

Toward Worldwide Heparin Assay Harmonization: Identification of a Commutable Secondary Reference Material

Lisa N. van der Vorm,¹ Jan C.M. Hendriks,² Coby M. Laarakkers,^{1,3} Siem Klaver,¹ Andrew E. Armitage,⁴ Alison Bamberg,⁵ Anneke J. Geurts-Moespot,¹ Domenico Girelli,⁶ Matthias Herkert,⁷ Outi Itkonen,⁸ Robert J. Konrad,⁹ Naohisa Tomosugi,¹⁰ Mark Westerman,¹¹ Sukhvinder S. Bansal,¹² Natascia Campostrini,⁶ Hal Drakesmith,⁴ Marianne Fillet,¹³ Gordana Olbina,¹¹ Sant-Rayn Pasricha,⁴ Kelly R. Pitts,⁵ John H. Sloan,⁹ Franco Tagliaro,¹⁴ Cas W. Weykamp,¹⁵ and Dorine W. Swinkels^{1,3*}

BACKGROUND: Absolute plasma hepcidin concentrations measured by various procedures differ substantially, complicating interpretation of results and rendering reference intervals method dependent. We investigated the degree of equivalence achievable by harmonization and the identification of a commutable secondary reference material to accomplish this goal.

METHODS: We applied technical procedures to achieve harmonization developed by the Consortium for Harmonization of Clinical Laboratory Results. Eleven plasma hepcidin measurement procedures (5 mass spectrometry based and 6 immunochemical based) quantified native individual plasma samples ($n = 32$) and native plasma pools ($n = 8$) to assess analytical performance and current and achievable equivalence. In addition, 8 types of candidate reference materials (3 concentrations each, $n = 24$) were assessed for their suitability, most notably in terms of commutability, to serve as secondary reference material.

RESULTS: Absolute hepcidin values and reproducibility (intrameasurement procedure CVs 2.9%–8.7%) differed substantially between measurement procedures, but all were linear and correlated well. The current equivalence (intermeasurement procedure CV 28.6%) between the methods was mainly attributable to differences in calibration and could thus be improved by harmonization with a common calibrator. Linear regression analysis and standardized residuals showed that a candidate

reference material consisting of native lyophilized plasma with cryolyoprotectant was commutable for all measurement procedures. Mathematically simulated harmonization with this calibrator resulted in a maximum achievable equivalence of 7.7%.

CONCLUSIONS: The secondary reference material identified in this study has the potential to substantially improve equivalence between hepcidin measurement procedures and contributes to the establishment of a traceability chain that will ultimately allow standardization of hepcidin measurement results.

© 2016 American Association for Clinical Chemistry

Bioactive hepcidin is a 25-amino-acid hepatic peptide hormone controlling physiological iron homeostasis (1, 2) by causing degradation of the iron-exporter ferroportin, thereby inhibiting absorption of dietary iron from the gut and iron release from storage sites (3). As a key player in iron metabolism, hepcidin holds great promise as the target of therapy and biomarker for diagnosis and monitoring of iron disorders (2). Hepcidin concentrations may be used for diagnosis of iron-refractory iron deficiency anemia (4, 5), differentiation between iron deficiency anemia and anemia of chronic disease (2, 6, 7), guidance in safe iron supplementation (8), and as a companion diagnostic in novel therapies (9, 10).

¹ Department of Laboratory Medicine and ² Department of Health Evidence, Radboud University Medical Center, Nijmegen, the Netherlands; ³ Heparinanalysis.com, Nijmegen, the Netherlands; ⁴ MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK, and Blood Theme, NIHR Oxford Biomedical Research Centre, Oxford, UK; ⁵ Corgenix Medical Corporation, Broomfield, CO; ⁶ Department of Medicine, University of Verona, Verona, Italy; ⁷ DRG Instruments, Marburg, Germany; ⁸ Helsinki University Central Hospital, Laboratory Division HUSLAB, Helsinki, Finland; ⁹ Eli Lilly and Company, Indianapolis, IN; ¹⁰ Division of Advanced Medicine, Medical Research Institute, Kanazawa Medical University, Ishikawa, Japan; ¹¹ Intrinsic LifeSciences, La Jolla, CA; ¹² Institute of Pharmaceutical Sciences, King's College London, London, UK; ¹³ Department of Analytical Pharmaceutical Chemistry, Institute of Pharmacy, University of Liège, Liège, Belgium; ¹⁴ Department of Diagnostics and

Public Health, University of Verona, Italy; ¹⁵ Department of Clinical Chemistry, Queen Beatrix Hospital, Winterswijk, the Netherlands.

* Address correspondence to this author at: Department of Laboratory Medicine, Translational Metabolic Laboratory (830), Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands. Fax +31 (0)24-3668754; e-mail dorine.swinkels@radboudumc.nl.

Previous presentations of the work: Poster presentation, Biolron, September 6–10, 2015, Hangzhou, China. Oral presentation, European Iron Club (EIC), April 7–10, 2016, Innsbruck, Austria.

Received March 8, 2016; accepted April 7, 2016.

Previously published online at DOI: 10.1373/clinchem.2016.256768

© 2016 American Association for Clinical Chemistry

Two prior round robin studies found large differences in absolute hepcidin concentrations between measurement procedures (11, 12). This renders reference intervals and decision limits measurement procedure dependent, complicating interpretation of results (11). To facilitate the use of hepcidin in clinical practice and research, the equivalence between measurement procedures needs to be improved. Ideally, measurement procedures would be standardized, i.e., a primary reference measurement procedure and material are available, allowing traceability to the internationally recognized International System of Units (13). Although various synthetic hepcidin standards are commercially available, several do not contain 100% hepcidin and may not be in their native circulating (folding) state (14, 15), precluding their use as a higher order primary reference material (RM)¹⁶ (11). Also, no primary reference measurement procedure has been defined at present. In this study we aimed for a lower category of metrological traceability: harmonization. Harmonization is accomplished when results are equivalent by being traceable to a secondary RM (i.e., calibrator) to which a value is assigned by a consensus approach, such as the mean of all or multiple methods (16). To this end, we applied technical procedures toward harmonization as recently developed by the International Consortium for Harmonization of Clinical Laboratory Results (ICHCLR) (17). The current study comprised assessment of (a) the analytical performance of participating measurement procedures, (b) the current and achievable degree of equivalence between measurement procedures, (c) the suitability of 8 different candidate RMs as potential calibrators, and (d) mathematically simulated success of harmonization with the selected candidate RM, expressed as achievable equivalence and placed in the context of the total allowable error (TEa).

