

Judith Taibon*, Tobias Santner, Neeraj Singh, Sara Cheikh Ibrahim, Galina Babitzki, Daniel Köppl, Alexander Gaudl, Andrea Geistanger, Uta Ceglarek, Manfred Rauh and Christian Geletneky

An isotope dilution-liquid chromatography-tandem mass spectrometry (ID-LC-MS/MS)-based candidate reference measurement procedure (RMP) for the quantification of aldosterone in human serum and plasma

<https://doi.org/10.1515/cclm-2022-0996>

Received October 5, 2022; accepted March 3, 2023;

published online March 24, 2023

Abstract

Objectives: An isotope dilution-liquid chromatography-tandem mass spectrometry (ID-LC MS/MS)-based candidate reference measurement procedure (RMP) for aldosterone quantification in human serum and plasma is presented.

Methods: The material used in this RMP was characterized by quantitative nuclear magnetic resonance (qNMR) to assure traceability to SI Units. For liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis a two-dimensional heart cut LC approach, in combination with an optimal supported liquid extraction protocol, was established for the accurate analysis of aldosterone in human serum and plasma in order to minimize matrix effects and avoid the co-elution of interferences. Assay validation was performed according to current guidelines. Selectivity and specificity were assessed using spiked serum; potential matrix effects were examined by a post column infusion experiment and the comparison of standard line slopes. An extensive protocol over 5 days was applied to determine precision, accuracy and trueness. Measurement uncertainty was evaluated according to the Guide to the Expression of Uncertainty in Measurement (GUM), for which three

individual sample preparations were performed on at least two different days.

Results: The RMP allowed aldosterone quantification within the range of 20–1,200 pg/mL without interference from structurally-related compounds and no evidence of matrix effects. Intermediate precision was $\leq 4.7\%$ and repeatability was 2.8–3.7% for all analyte concentrations. The bias ranged between –2.2 and 0.5% for all levels and matrices. Total measurement uncertainties for target value assignment (n=6) were found to be $\leq 2.3\%$; expanded uncertainties were $\leq 4.6\%$ (k=2) for all levels.

Conclusions: The RMP showed high analytical performance for aldosterone quantification in human serum and plasma. The traceability to SI units was established by qNMR content determination of aldosterone, which was utilized for direct calibration of the RMP. Thus, this candidate RMP is suitable for routine assay standardization and evaluation of clinical samples.

Keywords: aldosterone; isotope dilution-liquid chromatography-tandem mass spectrometry; qNMR; reference measurement procedure; SI units; traceability.

Introduction

Aldosterone is a mineralocorticoid steroid hormone that is produced in the zona glomerulosa of the adrenal cortex, and regulated primarily by the renin-angiotensin system through angiotensin II [1–4]. Aldosterone plays an important role in controlling blood volume and blood pressure, and promotes reabsorption of sodium and water and excretion of potassium from the renal tubules [1–4]. Concentration of aldosterone in serum or plasma is routinely measured in the diagnosis of diseases associated with hyperaldosteronism and hypoaldosteronism [3]; therefore, accurate measurement of aldosterone is important for reliable and timely diagnosis.

*Corresponding author: Judith Taibon, PhD, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany, Phone: +49 8856 6012941, E-mail: judith.taibon@roche.com

Tobias Santner, Neeraj Singh, Sara Cheikh Ibrahim, Galina Babitzki, Andrea Geistanger and Christian Geletneky, Roche Diagnostics GmbH, Penzberg, Germany

Daniel Köppl and Manfred Rauh, Clinic for Children and Adolescents, University Hospital Erlangen, Erlangen, Germany

Alexander Gaudl and Uta Ceglarek, Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital Leipzig, Leipzig, Germany

As a consequence a reference measurement system is required to improve measurement accuracy and reduce variability between laboratories, which generally entails a reference measurement procedure (RMP) and higher order reference materials. The latter are usually provided by National Measurement Institutes [NMI] [2, 5]. The Joint Committee on Traceability in Laboratory Medicine (JCTLM) is responsible for maintaining a database of current higher order reference materials (International Organization for Standardization [ISO] 15194), procedures (ISO 15193) and services (ISO 15195) [5, 6]. Additionally, the JCTLM together with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) hosts the External Quality Assessment scheme for Reference Laboratories in Laboratory Medicine (RELA scheme), which offers a range of analytes including aldosterone and enables reference laboratories and candidate reference laboratories to demonstrate their competence and assesses the equivalence of different reference measurement systems [7, 8]. In this role, the JCTLM and IFCC promote the development and implementation of ID-LC-MS/MS-based reference methods to improve standardization of all clinical assays [6, 8]. Until now, for aldosterone only isotope dilution gas chromatography based RMPs and services are listed within the JCTLM database [9–11]. Recently, an isotope dilution-liquid chromatography-tandem mass spectrometry (ID-LC-MS/MS)-based candidate RMP for the measurement of aldosterone in human plasma has been reported, however, no International System of Units (SI units) traceable primary reference material was used in this method [2]. There is, therefore, a need for an easy-to-implement candidate RMP that is able to quantify the concentration of aldosterone in human samples that is directly calibrated using a characterized SI traceable reference material. Since, there is no primary reference material available for aldosterone, we have developed an in-house quantitative nuclear magnetic resonance (qNMR) protocol for the target value assignment, as qNMR is one of the optimal methodologies to assign the absolute content to reference materials and is established as a primary ratio method by the NMIs [12]. The powerful structure elucidation characteristics coupled with the linear response to the amount (count) of the analyte and direct traceability to the SI unit (kilogram) via the qNMR standards, provides a unique ability for determination of the amount or ‘counts’ of a particular analyte. Additionally, the highest order qNMR internal standards are traceable directly to the National Institute of Standards and Technology (NIST) benzoic acid 350b (Coulometric) and/or NIST PS1 (benzoic acid; first primary qNMR standard) [13].

In view of this, we aimed to establish a concept for the development and validation of RMPs that meet the

requirements of the guidelines ISO 15193 and ISO 15194. Here, the methodology is described in detail, including all technical information and the calculation of the measurement uncertainty, allowing easy transferability of the methods to a second laboratory. The validation of the method was carried out with a strong focus on the two most important quality parameters: precision and trueness. The calculation of the associated measurement uncertainty is based on the Guide to the Expression of Uncertainty in Measurement (GUM) [14], and is also described in detail herein.

In addition, the RMP must also fulfill pre-determined acceptance criteria for bias (B) and imprecision (coefficient of variation [CV]) based on the biological variation set out by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) database [15]. For aldosterone, the imprecision of the RMP (CV_{ref}) should be half of the imprecision of routine assays (CV_{rou}) and bias of the RMP (B_{ref}) should be one third of the bias of routine assays (B_{rou}) [16]. Based on data from EFLM database for $CV_{rou} \leq 18.3\%$ and $B_{rou} \leq 12.6\%$ [15], the acceptance criteria for the RMP should be defined as $CV_{ref} \leq 9.2\%$ and $B_{ref} \leq 4.2\%$ [16].

