be seen, the results also show a wide range of reported concentrations, although differences in proportionality (the ratio of the value for one sample to that for the other) are less frequent than for CRP.

It is obvious that the introduction of CRM 470/ERM-DA470 resulted in a substantial reduction in between-manufacturer and between-laboratory variability for several serum proteins,

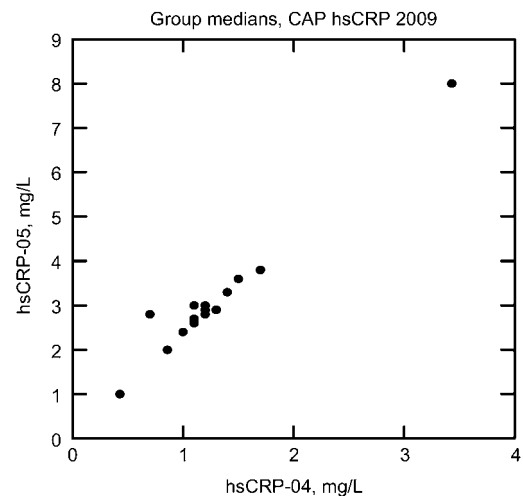


Figure 3 Youden plot of C-reactive protein values for two specimens in the 2009 College of American Pathologists surveys. Each dot represents the median value obtained by all laboratories using the same commercial system (same method, platform, and calibrators).

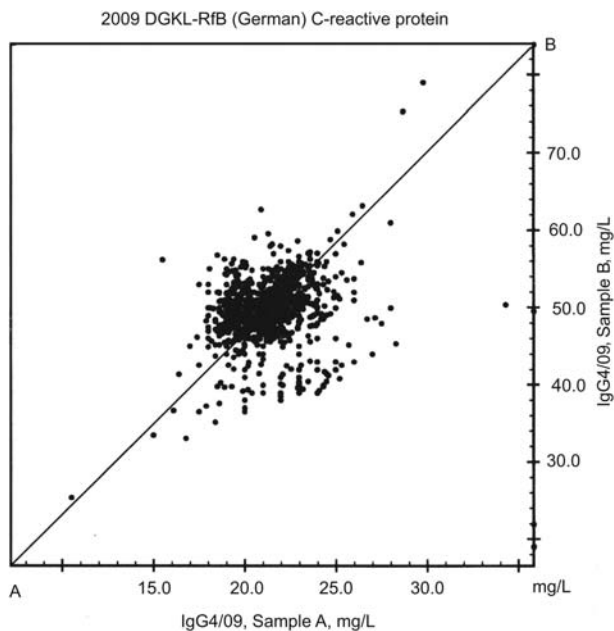


Figure 4 Youden plot of C-reactive protein values obtained by individual laboratories for two specimens in the 2009 German (DGKL-RfB) surveys.

IgG4/09 refers to the sample mailing, not the evaluated analyte. Outliers are not included in the graph. (Adapted from reference 33, with permission).

including AAT, HPT, TRF, the three major immunoglobulins, and the complement components. However, recent data show slight increases in variability when compared to results from the later 1990s, and TTR, CER, C4, and CRP continue to show substantial between-manufacturer differences.

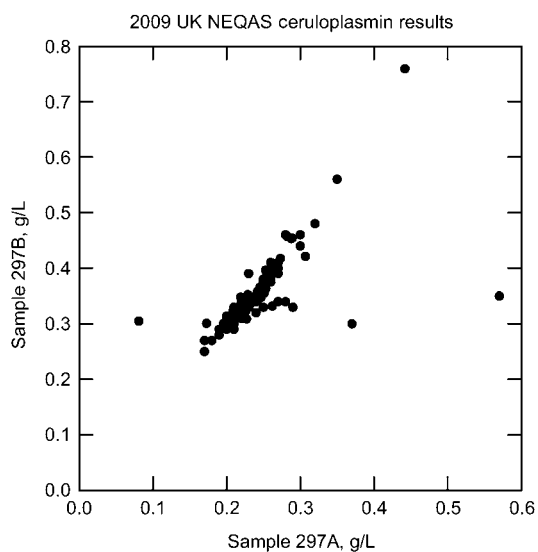


Figure 5 Youden plot of ceruloplasmin values obtained by individual laboratories for two specimens in the 2009 UK NEQAS surveys.