Methods

STUDY DESIGN OVERVIEW

The current study represented the “assessment” part of the integrated harmonization protocol of the ICHCLR (17). To determine if harmonization was technically achievable, analytical performance of hepcidin measurement procedures in terms of reproducibility and linearity was assessed by triplicate measurements of 32 individual blinded native human plasma samples and of 3 samples that served as a linearity panel. Commutability of candi-

date RMs was evaluated from results for the same 32 individual samples combined with results of the candidate RMs themselves. To determine their stability, candidate RMs were stored at 2–5 different temperatures and measured periodically for 6 months. Finally, we mathematically simulated harmonization with the identified commutable candidate RM to evaluate the potential effect on the equivalence of measurement procedures.

INCLUDED MEASUREMENT PROCEDURES

Methods were selected to represent a variety of methodological characteristics, i.e., analytical platforms, extraction procedures, and standards. Ten laboratories, housing 11 worldwide leading hepcidin measurement procedures [5 mass spectrometry (MS), 6 immunochemical (IC)], agreed to participate in the current study. Relevant measurement procedure characteristics are presented in Table 1.

PLASMA COLLECTION

We obtained heparin plasma samples covering a broad range of clinically observed hepcidin concentrations [<0.5 nmol/L to >50 nmol/L, determined by MS method 1 (MS-1, Table 1)] from 32 individuals. High (≥ 8.0 nmol/L), medium (3.1–7.9 nmol/L), and low (≤ 3.0 nmol/L) hepcidin samples were collected from patients and healthy volunteers and treated as described in the Data Supplement that accompanies the online version of this article at <http://www.clinchem.org/content/vol62/issue7>. After centrifugation or homogenization, the plasma was pipetted into (a) a 12-mL polypropylene tube for sample set preparation and (b) a 0.5-mL polypropylene cryovial for hepcidin measurement by MS-1 to verify classification as “low,” “medium,” or “high” hepcidin. Blank plasma for the candidate RMs spiked with synthetic hepcidin-25 was collected from phlebotomy of an iron depleted juvenile (HJV)-hemochromatosis patient at the St. Antonius Hospital (Utrecht, the Netherlands) and divided into aliquots. All samples were stored at -80 °C until the day of measurement (0.5-mL cryovials) or sample set preparation (12-mL tubes and blank plasma aliquots).

ETHICS APPROVAL

Before blood collection, all volunteers and patients with hemochromatosis provided informed consent. Intensive care patient samples were anonymized immediately after collection of leftover material. Use of plasma from patients and healthy volunteers for this study conformed to the code for proper secondary use of human tissue in the Netherlands and the Declaration of Helsinki, respectively. The study was approved by the Ethics Committee of the Radboudumc, Nijmegen, the Netherlands.

¹⁶ Nonstandard abbreviations: RM, reference material; ICHCLR, International Consortium for Harmonization of Clinical Laboratory Results; TEa, total allowable error; MS, mass spectrometry; IC, immunochemical; CLP, cryolyoprotectant.

Table 1. Methodological characteristics of participating plasma hepcidin MPs.

ID	MP ^a	Extraction	Standard	Manufacturer standard	Reference
MS-1	MALDI-TOF MS	WCX	Heavy-isotope labeled synthetic hepcidin-25	Peptide Inst. (JP)	(18)
MS-3	LC-MS/MS	Reversed phase	Heavy-isotope labeled synthetic hepcidin-25	Peptide Inst. (JP)	(27)
MS-4	SELDI-TOF MS	IMAC	Synthetic hepcidin-24	Peptide Int. (US)	(28) ^b
MS-5	LC-MS/MS	Reversed phase	Heavy-isotope labeled synthetic hepcidin-25	Peptide Int. (US)	(29) ^c
MS-6	LC-MS/MS	HLB	Heavy-isotope labeled synthetic hepcidin-25	Peptide Inst. (JP)	(22)
IC-2	cELISA	None	Synthetic hepcidin-25	Peptide Inst. (JP)	(30)
IC-3	Dual mAb sELISA	None	Synthetic hepcidin-25	Peptide Inst. (JP)	(31)
IC-4	cELISA	None	Synthetic hepcidin-25	Bachem	(32, 33) ^d
IC-5	cELISA	None	Synthetic hepcidin-25	Peptide Int. (US)	(34)
IC-6	cELISA	None	Synthetic hepcidin-25	Bachem	(35) ^e
IC-7	Direct CL ELISA	None	Synthetic hepcidin-25	Bachem	TBP

^a MP, measurement procedure; MS, mass spectrometry-based MP; IC, immunochemical-based MP; c, competitive; s, sandwich; mAb, monoclonal antibody; CL, chemiluminescence; WCX, weak-cation exchange; IMAC, immobilized metal affinity capture; HLB, hydrophilic lipophilic balanced reversed phase; TBP, to be published in the near future.

^b MS-4 was initially included in the study and raw data are available in Supplemental Table 1, but was left out of further analyses and will not be included in future harmonization efforts because this MP will soon no longer be available at that laboratory and will be replaced by MS-5. MPs with IDs MS-2, MS-7, and IC-1 are missing because they were not ready to measure the samples at the time of sample send out.