In this regard, an ID-LC-MS/MS-based candidate RMP for aldosterone quantification in human serum and plasma is presented, which (i) is metrologically traceable, (ii) allows the true target value assignment, (iii) can be used for the standardization and evaluation of routine tests and (iv) enables the evaluation of patient samples to ensure traceability of the individual patient results.

Materials and methods

A detailed account of the methods, including a full list of materials and equipment used, can be found in the Supplementary Material 1.

Chemicals and reagents

The following chemicals and reagents were used: acetonitrile, methanol and formic acid (all LC-MS grade, Biosolve, Valkenswaard, The Netherlands), ammonium fluoride ($\geq 99.99\%$ trace metals basis), deuterated acetonitrile (CD_3CN) and tecnazene (TraceCert®, Batch Nr. BCBW6288) (all Sigma-Aldrich, Taufkirchen, Germany) and ethyl acetate (Merck, Darmstadt, Germany). Water was purified using a Millipore Milli-Q IQ 7000 system (Merck). Aldosterone (product code A9477, Lot Nr. MKCJ7855) and its deuterated internal standard (ISTD), aldosterone-9,11,12,12- d_4 (100 $\mu\text{g/mL}$ in acetonitrile), were obtained from Sigma Aldrich.

Steroid-free human serum and human serum (multi-individual pooled) were obtained from Roche Diagnostics GmbH (Penzberg, Germany). Human plasma (multi-individual pooled) was obtained from Roche Diagnostics GmbH (Mannheim, Germany) for use as the plasma matrix. The patient samples were anonymized, residual samples, in accordance with the Declaration of Helsinki.

General requirements for laboratory equipment

All equipment was calibrated and certified by the manufacturer. The minimum sample weight for the ultra-microbalance used (XP6U, Mettler Toledo, Columbus, Ohio, USA) was determined according to the United States Pharmacopeial Convention (USP) guidelines (USP Chapters 41 and 1251). Direct displacement pipettes were used to measure organic solvents and serum. Volumetric glassware (Class A volumetric flasks) meeting the requirements of ISO 1042 and USP must be used for the preparation of stock and spike solutions.

qNMR for determining the purity of the standard materials

Single-Pulse- $^1\text{H}\{^{13}\text{C}\}$ NMR (Supplementary Material 2, Supplementary Figure 1) was utilized for the quantitation (Olefinic ^1H ; $\delta=5.62$ ppm) with an inter-scan delay of 70 s. An extensive analysis of aldosterone diastereomers, the quantification signal, NMR parameters and the solvents utilized, was published previously [12]. Details of NMR acquisition and free induction decay (FID) processing parameters can be found in the Supplementary Material 2.

Preparation of calibrators and quality control (QC) samples

For the preparation of matrix-based calibrators, two individual primary stock solutions were prepared and further diluted to produce working and spike solutions. To prepare each stock solution, 5 mg ($\pm 5\%$) of aldosterone was weighed using an ultra-microbalance (XP6U/M, Mettler Toledo, readability of 0.0001 mg) and dissolved in 50 mL acetonitrile in a volumetric flask. The concentration of each primary stock solution was calculated according to the purity of the reference material ($94.7\% \pm 0.2\%$, determined by qNMR). Each primary stock solution was diluted further with acetonitrile to achieve working solutions with a concentration of 1,000 ng/mL. The working solutions were used to prepare eight calibrator spike solutions of different concentrations in 50% methanol. Final matrix-based calibrator levels ranging from 20.0–1,200 pg/mL (conversion factor pg/mL to nmol/L: 0.002774) were prepared with a 1 + 49 dilution (v/v) in steroid-free serum matrix. Three levels of matrix-based QC material were prepared in the same way as the calibrator levels, using a third, independent primary stock solution and calibrator spike solutions. Final concentration levels of the QC material were 60.0, 150 and 1,000 pg/mL, respectively. A detailed, stepwise instruction of calibrator and QC material preparation including information on pipettes, tips and volumetric flasks used is described in Supplementary Material 1. All relevant information for the type B estimation of measurement uncertainty is described in the Supplementary Material 3.

To monitor systematic drifts, two native patient samples close to the medical decision point (MDP: 150 pg/mL) were used to generate a control chart. Acceptable results were within two standard deviations (SDs). All samples (spiked and native material) were stored at -80°C for a maximum of 11 weeks.

Internal standard (ISTD) solution

For preparation of the ISTD stock solution, 20 μL of the 100 $\mu\text{g/mL}$ aldosterone-9,11,12,12- d_4 solution was pipetted into a 20 mL volumetric

flask and filled up to the calibration mark with acetonitrile. Final ISTD solution with a concentration of 200 pg/mL was prepared by pipetting 200 μL of this stock solution into a 10 mL volumetric flask and adding 50% methanol (v/v) up to the calibration mark.

Sample preparation

As sample matrix native serum, plasma (lithium [Li]-heparin plasma, K_2EDTA plasma and K_3EDTA plasma) and steroid-free serum serving as surrogate matrix can be used.

400 μL of sample specimen (either native sample, calibrator, or QC material) was transferred into a 1.5 mL polypropylene tube (Sarstedt, Nümbrecht, Germany) and 40 μL ISTD solution was added. The sample was shaken using an overhead rotation mixer (Sarstedt Sarmix M200) for 30 min at 60 rpm to ensure thorough equilibration; 500 μL of Milli-Q-water was added and the solution was mixed again. As a sample preparation procedure supported liquid extraction (SLE) was executed. Briefly, the mixture was loaded onto a SLE column (Phenomenex SLE Novum, 6cc tube) by applying a vacuum for 5 s. After 10 min equilibration, 2.5 mL ethyl acetate was added and eluted in a glass tube. The pooled fractions were evaporated using a Biotage TurboVap system (Uppsala, Sweden; 50°C , 1.0 bar, 20 min), reconstituted in 100 μL 30% acetonitrile, filtered using a SPIN-X centrifuge tube (Corning Incorporated, Salt Lake City, USA; 15 min, 10,000 rpm) and transferred to a high-performance liquid chromatography (HPLC)-vial with insert (both VWR International GmbH, Darmstadt Germany).

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

An Agilent 1290 Infinity II LC system equipped with two binary pumps, a thermostatted autosampler and a column compartment (all Agilent, Santa Clara, California, USA) was used for chromatographic separation. Analyte detection was performed using an AB Sciex Q-Trap 6500+ mass spectrometer with a Turbo V ion source (all AB Sciex, Framingham, Massachusetts, USA).