Conclusions

Since the release of the first international Reference Preparation ERM-DA470, QA programs worldwide have shown significant improvements in comparability of serum protein measurements. The production and characterization of the new reference material, ERM-DA470k/IFCC, has demonstrated that it is possible to reproduce the earlier established procedures and hopefully maintain standardization. The value assignment procedures have now been used for several different proteins and have demonstrated their robustness. In this way, it has been possible to transfer trueness reliably from the pure protein preparation to the final patient result, maintaining full traceability through all the steps. This goal has been achieved through the combined efforts of the Institute for Reference Materials and Measurements and the C-PP of the IFCC Scientific Division, representing a model of fruitful and synergistic collaboration. The past and present C-PP members sincerely hope that efforts will continue at all levels to improve the harmonization of serum protein assays in order to offer clinicians reliable tools for improving patient care.

Acknowledgements

We would like to thank Thomas Ledue (Foundation for Blood Research, USA) and Sharon Burr (College for American Pathologists, USA) for their aid in obtaining the CAP data and permission for their use; Dr. Rolf Kruse (DGKL-RfB, Germany) for permission to use the German data; and Dr. David Bullock for supplying the UK NEQAS data from 2009. We thank Heinz Schimmel, Anja Kessler and Mauro Panteghini for thoughtful discussions. Partly supported by EURAMY project (European Community's Sixth Framework Program); Fondazione Cariplo (Milano, Italy); Nobel Project, Transcriptomics and Proteomics Approaches to Diseases of High Sociomedical Impact: A Technology-Integrated Network; and Ricerca Finalizzata Malattie Rare, Ministero della Salute, Istituto Superiore di Sanità (526D/63) and Ministry of University and Research (MIUR) (PRIN 2007AE8FX2_003).

Conflict of interest statement

Authors' conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

References

1. Panteghini M. Traceability as a unique tool to improve standardization in laboratory medicine. *Clin Biochem* 2009;42:236–40.
2. Rowe DS, Anderson SG, Grab B. A research standard for human serum immunoglobulins IgG, IgA and IgM. *Bull World Health Org* 1970;42:535–52.
3. Rowe DS, Grab B, Anderson SG. An international reference preparation for human serum immunoglobulins G, A and M: content of immunoglobulins by weight. *Bull World Health Org* 1972;46:67–79.