^c Adapted with minor modifications by the same research group of MS-4. Peptide Inst. (JP), Peptide Institute, Osaka, Japan; Peptide Int (US), Peptide International, Louisville, KY.

^d Commercial assay from Bachem (Hepcidin-25 EIA Kit cat no. S1337, lot nos. A14999 and A15103, Torrance, CA, USA). The lab executing method IC-4 deviated from the kit protocol by incorporating a 9-point, 2-fold instead of a 6-point, 4-fold serial dilution standard curve and by using batch-prepared frozen aliquots of antiserum, standards, and biotinylated-hepcidin tracer stocks prepared to cover the entire set of plates instead of reconstituting these reagents separately for each plate.

^e Commercial assay from DRG (Hepcidin 25 bioactive HS, EIA-5782, batch no. 314K045, Marburg, Germany).

SAMPLE SET PREPARATION AND MEASUREMENTS

Within 6 weeks after collection, plasma samples were transported on dry ice to the Queen Beatrix Hospital (Winterswijk, the Netherlands) where sample sets were prepared the same day (9 sets containing 250 μ L samples, 2 sets containing 600 μ L samples for MS-5 and MS-6; see online Supplemental Table 1). In total, 64 samples were prepared for each measurement procedure. All were blinded and randomized, i.e., low, intermediate and high hepcidin samples were distributed randomly over sample IDs. In addition to the 32 samples collected from individuals, 5 samples comprised pools composed of the individual samples in different proportions to serve as a linearity panel (L1–L5, containing 100/0, 75/25, 50/50, 25/75, and 0/100 proportions of mixtures of the 8 lowest and highest hepcidin samples, respectively). The remaining 24 samples were 8 different candidate RMs to study the effect of 3 variables on commutability, namely (*a*) type of hepcidin, either native plasma (pooled to 3 levels: “low,” “medium,” and “high”), or blank plasma spiked with synthetic human hepcidin-25; (*b*) preservation, either freezing or lyophilization (see online Supplemental Table 2); (*c*) addition of cryoprotectant (CLP), that will be separately patented by Radboudumc. In addition,

a selection of high hepcidin (mean 15.0 nmol/L, assigned by MS-1) native plasma samples were pooled to prepare 148 aliquots of all 4 types of native candidate RMs for long-term stability testing.

Cryovials were placed at -80 °C and the samples to be lyophilized in the freeze-dryer (Zirbus Technology) for 63 h. Two weeks after lyophilization, sample sets were shipped on dry ice to the participating laboratories, where they arrived in frozen condition. Laboratories were requested to prepare samples according to provided instructions and to measure triplicates of each sample within 1 run. All laboratories reported results within 4 weeks after receipt of the samples.

STABILITY STUDY

Stability of the native candidate RMs was studied by comparing the hepcidin concentration in aliquots stored at -80 °C, -20 °C, 4 °C, 20 °C, and 37 °C. These measurements took place at baseline and subsequently after 1 week, 1, 3, and 6 month(s) of storage and will be continued for a maximum of 5 years or until the hepcidin concentration is below the lower limit of detection of MS-1 (0.5 nmol/L). Measurements were only performed by MS-1 for practical feasibility and because of its well-

studied measurement procedure performance. The intraassay and interassay CVs of this measurement procedure were 2.1–3.5% and 4.6–8.3%, respectively (18).

Statistics

EXCLUSION OF OUTLIER MEASUREMENTS AND RESULTS

BELOW THE LOWER LIMIT OF DETECTION

Outlier measurements (maximum 2 per method; see online Supplemental Table 3) were left out of the analyses. Furthermore, only samples with hepcidin values measurable by all methods (above the lower limit of detection) were included in the calculation of CVs.

ANALYTICAL PERFORMANCE OF MEASUREMENT PROCEDURES

On the basis of the logarithmically transformed results of triplicate measurements of the 32 individual samples, reproducibility of measurement procedures was determined as the intrameasurement procedure CV (%) and mutual correlations were calculated as the Spearman rank coefficient between each combination of 2 measurement procedures. Linearity was assessed from the linearity panel results plotted against the mixture ratios (see online Supplemental Table 3, L samples), and expressed as the intercept a , slope b , and Pearson correlation coefficient (r) with its CV (%).

EQUIVALENCE BETWEEN MEASUREMENT PROCEDURES AND COMMUTABILITY OF CANDIDATE RMs

Linear regression between results of each measurement procedure (y axis) and the mean of all measurement procedures (x axis) was performed to explore the current degree of equivalence, summarized in terms of the Spearman rank correlation coefficient r , slope a , and intercept b of the ordinary least squares (OLS) regression line. Current and achievable equivalence were quantified as the mean intermeasurement procedure CV (%), calculated from the means of the logarithmically transformed results of the 32 individual samples. We statistically assessed commutability of the candidate RMs according to the Clinical and Laboratory Standards Institute (CLSI) C53-A protocol (19). To this end, we calculated standardized residuals to the OLS regression line for each of the candidate RM samples, which were defined as non-commutable when they exceeded -3 or 3 . Finally, the effect of harmonization with the identified commutable candidate RM on the equivalence between measurement procedures was mathematically simulated by value reassignment on the basis of the regression of the candidate RM results per measurement procedure against the mean result of all measurement procedures. Details of the value reassignment process are described in the online Supplemental Materials file.

To place the achievable equivalence of measurement procedures in the context of biological variation in hepcidin concentrations, we calculated limits for total allowable error with $TEa\% = f1(0.25 CV_I) + f2(CV_I^2 + CV_G^2)^{1/2}$ in which CV_I is the between-day intraindividual CV, CV_G is the interindividual variation and $f1$ and $f2$ are factors for optimum (0.25 and 0.125), desirable (0.5 and 0.250), and minimum (0.75 and 0.375) TEa (20). Values for CV_I and CV_G were derived from Murphy et al. (21) and are the mean CVs over 5 individuals and over women and men, respectively.