Chromatographic separation of aldosterone was achieved using a two-dimensional heart-cut LC approach with a combination of two orthogonal stationary phases and mobile phases to increase peak capacity and selectivity. Detailed operation instructions including all relevant LC and MS parameters as well as the system setup information are described in the Supplementary Material 1.

In brief, 20 μL of the sample was injected and aldosterone was isolated in the first dimension (Waters Acquity BEH C18 1.7 μm , 50×2.1 mm) at a retention time of 2.8 min with a window of 0.17 min. The sample was transferred to the second dimension (Restek Raptor Biphenyl 2.7 μm , 100×2.1 mm) via a 100 μL -loop and aldosterone was eluted at 7.5 min. Column temperatures were kept at 50°C . Mobile phases in the first dimension consisted of Milli-Q-water (A1) and acetonitrile (B1), and 0.2 mM ammonium fluoride in Milli-Q-water (A2) and methanol (B2) in the second dimension. In both dimensions, chromatographic separation was achieved using individual gradient programs over 9 min with flow rates of 0.3 mL/min and 0.5 mL/min, respectively.

Aldosterone was detected in multiple reaction monitoring (MRM) mode using a Sciex QTrap 6500+ mass spectrometer operating in negative electrospray ionization mode. The quantifier transition (aldosterone 359.1 $m/z \rightarrow 189.0$ m/z) serves as basis for the quantitation and is

associated with the corresponding transition of the ISTD (aldosterone- d_4 363.1 $m/z \rightarrow$ 190.0 m/z). The additional qualifier transition (aldosterone 359.1 $m/z \rightarrow$ 331.2 m/z and aldosterone- d_4 363.1 $m/z \rightarrow$ 335.2 m/z) monitored to screen for unknown interferences in clinical samples by application of the “branching ratio concept”.

System suitability test (SST)

To test system performance and assure the long-term stability of the method, an SST was established to examine sensitivity and chromatographic resolution before every sequence. Two levels (SST1 and SST2) were prepared in 30% acetonitrile both containing aldosterone and 18-hydroxycorticosterone, an interference which is not separated in the first dimension and therefore transferred to the second dimension (Figure 1). SST1 and SST2 aldosterone concentrations corresponded to the analyte concentration within the processed calibrator level 1 and 8, respectively.

To pass the SST, the signal-to-noise ratio (S/N) of the quantifier transition was required to be ≥ 100 for SST1 and $\geq 7,000$ for SST2. S/N was calculated using the Analyst software (AB Sciex, Framingham, Massachusetts, USA) as peak height divided by noise. To calculate noise, the software uses the SD (with a mean of zero) of all chromatographic data points between the specified background. A retention time of 7.5 min (± 0.5 min) was required and aldosterone had to be baseline separated from 18-hydroxycorticosterone with a chromatographic resolution of $R_s > 1.5$.

To examine potential carryover, the injection of SST2 was followed by two solvent blanks. The analyte peak area observed in the first blank was required to be $\leq 20\%$ of the analyte peak area of SST1 to pass.

Calibration and structure of analytical series

The assay was calibrated using the calibrators prepared as described in the calibrator and quality control (QC) sample preparation section. Calibrator levels were measured in increasing concentration at the

beginning and the end of the analytical series. The calibration functions were obtained by linear regression of the area ratios of the analyte and the ISTD (y) against the analyte concentration (c_A) was calculated using the following function: $y = ac_A + b$.

The individual sequence setup varied due to specific measurement requirements and the intended use of the RMP (see Supplementary Material 1). If reference values were assigned, the number of sample preparations ($n=x$) was dependent on the desired measurement uncertainty, and samples were measured on at least two different days. If a method comparison study was performed, or complaint samples were measured, samples were prepared ($n=1$) and measured.

Data processing

For processing of the raw data file, the Analyst software (v1.6.2 or 1.7) was used with Intelli Quan as Quantitation Integration Algorithm. For peak integration, a smoothing width of three points, a peak splitting factor of two, and noise percentage of 99% were used. The calibration curve was linear with a $1/x$ weighting; the origin was ignored, and for the response, the area ratio was obtained.

Method validation

Assay validation and determination of measurement uncertainty were performed according to the Clinical & Laboratory Standard Institute Guidelines C62A “Liquid Chromatography-Mass Spectrometry Methods” [17], the International Conference on Harmonization guidance document “Harmonized Tripartite Guideline Validation of Analytical Procedures: Text and Methodology Q2 (R1)” [18] and the GUM [14].

Selectivity/specificity

Selectivity was determined by spiking aldosterone ISTD (aldosterone-9,11,12,12- d_4) and relevant structurally related compounds, including

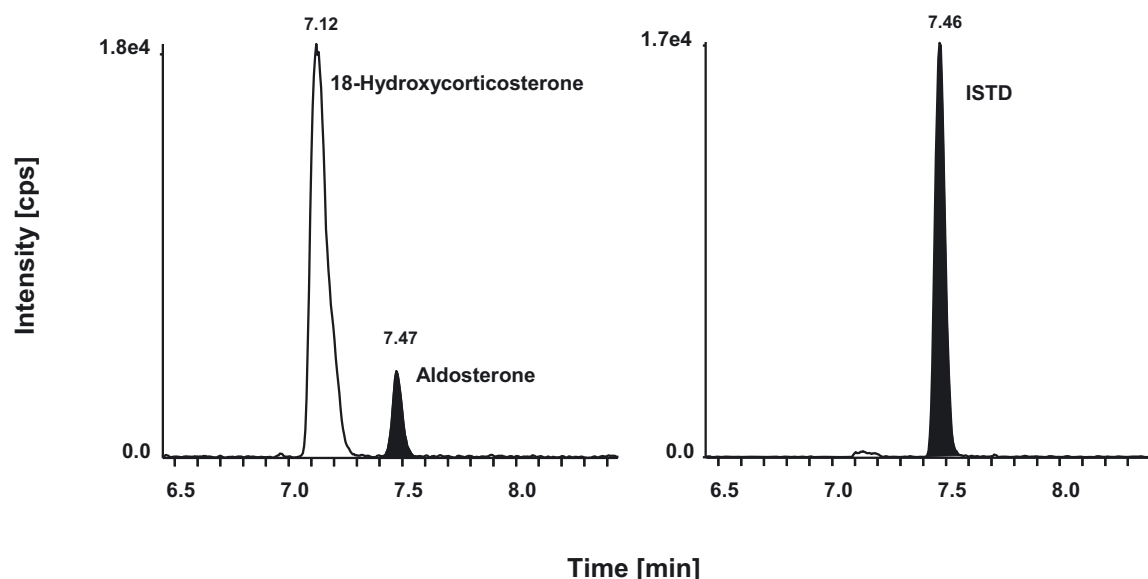


Figure 1: Representative analyte ion trace of the SST1. The interference was baseline separated in the second dimension using a Restek Raptor Biphenyl column (2.7 μm , 100×2.1 mm). The mobile phase used in the second dimension consisted of 0.2 mM ammonium fluoride in Milli-Q-water and methanol, and was chosen as ammonium fluoride enhances the sensitivity of the method.