4. Reimer CB, Smith SF, Hannon WH, Ritchie RF, van Es L, Becker W, et al. Progress towards international reference standards for human serum proteins. *J Biol Stand* 1978;6:133–58.
5. Reimer CB, Smith SJ, Wells TW. Collaborative calibration of three established reference preparations for specific proteins in human sera as secondary standards for IgA, IgM, and IgG. *J Biol Stand* 1981;9:393–400.
6. Whicher JT, Hunt J, Perry DE, Hobbs JR, Fifield R, Keyser J, et al. Method-specific variations in the calibration of a new immunoglobulin standard suitable for use in nephelometric techniques. *Clin Chem* 1978;24:531–5.
7. Chambers RE, Bullock DG, Whicher JT. Improved between-laboratory agreement for specific protein assays in serum following introduction of a common reference preparation (SPS-01) demonstrated in an external quality assessment scheme. *Clin Chim Acta* 1987;164:189–200.
8. Bullock DG, Dumont G, Vassault A, Aguzzi F, Chambers RE, Milford Ward A, et al. Immunochemical assays of serum proteins: a European external quality assessment survey and the effects of calibration procedures on interlaboratory agreement. *Clin Chim Acta* 1990;187:21–35.
9. Whicher JT, Ritchie RF, Johnson AM, Baudner S, Bienvenu J, Blirup-Jensen S, et al. New international reference preparation for proteins in human serum (RPPHS). *Clin Chem* 1994;40:934–8.
10. Baudner S, Bienvenu J, Blirup-Jensen S, Carlström A, Johnson AM, Milford Ward A, et al. The certification of a matrix reference material for immunochemical measurement of 14 human serum proteins CRM470EUR 15423 EN and 16882 N: European Community, Luxembourg, 1993.
11. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J, et al. International staging system for multiple myeloma. *J Clin Oncol* 2005;23:3412–20.
12. Zegers I, Schreiber W, Linstead S, Lammers M, McCusker M, Munoz A, et al. Development and preparation of a new serum protein reference material: feasibility studies and processing. *Clin Chem Lab Med* 2010;48:805–13.
13. Zegers I, Schreiber W, Munoz A, Sheldon J, Merlini G, Itoh Y, et al. Certification of proteins in the human serum reference material ERM-DA470k/IFCC EUR 23431 EN, European Communities, Luxembourg, 2008, ISBN 978-92-79-094903. In: European Communities, ed., Vol. Luxembourg, 2008:1–54.
14. Zegers I, Charout-Got J, Rzychon M, Trapmann S, Emons H, Schimmel H, et al. Certification of C-reactive protein in reference material ERM-DA472/IFCC EUR 23756 EN, European Communities, Luxembourg, 2009, ISBN 978-92-79-11326-0. In: Communities E, ed., Vol. Luxembourg, 2009:1–28.
15. ISO. ISO Guide 34: General requirements for the competence of reference material producers. 2009.
16. Zegers I, Keller T, Schreiber W, Sheldon J, Albertini R, Merlini G, et al. Characterisation of the new serum protein reference material ERM-DA470k/IFCC: value assignment by immunoassay. *Clin Chem* 2010; Oct 5. [Epub ahead of print].
17. Blirup-Jensen S, Johnson AM, Larsen M. Protein standardization IV: Value transfer procedure for the assignment of serum protein values from a reference preparation to a target material. *Clin Chem Lab Med* 2001;39:1110–22.
18. Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement. *Clin Chem Lab Med* 2010;48:7–10.
19. VIM. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). JCGM 200:2008. Joint Committee for Guides in Metrology. http://www.bipm.org/utls/common/documents/jcgm/JCGM_200_2008.pdf.
20. Blirup-Jensen S. Protein standardization I: protein purification procedure for the purification of human prealbumin, orosomucoid and transferrin as primary protein preparations. *Clin Chem Lab Med* 2001;39:1076–89.
21. Blirup-Jensen S, Grubb A, Lindstrom V, Schmidt C, Althaus H. Standardization of Cystatin C: development of primary and secondary reference preparations. *Scand J Clin Lab Invest Suppl* 2008;241:67–70.
22. Blirup-Jensen S. Protein standardization II: dry mass determination procedure for the determination of the dry mass of a pure protein preparation. *Clin Chem Lab Med* 2001;39:1090–7.
23. Blirup-Jensen S. Protein standardization III: method optimization basic principles for quantitative determination of human serum proteins on automated instruments based on turbidimetry or nephelometry. *Clin Chem Lab Med* 2001;39:1098–109.
24. Blirup-Jensen S, Johnson AM, Larsen M. Protein standardization V: value transfer. A practical protocol for the assignment of serum protein values from a Reference Material to a Target Material. *Clin Chem Lab Med* 2008;46:1470–9.
25. Ledue TB, Johnson AM, Cohen LA, Ritchie RF. Evaluation of proficiency survey results for serum immunoglobulins following the introduction of a new international reference material for human serum proteins. *Clin Chem* 1998;44:878–9.
26. Johnson AM, Whicher JT, Ledue TB, Carlstrom A, Itoh Y, Petersen PH. Effect of a new international reference preparation for proteins in human serum (certified reference material 470) on results of the College of American Pathologists Surveys for plasma proteins. *Arch Pathol Lab Med* 2000;124:1496–501.
27. Goodall SR. Advances in plasma protein standardization. *Ann Clin Biochem* 1997;34:582–7.
28. Ledue TB, Johnson AM. Commutability of serum protein values: persisting bias among manufacturers using values assigned from the certified reference material 470 (CRM 470) in the United States. *Clin Chem Lab Med* 2001;39:1129–33.
29. Johnson AM, Whicher JT. Effect of certified reference material 470 (CRM 470) on national quality assurance programs for serum proteins in Europe. *Clin Chem Lab Med* 2001;39:1123–8.
30. Beetham R, White P, Riches P, Bullock D, MacKenzie F. Use of CRM 470/RPPHS has not achieved true consensus for ceruloplasmin measurement. *Clin Chem* 2002;48:2293–4.
31. Infusino I, Valente C, Dolci A, Panteghini M. Standardization of ceruloplasmin measurements is still an issue despite the availability of a common reference material. *Anal Bioanal Chem* 2010;397:521–5.
32. Bioanalytik RF. http://www.dgkl-rfb.de/4daction/rv_dialog_E/00000000000000000000.