SOFTWARE

All statistical analyses were performed with IBM SPSS statistics software version 20.0 (IBM Corp.).

Results and Discussion

ASSESSMENT OF HARMONIZATION POTENTIAL

Absolute mean hepcidin concentrations measured in the 64 samples differed considerably between measurement procedures, illustrated by a mean factor 7 difference between IC-3 and IC-5 (see online Supplemental Table 3). This confirmed the current lack of equivalence observed in our previous interlaboratory comparison studies (11, 12).

Analytical performance of the 11 hepcidin measurement procedures is presented in Table 2. Reproducibility, expressed as intrameasurement procedure CV, varied widely, ranging from 2.9% for MS-6 to 8.7% for IC-4. Contrary to the large absolute differences, results of all measurement procedures correlated well, as shown by the Spearman rank correlations ranging from 0.96–1.00 for all pair-wise combinations (see online Supplemental Table 4). These high correlations suggested calibration, not heterogeneity of the measurand, as the major cause of the current lack of equivalence. Pearson correlation coefficients showed that all measurement procedures had a linear response, which was a prerequisite characteristic for harmonization (Table 2). Hence, we concluded that the comparable analytical performance of the 10 measurement procedures, in terms of reproducibility, linearity, and mutual correlations, allowed harmonization of their results.

CURRENT EQUIVALENCE BETWEEN MEASUREMENT PROCEDURES

The current degree of equivalence is represented by the linear regression relationship between results of each of the measurement procedures and the mean results of all 10 measurement procedures (Table 2). Again, correlation coefficients indicated a linear relationship for all measurement procedures, facilitating harmonization. Slopes of regression lines were different from 1.00 for 9 out of 10 measurement procedures, indicating that dif-

Table 2. Assessment of analytical characteristics of hepcidin measurement procedures: reproducibility, linearity, and current equivalence of measurement procedures.

ID	Reproducibility		Regression relationship compared to all MPs mean		
	Intra-MP CV, % ^a	r ^b	r ^c	Slope (95% CI)	Intercept (95% CI)
MS-1	5.0	1.000	0.971	0.598 (0.537–0.659)	1.31 (–1.37 to 1.63)
MS-3	3.6	1.000	0.968	0.861 (0.771–0.952)	–0.48 (–2.66 to 1.70)
MS-5	8.2	0.991	0.969	1.054 (0.941–1.167)	–1.17 (–3.99 to 1.65)
MS-6	2.9	0.993	0.934	0.547 (0.489–0.658)	0.25 (–1.71 to 2.22)
IC-2	5.0	0.999	0.992	0.764 (0.726–0.802)	0.01 (–0.89 to 0.92)
IC-3	4.4	1.000	0.975	0.446 (0.408–0.483)	–0.31 (–1.16 to 0.54)
IC-4	8.7	0.989	0.997	1.569 (1.523–1.614)	–0.10 (–1.95 to 1.85)
IC-5	7.7	1.000	0.997	3.184 (3.086–3.281)	–0.85 (–2.91 to 1.21)
IC-6	5.9	0.994	0.991	0.711 (0.674–0.748)	0.73 (–0.05 to 1.52)
IC-7	5.7	0.997	0.956	0.744 (0.650–0.838)	0.37 (–1.94 to 2.68)

^a Intrameasurement procedure CV calculated from the SD of the logarithmically transformed results (SD_i) of the 32 native individual samples (“S” in online Supplemental Table 1).

^b Linearity as assessed by the linearity panel samples (“L” in online Supplemental Table 1), expressed as Pearson correlation coefficient *r*.

^c Pearson correlation coefficient *r* of the regression relationship between results of each individual measurement procedures with the all measurement procedures mean.

MP, measurement procedure; MS, mass spectrometry-based measurement procedure; IC, immunochemical-based measurement procedure.

ferences in absolute measurement results could be decreased by harmonization. The current equivalence, expressed as intermeasurement procedure CV (%), was 28.6%. Contributions of imprecision and nonlinearity were relatively small, 10.8% and 6.7%, respectively (see online Supplemental Table 5). Nonetheless, optimization of these 2 aspects may well improve the equivalence between measurement procedures. Importantly, 75% of the intermeasurement procedure CV could be attributed to differences in calibration and could thus be improved by harmonization. The remaining 7.7% of variation was due to heterogeneity in the measurand. The measurement procedures differed in the degree to which they measured smaller hepcidin-20, -22, and -24 isoforms, which were potentially included in hepcidin concentrations measured some IC but not by MS measurement procedures (2). These isoforms mainly, but not exclusively, occur in blood in disease states associated with increased hepcidin-25 concentrations (2). Hence, they might have been present in the high-hepcidin samples in this study. Moreover, previous studies have shown that prolonged (>1 day) exposure to room temperature results in ex vivo processing of hepcidin-25 to its smaller isoforms (18, 22). However, because samples were only at room temperature for approximately 8 h during sample set preparation, the contribution to isoform formation was likely negligible. Furthermore, it is unclear whether measurement procedures measure free hepcidin or hepcidin bound to α 2-macroglobulin, which might contribute to heterogeneity to an unknown extent (23–

25). Regardless of the cause of heterogeneity, it can only be reduced by assay modification, not by harmonization. Therefore, the maximum achievable equivalence was estimated to be the 8.0% intermeasurement procedure CV that remained from heterogeneity after optimization of all other factors. All in all, these results indicate that harmonization can potentially greatly improve the equivalence between these 10 measurement procedures.

IDENTIFICATION OF A COMMUTABLE, STABLE

SECONDARY RM

Next, we assessed candidate RMs for their suitability to achieve this 75% reduction in intermeasurement procedure CV. The major requirement for a secondary RM is commutability. Two approaches were applied to determine commutability: 1 based on statistics and 1 based on “suitability for purpose,” as discussed under Simulated Harmonization.