cortisol, cortisone, prednisone, prednisolone and 18-hydroxycorticosterone, in steroid-free human serum and a native human serum pool. To examine possible interfering matrix signals for the analyte quantifier and qualifier transition, both the steroid-free human serum and the native human serum pool were checked at the expected retention time. In addition the ion ratio between quantifier and qualifier transition (Quant/Qual) in native patient samples (measured within the method comparison study) was compared to the Quant/Qual ratio in neat SST samples (SST1 and SST2) to confirm that no interfering compounds were present. In addition, steroid-free human serum was spiked with deuterated ISTD to evaluate residual unlabeled analyte within the stable isotope-labeled ISTD. The estimation of residual unlabeled analyte was performed by comparing the absolute areas of the analyte signal within the internal standard and the analyte signal of the lowest calibrator level. The amount of unlabeled analyte in the internal standard must not exceed 20% of the amount of the lower limit of the measuring interval (20 pg/mL), corresponding to the concentration level of the lowest calibrator.

Matrix effects

To determine possible matrix effects, a qualitative post-column infusion experiment and a quantitative experiment based on the comparison of standard line slopes were performed.

In the qualitative post-column infusion experiment, a solution of 50 pg/mL aldosterone in 30% acetonitrile was infused at a flow rate of 7 μ L/min via a T-piece into the HPLC post-column eluent prior to entering the MS/MS system to generate a stable analyte background signal. A processed matrix sample was injected and the change in background signal was acquired. The following matrices were analyzed: neat solution (30% acetonitrile), serum matrix pool, steroid-free human serum and plasma (Li-Heparin, K₂EDTA and K₃EDTA). Any change in the selected reaction monitoring analyte signal would indicate a matrix component-mediated effect on the ionization yield of the analytes.

In the quantitative experiment, a comparison of standard line slopes was performed by comparing the following matrices: neat solution (30% acetonitrile), a native human serum pool, steroid-free human serum and Li-Heparin plasma [19]. Since the calibrator samples were prepared by spiking the analyte prior to sample preparation, the peak areas of the analyte and IS reflect a combination of the recovery efficiency and the effect of the matrix on ionization (process efficiency [PE]). Although the absolute peak areas of the analytes at the same concentration may vary in different matrices, the ratios of analyte/IS and the deriving slopes must not be affected. Thus, the slopes of the calibration curves in different matrices are a good measure of the relative matrix effect [19]. Calibrator levels were prepared as described in the section preparation of calibrators and quality control (QC) samples. Neat samples were diluted to the final concentration of processed calibrator levels whereas matrix calibrator levels were prepared by SLE. All samples were measured once and slopes and coefficients of determination were compared.

In addition, matrix effect was evaluated by comparing the MS response (area count ratios of analyte to IS) in all matrices to the MS response of the neat samples. The sample matrix effect (ME) was then calculated with the following equation: $ME\% = B/A \times 100$, where B is the area count ratios of analyte to IS obtained from samples in matrix, and A is the area count ratios in matrix-free samples [20, 21].

Linearity

For the determination of linearity, calibration curves from three individual weighings were prepared as described in the calibrator and quality control (QC) sample preparation section. Calibration range was extended by $\pm 20\%$, and two additional spike solutions were prepared to obtain spiked serum samples with a final concentration of 16.0 pg/mL and 1,440 pg/mL aldosterone. The peak area ratio of analyte to the corresponding ISTD was plotted against the respective analyte concentration (pg/mL). Correlation coefficients and residuals for each curve were determined. The linearity of the method was proven based on the recovery of serially diluted samples (samples 1–11) using the preferred regression model for calculation. Sample 1 was a native patient sample with an analyte concentration close to the lowest calibrator level; sample 11 was a native patient sample near the expected upper measuring range. Samples 1 and 11 were diluted to prepare a set of serially diluted samples, referred to herein as samples 2–10. Using these two samples, nine mixtures were diluted: 9 + 1 v/v, 8 + 2 v/v, 7 + 3 v/v, 6 + 4 v/v, 5 + 5 v/v, 4 + 6 v/v, 3 + 7 v/v, 2 + 8 v/v and 1 + 9 v/v. Recovery was reported as the percentage of recovery of the measured concentration relative to the nominal concentration of the sample pool.

Lower limit of the measuring interval (LLMI) and limit of detection (LOD)

The LLMI was determined using spiked samples (steroid free human serum) at the lowest calibration level (approximately 20.0 pg/mL) in accordance with the performance specifications of RMPs. Samples were prepared six-fold and injected once for the determination of precision and trueness.

The LOD is estimated by using both the measured LoB and test replicates of a sample with a known low concentration of the analyte. As low concentration sample calibrator level 1 (n=10 independent samples) was used. For LOB calculation blank samples (steroid free human serum, n=10 independent samples) were used to calculate the individual mean blank signal and its standard deviation. The mean and SD of the low concentration sample is then calculated according to $LOD = LOB + 1.645 * (SD_{\text{low concentration sample}})$ [22].

Precision, trueness and accuracy

Precision was evaluated by performing a 5-day validation experiment using two individual calibrator preparations for two different measurement sequences (part A and part B) on each day. To estimate total variability of the method (type A uncertainty), between-injection variability, between-preparation variability, between-calibration variability and between-day variability were determined using an ANOVA-based variance-components analysis.

Three spiked serum samples (60.0, 150 and 1,000 pg/mL of aldosterone) covering the measuring range and two native patient samples (approximately 80.0 pg/mL and 200 pg/mL of aldosterone, covering the MDP) were prepared in triplicate for each part (A and B) and injected twice (n=12 measurements per day; n=60 measurements over 5 days). Sample measurement for each part (A and B) was performed in the following sequence: calibrators, spiked samples, native samples and calibrators. Spiked human serum samples (steroid-free serum) were

prepared as described in the calibrator and quality control (QC) sample preparation section using independent stock and spike solutions.

Data evaluation was done using Biowarp, an internal statistics program based on the VCA Roche Open Source software package in R [23]. Alternatively, any commercially available statistics software, e.g., Analyse-it [24] can be used. Repeatability included between-injection variability and between-preparation variability whereas intermediate precision included between-calibration variability and between-day variability. Repeatability and intermediate precision are expressed as SD and CV. As independent calibrator production, including preparation of stock, spike solutions and matrix-based calibrator levels, was not included in this setup, this variability does not correspond to the total variability and is evaluated as type B uncertainty, as described further below.