For the statistical commutability-assessment, standardized residuals to the regression line for each of the candidate RM samples were calculated, which were accepted as commutable when not exceeding -3 or 3 . Fig. 1 presents a summary of this analysis. Online Supplemental Table 6 presents the exact range (minimum–maximum) of the 45 standardized residuals per candidate RM sample. All candidate RMs were commutable for at least some measurement procedures, but standardized residuals of most candidate RMs varied widely over all measurement procedures. Candidate RMs based on native plasma were generally

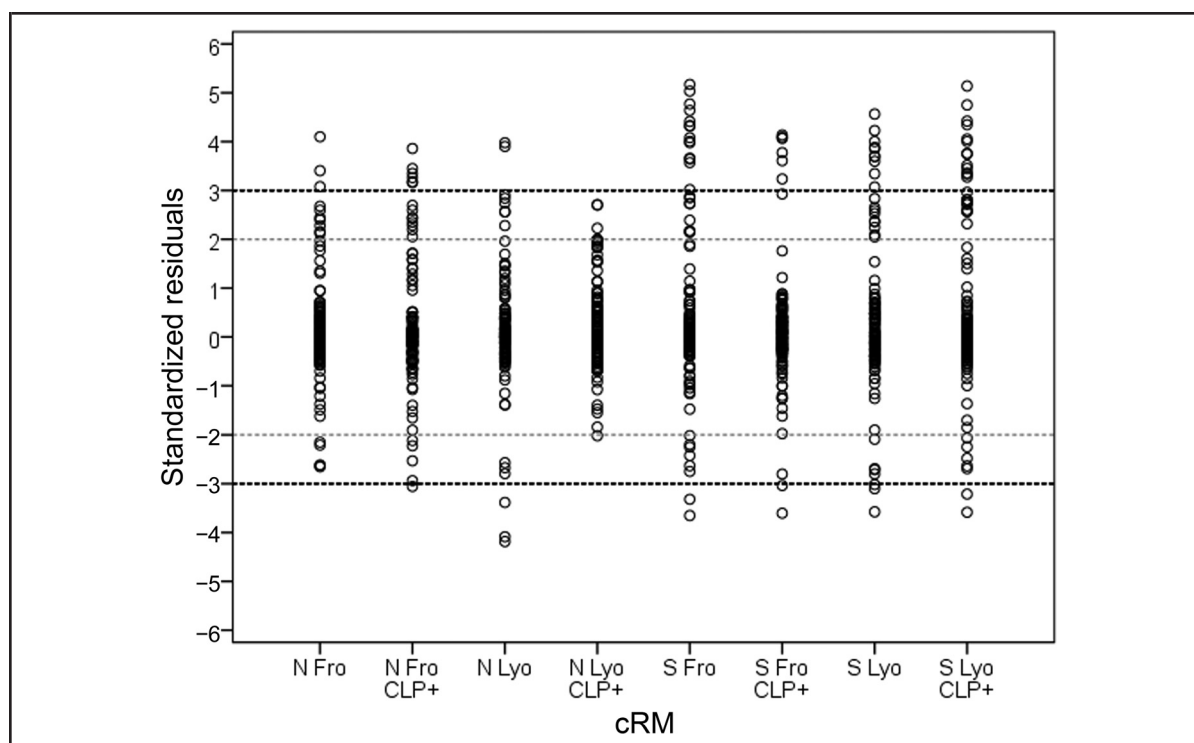


Fig. 1. Standardized residuals of the candidate RM samples.

Circles represent standardized residuals relative to the Passing–Bablok linear regression line of each candidate RM sample in 45 Bland–Altman plots for all pairwise measurement procedure combinations (4 exemplary plots are presented in online Supplemental Fig. 1).

Horizontal dashed lines represent CLSI C53-A acceptance criteria for commutability. A candidate RM was considered not commutable when the standardized residual exceeded -3 or 3 . N, native; S, synthetic; Fro, frozen; Lyo, lyophilized; CLP+, with cryolyoprotectant.

more commutable than those based on blank plasma spiked with synthetic hepcidin. This corroborates our previous round robin study on hepcidin, in which samples spiked with synthetic hepcidin-25 from 2 different manufacturers were also noncommutable with native samples (11). Several explanations for the different behavior of synthetic and native hepcidin were previously proposed, for instance, matrix effects, the presence of a nonnative analyte other than hepcidin-25, and inadequate specificity of some measurement procedures for the synthetic hepcidin-25 due to the heterogeneity.

In addition, our data show that addition of CLP improves commutability, most notably for native candidate RMs. This is in line with results from a previous external quality assessment study for lysosomal enzymes measurement procedures, which demonstrated that addition of CLP improved the stability of lyophilized RM (26). This beneficial effect may be attributed to protection of native hepcidin against denaturation during freezing/thawing, lyophilization, and aging. The effect of CLP was less pronounced or nearly

absent for the synthetic frozen and lyophilized candidate RMs, respectively. Most importantly, native lyophilized plasma with CLP was commutable for all 10 measurement procedures. For this candidate RM 97% of standardized residuals were between -2 and 2 , fulfilling the most stringent acceptance criterion for commutability in the CLSI C53-A protocol. Therefore, we concluded that this candidate RM is suitable as a secondary RM for harmonization of these 10 hepcidin measurement procedures.

A second important characteristic of a secondary RM is stability. This was investigated by periodically measuring hepcidin concentrations in aliquots of the candidate RMs stored for 6 months at temperatures ranging from -80 °C to 37 °C. With the exception of the native lyophilized candidate RM without CLP, no substantial decrease in hepcidin concentrations occurred for all candidate RMs at any of the tested temperatures (see online Supplemental Table 7). Of particular importance for this harmonization effort, the selected lyophilized candidate RM with CLP was shown to fulfill the requirement of robust stability.