Accuracy and trueness were assessed using the certified secondary reference material from the National Metrology Institute of Japan (NMIJ, CRM 6402-b) as well as leftover samples from the RELA Scheme from the years 2016 and 2018 and two samples with reference values (93.9 and 437 ng/mL) provided by the Reference Institute of Bioanalytics (RfB). All samples were prepared threefold over two different days. In addition accuracy and trueness were evaluated using three spiked steroid-free human serum samples (approximately 60.0, 150 and 1,000 pg/mL aldosterone), and a native serum pool with low endogenous aldosterone concentration (approximately 34.0 pg/mL) fortified with the same amount of aldosterone. To prove the method was suitable for plasma matrix, the same concentration levels were spiked into Li-Heparin plasma (endogenous aldosterone concentration approximately 35.0 pg/mL). Validity of dilution was performed using two spiked serum samples (steroid-free human serum) at approximately 1,500 and 2,000 pg/mL.

All samples were prepared in triplicate for part A and part B ($n=6$ measurements) on one day. Accuracy was evaluated as closeness of agreement between the test result and the accepted reference value, whereas trueness was evaluated as closeness of agreement between the average value obtained from a series of test results and an accepted reference value. For the spiked native serum pool and Li-Heparin plasma, accuracy and trueness were evaluated related to the final concentration of endogenous and spiked analyte.

Sample stability

The stability of the processed samples at three different concentration levels (60.0, 150 and 1,000 pg/mL aldosterone) on the auto sampler was investigated at 4–8 °C for 10 days. Samples were re-measured on days 3, 4 and 10 after the initial injection. Samples from the precision experiment were used. Recoveries were calculated by comparing the measured value with the nominal concentration ($t=0$ h).

Stability of spiked solutions and matrix-based spiked calibrator and QC material stored at –80 °C were evaluated at three concentration levels (60.0, 150 and 1,000 pg/mL aldosterone) for approximately 12 weeks.

Equivalence of results between independent laboratories

To assess the agreement of the RMP between two independent laboratories (Laboratory 1: Roche Diagnostics GmbH, Penzberg, Germany; Laboratory 2: Clinic for Children and Adolescents, University Hospital Erlangen, Germany), a method comparison study was performed on $n=155$ native residual, anonymized patient samples. The RMP was transferred to Laboratory 2 and the system was set up as described in

the Supplementary Material 1. A 3-day precision experiment was performed at Laboratory 2. Spiked samples were provided by Laboratory 1. Both laboratories prepared their own calibrator levels using aldosterone (characterized by qNMR) as primary reference material.

Adaptations to the sample preparation protocol were made at Laboratory 2 in order to use the 96-well plate format. All steps before the SLE sample preparation were done in one single 96-well deep well plate made from polypropylene. In brief, a 400 μ L sample and 40 μ L ISTD were incubated before adding 360 μ L of water. The diluted sample was mixed and sample volume was split in two parts for extraction. The extraction process was carried out using two 96-well plates (Phenomenex Novum SLE) instead of using one SLE tube per sample. The eluate of both parts was collected in one set of glass inserts (Hirschmann, Eberstadt, Germany). After evaporation, the samples were reconstituted in 100 μ L, mixed, sealed, and measured as described above.

A second method comparison study was performed between Laboratory 1 and Laboratory 3 (the Institute of Laboratory Medicine at Leipzig University Hospital), to demonstrate comparability between the RMP and an established routine LC-MS/MS method using commercially available calibration material. Laboratory 3 analyses were done using the online solid phase extraction LC-MS/MS method published by Gaudl et al. [25].

Uncertainty of measurements

Uncertainty of measurements was determined according to the GUM [14] and considered the following steps: purity of the reference material, weighing of the analyte, preparation of stock, working, spike and calibrator solutions, preparation of internal standard solution, sample preparation of calibrators, measurement of calibrators and generation of the calibration curve, preparation of unknown samples as well as the measurement and evaluation of sample results. The estimation of the uncertainty for the preparation of calibrators (unc_{cal}) was performed as type B evaluation. All other aspects were evaluated in the precision experiment (unc_{prec}) as type A evaluation.

The total measurement uncertainty of the whole approach for a single measurement was estimated as a combination of the uncertainty of calibrator preparation (unc_{cal}) and the calculated uncertainty of the precision experiment (unc_{prec}). For the assignment of reference or target values, multiple sample preparations for each sample were performed on at least two different days and the result was calculated as the arithmetic mean ($n=x$). Total measurement uncertainty was calculated as combined uncertainty of the calibrator preparation (unc_{cal}) and uncertainty (SD) of the mean of measurement results (unc_{mean}). To avoid error underestimation, for each individual sample concentration level, unc_{cal} of the calibrator level with the highest uncertainty was chosen for this combination of uncertainties. The derived total uncertainty was multiplied by a coverage factor of $k=2$, which corresponds to an approximate confidence level of 95% assuming a normal distribution, to obtain an expanded uncertainty. A detailed description on the evaluation of measurement uncertainty is given in Supplementary Material 3.

Results and discussion

Traceability to SI units

The most important parameter for a reference measurement procedure, the traceability to the SI-unit kilogram,

has been established by the utilization of qNMR ISTDs, which are directly traceable to the NIST PS1 (primary qNMR standard) or NIST benzoic acid 350b primary standards. Sixfold experiments involving six individual weighings of the analyte and tecnazene, yielded a final content value of $94.7 \pm 0.2\%$ ($k=1$; Supplementary Material 2).

Specificity/selectivity

To minimize matrix effects and the co-elution of isobaric interferences (both known and unknown), a two-dimensional heart-cut LC approach for the accurate analysis of aldosterone in human serum was chosen. The combination of two orthogonal stationary and mobile phases increased peak capacity. The selective isolation of the compound of interest from the first dimension via a standard 2-switch 10-port-valve allowed measurement of the compound of interest with high specificity.

Using a Waters Acquity BEH C18 column in the first dimension in combination with a mobile phase consisting of Milli-Q-water and acetonitrile allowed the separation of cortisol, cortisone, prednisone and prednisolone, whereas 18-hydroxycorticosterone was baseline separated in the second dimension using a Restek Raptor Biphenyl column (Pennsylvania, USA). The mobile phase used in the second dimension consisted of 0.2 mM ammonium fluoride in Milli-Q-water and methanol and was chosen because ammonium fluoride enhances the sensitivity of the method [26]. Using this two-dimensional approach, no interferences were observed in all tested matrices. The Quant/Qual ratio was within the 20% criterion related to the mean ratio of the native SSTs for all native serum samples measured in the method comparison study. The ISTD was checked for residual unlabeled analyte and showed an amount of 6% of the amount of the lowest calibrator. This amount is far below the maximum allowable 20% and can therefore be neglected.

Matrix effects

Possible ion suppression or enhancement effects from different matrices (neat solution, serum and plasma matrices) were evaluated by performing a qualitative post-column infusion experiment and a quantitative experiment based on the comparison of standard line slopes.

Ion suppression or enhancement in the region of the retention time of aldosterone or its ISTD was not seen in the tested matrices performing the post-column infusion experiment. For the comparison of standard line slopes, slopes were found to be 0.0074 (95% confidence interval [CI]

0.0073–0.0075) for the native serum matrix, 0.0075 (95% CI 0.0074–0.0076) for the neat solution, 0.0071 (95% CI 0.0069–0.0074) for the surrogate serum matrix, and 0.0073 (95% CI 0.0071–0.0075) for the plasma matrix. The CIs of the slopes overlapped, which leads to the assumption that they are not significantly different from each other, supporting the absence of matrix effects. Correlation coefficients were ≥ 0.999 independent of the matrix used for calibration. The evaluation of sample ME % showed no effect in surrogate serum matrix (mean ME 98%) and indicated a slight ion suppression for the native serum matrix as well as in the plasma matrix resulting in a mean ME of 95 and 93%, respectively. However, the comparison of standard line slopes along with our evaluations on the accuracy and precision of the method showed no significant influence from this effect, thereby proving the method to be matrix independent.

Linearity

When assessing linearity, the residuals were randomly and equally distributed in a linear and quadratic regression model, so a linear regression model was chosen for assay calibration. The correlation coefficients were $r \geq 0.999$ for all individual calibration curves.

The linearity of the method was confirmed using serially diluted samples. The measurement results showed a linear dependence with a correlation coefficient of 0.999. The recovery of the diluted samples relative to the expected concentrations ranged from 97–107% for the lowest level.

Lower limit of measuring interval (LLMI) and limit of detection (LOD)

The LLMI was determined using a spiked matrix sample with a concentration of 19.2 pg/mL (Figure 2). The relative deviation ($n=6$) was -1.1% and CV was 3.6%. Thus, the recommended acceptance criteria of precision ($CV \leq 9.2\%$) and bias ($\leq 4.2\%$) were fulfilled. The LOD was estimated as 4.93 pg/mL.

Precision, trueness and accuracy

A 5-days validation experiment was performed to estimate total variability of the RMP. Three levels of spiked samples (low, mid, high) and two native patient samples with concentrations within the therapeutic/reference range were analyzed. Variability components were estimated using an

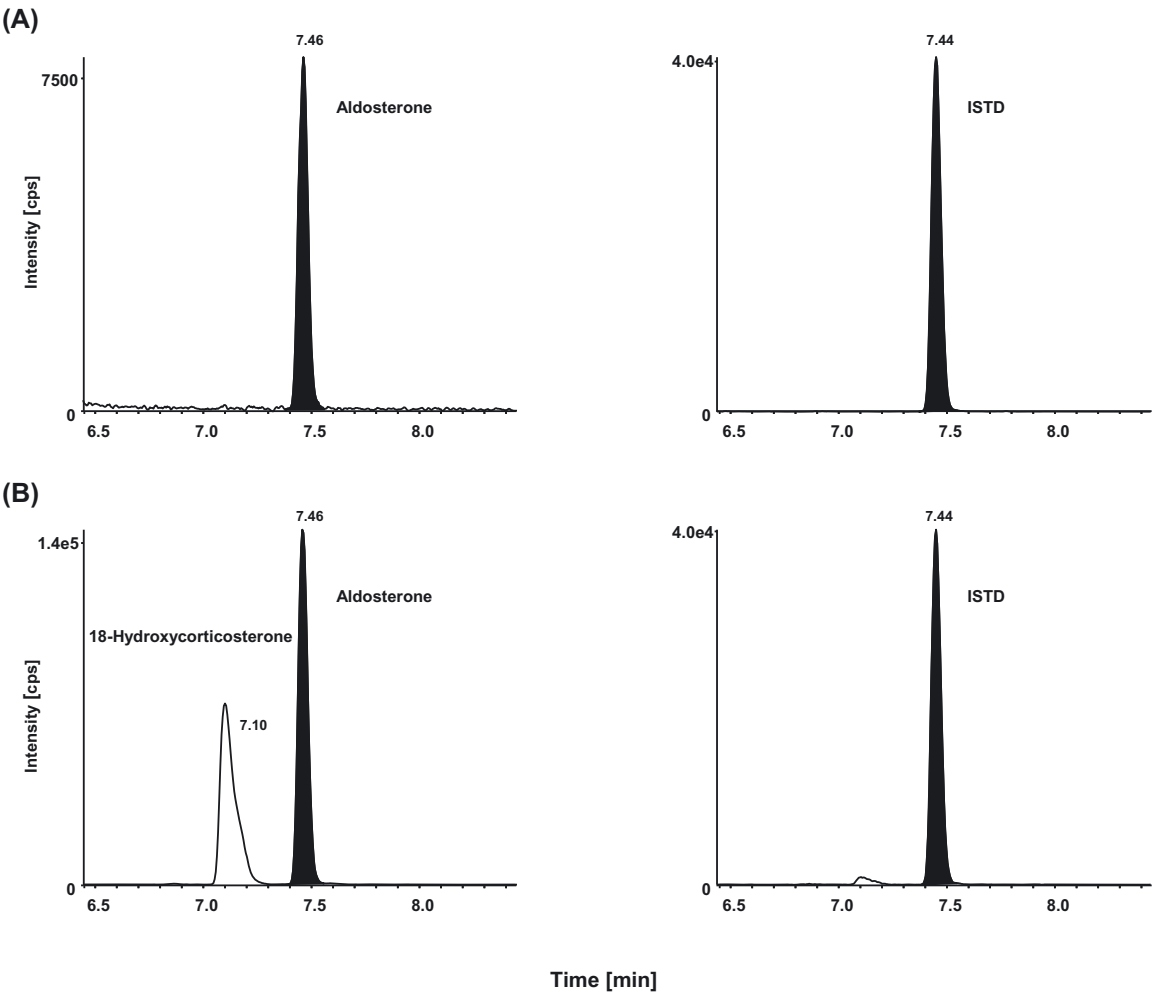


Figure 2: Aldosterone LC-MS/MS derived analytical readouts. (A) Chromatogram of the spiked matrix calibrator level 1 with a concentration of 20 pg/mL, analyte (left) and internal standard (right); (B) chromatogram of a native patient sample with a concentration of 370 pg/mL, analyte (left) and internal standard (right).

ANOVA-based variance component analysis. For a better interpretability, results are given as CV. Repeatability CV was 2.8–3.7% over all concentration levels (Table 1).

Intermediate precision, including variations in between-day -calibration, -preparation and -injection was ≤4.7%. Accuracy and trueness was demonstrated using the certified secondary reference NMIJ CRM 6402-b. All three

Table 1: Detailed precision performance obtained by VCA analysis (n=60 measurements).