Table 3. Mathematically simulated harmonization of hepcidin measurement procedures.

Equivalence (intermeasurement procedure CV), %			
Current			28.6
Achievable ^a			8.0
Simulated achievable equivalence ^b after calibration with each candidate RM			
Native	Fro	CLP –	14.7
		CLP +	13.2
	Lyo	CLP –	8.1
		CLP +	7.7
Synthetic	Fro	CLP –	14.8
		CLP +	12.7
	Lyo	CLP –	12.2
		CLP +	11.8

^a Achievable equivalence determined as the residual intermeasurement procedure CV from heterogeneity/nonspecificity of measurement procedures, which can only be reduced by assay modification, not by harmonization.

^b Achievable equivalence calculated as the intermeasurement procedure CV (%) after value reassignment of the 32 individual samples using the OLS regression equations of the respective candidate RM samples vs the mean of all measurement procedures for these samples. Fro, frozen; Lyo, lyophilized; CLP, cryoprotectant; –, absent; +, present.

SIMULATED HARMONIZATION

The effect of harmonization was mathematically simulated to assess whether the statistically commutable candidate RM is also commutable on the basis of the approach of suitability for purpose. This means that a candidate RM is designated commutable if the intermeasurement procedure CVs of individual patient samples are within the maximum allowable clinically meaningful limit after harmonization, using that RM as a calibrator. In the case of hepcidin, a clinically meaningful limit for the intermeasurement procedure CV has not been established, being a hormone whose concentrations can only currently be interpreted in the context of other laboratory values. In general, intermeasurement procedure CVs of individual samples would ideally be $\leq 10\%$. Table 3 presents the simulated mean intermeasurement procedure CVs of the 32 individual samples for each candidate RM. Simulated harmonization verified that calibration with the native lyophilized plasma with CLP lead to the greatest improvement of equivalence between measurement procedures, namely a mean intermeasurement procedure CV of 7.7% for 97% of the individual samples. Interestingly, calibration with the native lyophilized candidate RM without CLP was predicted to accomplish a mean intermeasurement procedure CV of 8.1%, and would thus also be considered commutable

on the basis of this suitability for purpose approach. Of note, optimization of imprecision and nonlinearity of the measurement procedures is required to accomplish these degrees of equivalence.

Another approach to assess whether the achieved intermeasurement procedure CV is adequate for the biological variation uses the criterion of TEa, combining bias and imprecision. Based on an intraindividual CV of 48.8% and interindividual CV of 154.1% (21), TEa of 23.3% (optimum), 46.5% (desirable), and 69.8% (minimum) were calculated. Before harmonization, results of few measurement procedures were within the optimum TEa limits and results of IC-5 were completely outside the minimum TEa limits (Fig. 2, left panel). Simulated results after harmonization (Fig. 2, right panel) with the native lyophilized calibrator with CLP were almost all within the limits for minimum TEa. Except for 2 of the IC-5 results, all were within the limits for desirable TEa and the majority within the window for optimum TEa, illustrating a major improvement compared to the current equivalence.

PRACTICAL REQUIREMENTS OF THE IDENTIFIED SECONDARY RM

Besides commutability and stability, the integrated harmonization protocol of the ICHCLR addresses the aspects of availability, sustainability, costs, and value assignment. The former 2 aspects represent a possible point of concern. In the current study, heparin plasma was chosen over serum as the matrix for the calibrators, mainly because plasma could be easily obtained from leftover material and phlebotomies. However, 2 out of 3 participating LC-MS/MS measurement procedures reported cloudiness in some of the undiluted samples (2% of samples for MS-5, 7% of samples for MS-6) that blocked the cartridges of the solid phase extraction plates. Thus, some LC-MS/MS systems experienced difficulties in processing (twice freeze-thawed) heparin plasma-based candidate RMs. However, because this phenomenon mainly occurred for high-hepcidin individual samples and not for the identified commutable candidate RM samples for any of the measurement procedures, we are confident to proceed with the harmonization effort focused on this plasma-based candidate RM.

Regarding costs, a major advantage of lyophilized material is that expensive shipment on dry ice is not required for its stability.

For the purpose of “true” value assignment of the identified secondary RM, a reference measurement procedure is required. In the absence of a reference measurement procedure for hepcidin, the current study used the mean of all measurement procedures in the analyses of current and achievable equivalence.

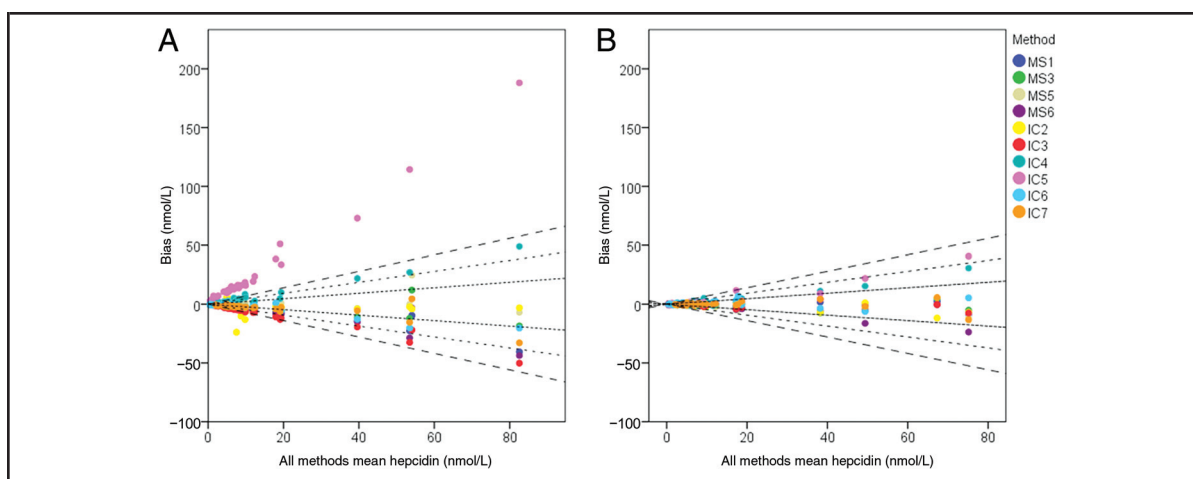


Fig. 2. Achievable equivalence expressed as TEa (A) before and (B) after calibration with the native lyophilized candidate RM with CLP.

The y axis shows the bias between the results of individual samples of each measurement procedure and the all measurement procedures' mean. The all measurement procedures' mean is on the x axis. Data points are mean results of 32 individual samples. Dashed lines represent limits for optimum (most narrow interval), desirable and minimum TEa. MS, mass spectrometry-based measurement procedure; IC, immunochemical-based measurement procedure. Color version of figure is available online.

Conclusions and Future of the Heparin Harmonization Process

The current harmonization study demonstrates that (a) harmonization of measurement procedures with good analytical performance in terms of reproducibility, linearity, and mutual correlations is technically possible, (b) native, lyophilized plasma with CLP is a commutable candidate RM with robust stability and (c) calibration of the measurement procedures with this candidate RM can achieve a reduction of the intermeasurement procedure CV from 28.6% to 7.7%.

The next step on the road toward harmonization is a second sample set send-out comprising native individual plasma samples as well as the native lyophilized candidate RM with CLP. The objectives will be (a) to prove commutability of a new, larger batch of the selected candidate RM, (b) to confirm the success of harmonization by this calibrator and (c) to make the calibrator available to third parties. If this proves successful, the candidate RM will be made available to third parties.

We anticipate this harmonization process will not only result in equivalent results of measurement procedures within the next few years, but should also be considered a stepping stone toward full standardization of hepcidin measurement procedures. To achieve this, a traceability chain minimally consisting of a primary RM (pure native hepcidin), a primary reference measurement procedure and secondary RM needs to be set up. Thus, the relevance of the current study goes beyond its poten-

tial to obtain equivalent results for all measurement procedures, because it also contributes to the future traceability chain by the identification of a potential secondary RM.

Ultimately, we believe the current and all future harmonization and standardization efforts will greatly impact the success of hepcidin as a biomarker by paving the way for (a) definition of generally accepted and usable reference intervals and decision limits, (b) application of consistent clinical decision limits for medical care and best practice guidelines and (c) pooling and comparison of data from various studies to facilitate medical research and its translation to the clinic.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: C.M. Laarakkers, hepcidinanalysis.com and Radboudumc; S. Klaver, hepcidinanalysis.com and Radboudumc; M. Herkert, DRG Instruments; RJ Konrad, Eli Lilly and Company; M. Westerman, Intrinsic LifeSciences; G. Olbina, Intrinsic LifeSciences; D.W. Swinkels, hepcidinanalysis.com and Radboudumc.

Consultant or Advisory Role: None declared.

Stock Ownership: R.J. Konrad, Eli Lilly and Company; M. Westerman, Intrinsic LifeSciences; G. Olbina, Intrinsic LifeSciences; K.R. Pitts, Corgenix, Inc.; J.H. Sloan, Eli Lilly and Company.

Honoraria: None declared.

Research Funding: None declared.

Expert Testimony: None declared.

Patents: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, and final approval of manuscript.

Acknowledgments: The authors thank all laboratories participating in this Hepcidin Harmonization Study. Furthermore, we thank Liesbeth Schröer-Janssen and all laboratory technicians of the MCA laboratory of the Queen Beatrix Hospital, Winterswijk, the Netherlands, for their assistance with the sample set preparation.