Variance source	CV, %				
	60.0 pg/mL	150 pg/mL	1,000 pg/mL	Native patient sample 1	Native patient sample 2
Intermediate	4.7	3.7	3.9	4.2	4.0
Between-day	2.7	0.9	1.5	0.0	1.5
Between-calibration	2.6	1.1	0.5	2.3	0.5
Repeatability	2.8	3.4	3.6	3.6	3.7
Between-preparation	0.0	2.5	2.6	3.0	3.2
Between-injection	2.8	2.4	2.5	2.0	1.9

CV, coefficient of variation; VCA, variance component analysis.

Table 2: External quality control for reference laboratories – RELA scheme.

	Laboratory	Sample A, nmol/L	Expanded uncertainty, nmol/L	Sample B, nmol/L	Expanded uncertainty, nmol/L	Method
Year 2016	A	0.729	0.007	1.332	0.013	ID/LC/MS/MS
	B	0.769	0.017	1.355	0.029	ID/LC/MS/MS
	C	0.755	0.022	1.217	0.031	ID/LC/MS/MS
	Roche	0.776	0.022	1.36	0.039	ID/LC/MS/MS
Year 2018	A	1.697	0.017	1.305	0.013	ID/LC/MS/MS
	B	1.703	0.034	1.322	0.032	ID/LC/MS/MS
	C	1.688	0.08	1.286	0.07	ID/LC/MS/MS
	D	1.73	0.035	1.36	0.03	ID/LC/MS/MS
	E	1.641	0.033	1.257	0.029	ID/LC/MS/MS
	Roche	1.70	0.034	1.34	0.046	ID/LC/MS/MS

Result line printed in bold indicates a JCTLM listed service [8]; conversion factor pg/mL to nmol/L: 0.002774.

levels (197 ± 21 , 383 ± 42 , 760 ± 40 pg/mL) were found with a bias of -1.4% (95% CI -2.6 to -0.1%) for the lowest level, 0.9% (95% CI -1.1 to 2.8%) for the mid-range level and 1.6% (95% CI 0.7 – 2.6%) for the highest level. Accuracy ranged from 97.2 to 104.3% over all levels. Reference samples provided by the Reference Institute for Bioanalytics (RfB) were found with a bias of -0.5% (95% CI -2.5 to 1.4%) at the lower level and 0.7 (95% CI -0.6 to 2.1%) at the higher level, whereas accuracy ranged from 97.4 to 102.8% . Expanded uncertainty over all samples measured in six replicates over two days was found to be between 2.0 and 2.8% ($k=2$). For two levels the bias differs statistically significant from zero. However,

it is within the range to be expected due to the uncertainty in the preparation of independent calibrator samples.

Also leftover samples from the RELA Scheme (years 2016 and 2018) were found in a good agreement with results from listed laboratories, of which the RfB serves as one JCTLM listed service (see Table 2).

In addition, trueness was evaluated using three levels of spiked steroid-free human serum, native serum as well as native Li-Heparin plasma matrix. The bias was -2.2 to 0.5% for all levels and matrices, with the exception of the highest level within the fortified native matrix pool, where the bias was -4.2% (Table 3).

Table 3: Bias evaluation results ($n=6$ measurements).

	Nominal concentration, pg/mL	Mean of measured concentration, pg/mL	Bias evaluation results		
			Bias, %	SD, %	95% CI
Surrogate matrix					
Level 1	57.3	57.6	0.4	3.5	−2.4 to 3.2
Level 2	143	144	0.4	1.2	−0.6 to 1.3
Level 3	956	951	−0.5	1.6	−1.8 to 0.8
Dilution 1	1,447	1,457	0.7	1.6	−0.6 to 2.0
Dilution 2	1,929	1,925	−0.2	1.5	−1.4 to 1.0
Li-Heparin plasma					
Level 1	92.8	90.8	−2.2	5.1	−6.3 to 2.0
Level 2	179	175	−2.1	2.1	−3.8 to −0.4
Level 3	991	971	−2.0	4.7	−5.8 to 1.8
Native serum					
Level 1	91.4	91.9	0.5	1.7	−0.8 to 1.9
Level 2	177	176	−1.0	2.6	−3.1 to 1.0
Level 3	990	948	−4.2	2.4	−6.1 to −2.3

Measurements were performed in triplicate within part A and part B ($n=6$ measurements). The bias and corresponding confidence intervals were calculated using the individual sample biases of $n=6$ preparations.

Dilution integrity was determined using two spiked samples at concentration levels of approximately 1,500 and 2,000 pg/mL, and the mean deviation was between -0.2 and 0.7% . Thus, the method is suitable for the analysis of samples up to a target concentration of 2,000 pg/mL (Table 3).

Overall, the method did not exhibit a statistically significant bias and was matrix independent. The proposed limits for an RMP of 9.2% CV and 4.2% bias were achieved.

Sample stability

Autosampler stability of processed samples at $4-8\text{ }^{\circ}\text{C}$ was determined at three different concentration levels (low, mid, high) and samples were stable for 9 days; relative bias was less than $\pm 2.2\%$ for all concentrations. Stability of neat spike solutions and spiked control samples (low, mid, high) at $-80\text{ }^{\circ}\text{C}$ was determined for 12 weeks and relative bias was less than $\pm 4.8\%$ for all concentrations.

Equivalence of results between independent laboratories

Anonymized native patient samples ($n=155$) were analyzed at Laboratories 1 and 2. In total, $n=60$ samples were either below or above the measurement range and were excluded from the method comparison; $n=2$ samples were identified as outliers and therefore not included.

Analysis of the data using Passing-Bablok showed very good agreement between the two laboratories, resulting in a regression equation with a slope of 1.02 (95% CI $1.01-1.04$) and an intercept of -1.2 (95% CI -2.1 to -0.51) (Figure 3A). Pearson's correlation coefficient was ≥ 0.999 . Bland-Altman analysis showed a mean bias of 0.2% , which is not statistically significant different from zero (95% CI -1.0 to 1.4%). The $\pm 1.96 \cdot \text{SD}$ range of the relative differences was between -11.2 and 11.5% , with a 95% CI interval of 4.1% , with a slightly higher scatter in the very low measuring range (Figure 3B). The performed 3-days precision experiment at Laboratory 2 resulted in very comparable CVs to those of Laboratory 1. The intermediate precision ranged between 1.4% for the highest concentration and 4.8% for the lowest concentration, and repeatability was found to be less than 2.0% , independent of the analyte

concentration. Both data scatter and data bias indicate that the proposed RPM is transferable between independent laboratories.

To demonstrate comparability to an established routine LC-MS/MS method [25], a method comparison study using the same native patient samples was performed. The routine assay was calibrated using a commercially available kit without any information regarding traceability to SI units. In total, $n=66$ native patient samples were included in the analysis.