References

- Ganz T. Hepcidin and iron regulation, 10 years later. *Blood* 2011;117:4425-33.
- Kroot JJ, Tjalsma H, Fleming RE, Swinkels DW. Hepcidin in human iron disorders: diagnostic implications. *Clin Chem* 2011;57:1650-69.
- Nemeth E, Tuttle MS, Powelson J, Vaughn MB, Donovan A, Ward DM, et al. Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science* 2004;306:2090-3.
- De Falco L, Silvestri L, Kannengiesser C, Moran E, Oudin C, Rausa M, et al. Functional and clinical impact of novel TMPRSS6 variants in iron-refractory iron-deficiency anemia patients and genotype-phenotype studies. *Hum Mutat* 2014;35:1321-9.
- Donker AE, Raymakers RA, Vlasveld LT, van Barneveld T, Terink R, Dors N, et al. Practice guidelines for the diagnosis and management of microcytic anemias due to genetic disorders of iron metabolism or heme synthesis. *Blood* 2014;123:3873-86.
- Theurl I, Aigner E, Theurl M, Nairz M, Seifert M, Schroll A, et al. Regulation of iron homeostasis in anemia of chronic disease and iron deficiency anemia: diagnostic and therapeutic implications. *Blood* 2009;113:5277-86.
- van Santen S, van Dongen-Lases EC, de Vegt F, Laarakkers CM, van Riel PL, van Ede AE, Swinkels DW. Hepcidin and hemoglobin content parameters in the diagnosis of iron deficiency in rheumatoid arthritis patients with anemia. *Arthritis Rheum* 2011;63:3672-80.
- Ganz T. Hepcidin and the global burden of iron deficiency. *Clin Chem* 2015;61:577-8.
- Camaschella C. Treating iron overload. *N Engl J Med* 2013;368:2325-7.
- Ruchala P, Nemeth E. The pathophysiology and pharmacology of hepcidin. *Trends Pharmacol Sci* 2014;35:155-61.
- Kroot JJ, van Herwaarden AE, Tjalsma H, Jansen RT, Hendriks JC, Swinkels DW. Second round robin for plasma hepcidin methods: first steps toward harmonization. *Am J Hematol* 2012;87:977-83.
- Kroot JJ, Kemna EH, Bansal SS, Busbridge M, Camprostrini N, Girelli D, et al. Results of the first international round robin for the quantification of urinary and plasma hepcidin assays: need for standardization. *Haematologica* 2009;94:1748-52.
- ISO. In vitro diagnostic medical devices—measurement of quantities in biological samples—metrological traceability of values assigned to calibrators and control materials—first edition. Geneva: ISO; 2003. ISO 17511:2003.
- Clark RJ, Tan CC, Preza GC, Nemeth E, Ganz T, Craik DJ. Understanding the structure/activity relationships of the iron regulatory peptide hepcidin. *Chem Biol* 2011;18:336-43.
- Jordan JB, Poppe L, Haniu M, Arvedson T, Syed R, Li V, et al. Hepcidin revisited, disulfide connectivity, dynamics, and structure. *J Biol Chem* 2009;284:24155-67.
- Miller WG, Myers GL, Gantzer ML, Kahn SE, Schönbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem* 2013;59:1291-3.
- Weykamp C, Eckfeldt J, Roberts W, Vesper H, Thienpont L, Burns C, et al. International Consortium for Harmonization of Clinical Laboratory Results. Toolbox of technical procedures to be considered when developing a process to achieve harmonization for a measurand. www.harmonization.net (Accessed December 2015).
- Laarakkers CM, Wiegerinck ET, Klaver S, Kolodziejczyk M, Gille H, Hohlbaum AM, et al. Improved mass spectrometry assay for plasma hepcidin: detection and characterization of a novel hepcidin isoform. *PLoS One* 2013;8:e75518.
- CLSI. Characterization and qualification of commutable reference materials for laboratory medicine: approved guideline. Wayne, PA: CLSI; 2010. CLSI document EP30-A (formerly C53-A).
- Fraser CG, Hyltoft Peterson P, Libeer JC, Ricos C. Proposal for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem* 1997;34:8-12.
- Murphy AT, Witcher DR, Luan P, Wroblewski VJ. Quantitation of hepcidin from human and mouse serum using liquid chromatography tandem mass spectrometry. *Blood* 2007;110:1048-54.
- Itkonen O, Parkkinen J, Stenman UH, Hämäläinen E. Preanalytical factors and reference intervals for serum hepcidin LC-MS/MS method. *Clin Chim Acta* 2012;413:696-701.
- Huang ML, Austin CJ, Sari MA, Rahmanto YS, Ponka P, Vyoral D, Richardson DR. Hepcidin bound to α 2-macroglobulin reduces ferroportin-1 expression and enhances its activity at reducing serum iron levels. *J Biol Chem* 2013;288:25450-65.
- Itkonen O, Stenman UH, Parkkinen J, Soliymani R, Baumann M, Hämäläinen E. Binding of hepcidin to plasma proteins. *Clin Chem* 2012;58:1159-60.
- Peslova G, Petrak J, Kuzelova K, Hrdy I, Halada P, Kuchel PW, et al. Hepcidin, the hormone of iron metabolism, is bound specifically to alpha-2-macroglobulin in blood. *Blood* 2009;113:6225-36.
- Ruijter GJ, Boer M, Weykamp CW, de Vries R, van den Berg I, Janssens-Puister J, et al. External quality assurance programme for enzymatic analysis of lysosomal storage diseases: a pilot study. *J Inher Metab Dis* 2005;28:979-90.
- Murao N, Ishigai M, Yasuno H, Shimomura Y, Aso Y. Simple and sensitive quantification of bioactive peptides in biological matrices using liquid chromatography/selected reaction monitoring mass spectrometry coupled with trichloroacetic acid clean-up. *Rapid Commun Mass Spectrom* 2007;21:4033-8.
- Camprostrini N, Castagna A, Zaninotto F, Bedogna V, Tessitore N, Poli A, et al. Evaluation of hepcidin isoforms in hemodialysis patients by a proteomic approach based on SELDI-TOF MS. *J Biomed Biotechnol* 2010;2010:329646.
- Wolff F, Deleers M, Melot C, Gulbis B, Cotton F. Hepcidin-25: Measurement by LC-MS/MS in serum and urine, reference ranges and urinary fractional excretion. *Clin Chim Acta* 2013;23:99-104.
- Kroot JJ, Laarakkers CM, Geurts-Moespot AJ, Grebenchtchikov N, Pickkers P, van Ede AE, et al. Immunochemical and mass-spectrometry-based serum hepcidin assays for iron metabolism disorders. *Clin Chem* 2010;56:1570-9.
- Butterfield AM, Luan P, Witcher DR, Manetta J, Murphy AT, Wroblewski VJ, Konrad RJ. A dual-monoclonal sandwich ELISA specific for hepcidin-25. *Clin Chem* 2010;56:1725-32.
- Pasricha SR, Atkinson SH, Armitage AE, Khandwala S, Veenemans J, Cox SE, et al. Expression of the iron hormone hepcidin distinguishes different types of anemia in African children. *Sci Transl Med* 2014;6:235re3.
- Prentice AM, Doherty CP, Abrams SA, Cox SE, Atkinson SH, Verhoef H, et al. Hepcidin is the major predictor of erythrocyte iron incorporation in anemic African children. *Blood* 2012;119:1922-8.
- Ganz T, Olbina G, Girelli D, Nemeth E, Westerman M. Immunoassay for human serum hepcidin. *Blood* 2008;112:4292-7.
- Zipperer E, Post JG, Herkert M, Kündgen A, Fox F, Haas R, et al. Serum hepcidin measured with an improved ELISA correlates with parameters of iron metabolism in patients with myelodysplastic syndrome. *Ann Hematol* 2013;92:1617-23.