Passing-Bablok regression yielded a regression equation with a slope of 0.97 (95% CI $0.92-1.02$) and an intercept of 0.65 (95% CI -2.22 to 3.43), excluding two noted outliers (total $n=64$) (Figure 4A). Corresponding Pearson's correlation coefficient was ≥ 0.992 . Bland-Altman analysis plot ($n=64$) showed very good agreement between the established routine assay and the RMP.

The resulted mean bias in the patient cohort was -1.6% (95% CI -4.4 to 1.1%) and does not differ statistically significant from zero. The $\pm 1.96 \cdot \text{SD}$ range of the relative differences was between -23.4 and 20.1% (95% CI interval = 9.5%) (Figure 4B), which is in the range to be expected for routine assays.

Uncertainty of results

The total measurement uncertainty of aldosterone for a single measurement was estimated as a combination of the uncertainty of calibrator preparation (unc_{cal}) and the uncertainty of the precision experiment (unc_{prec}) and was $\leq 4.8\%$ regardless of the concentration level and sample type (Table 4). To obtain an expanded uncertainty, the derived total uncertainty was multiplied by a coverage factor of $k=2$, which corresponds to an approximate confidence level of 95% assuming a normal distribution.

For the assignment of reference or target values, three individual sample preparations for each sample were performed on at least two different days and the result was calculated as the arithmetic mean ($n=6$). The total measurement uncertainty was estimated as a combination of uncertainty of calibrator preparation (unc_{cal}) and uncertainty of the mean of measurement results (unc_{mean}), resulting in total uncertainties $\leq 2.3\%$ and expanded uncertainties $\leq 4.6\%$ (coverage factor of $k=2$) (Table 5).

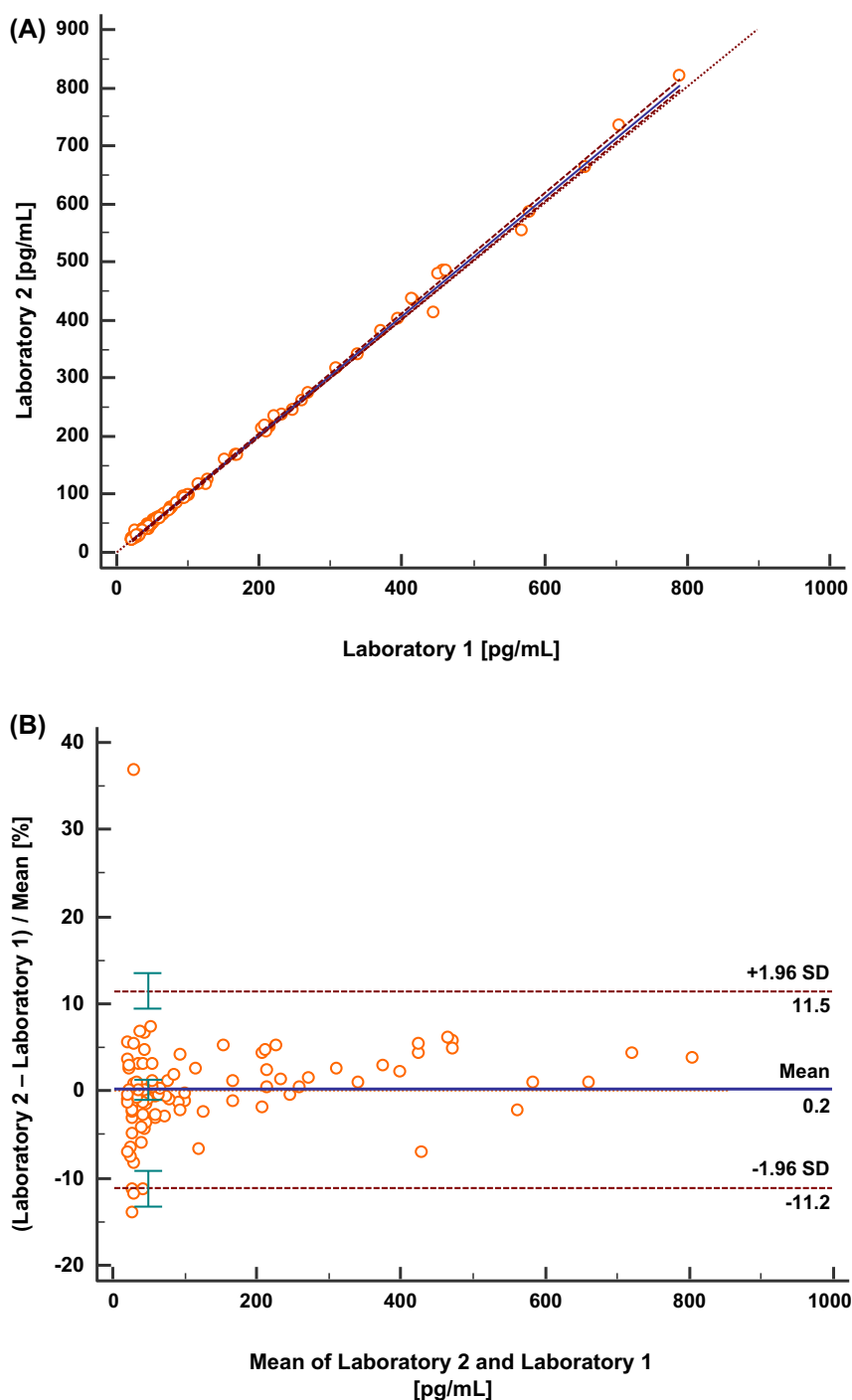


Figure 3: Results from the patient sample-based aldosterone method comparison study performed between two independent laboratories. (A) Passing-Bablok regression plot including the Pearson regression analysis for the method comparison study of the RMP ($n = 93$ patients, $n=1$ sample preparation) between the independent laboratories (Laboratory 1: Roche, Penzberg, and Laboratory 2: Clinic for Children and Adolescents, University Hospital Erlangen). Passing-Bablok regression plot (A); the regression analysis resulted in a regression equation $y = 1.02x - 1.2$, with a 95% confidence interval from 1.01 to 1.04 for the slope and from -2.1 to -0.51 for the intercept. The Pearson's correlation coefficient was ≥ 0.999 . (B) Bland-Altman plot for the method comparison study of the RMP ($n = 93$ patients, $n=1$ sample preparation) between two independent laboratories (Laboratory 1: Roche, Penzberg, and Laboratory 2: Clinic for Children and Adolescents, University Hospital Erlangen). The interlaboratory measurement bias was $+0.2\%$ (95% CI interval from -1.0 to 1.4%) and the 2S interval of the relative difference was 11.4% (95% CI interval = 4.1%). CI, confidence interval; RMP, reference measurement procedure; SD, standard deviation.

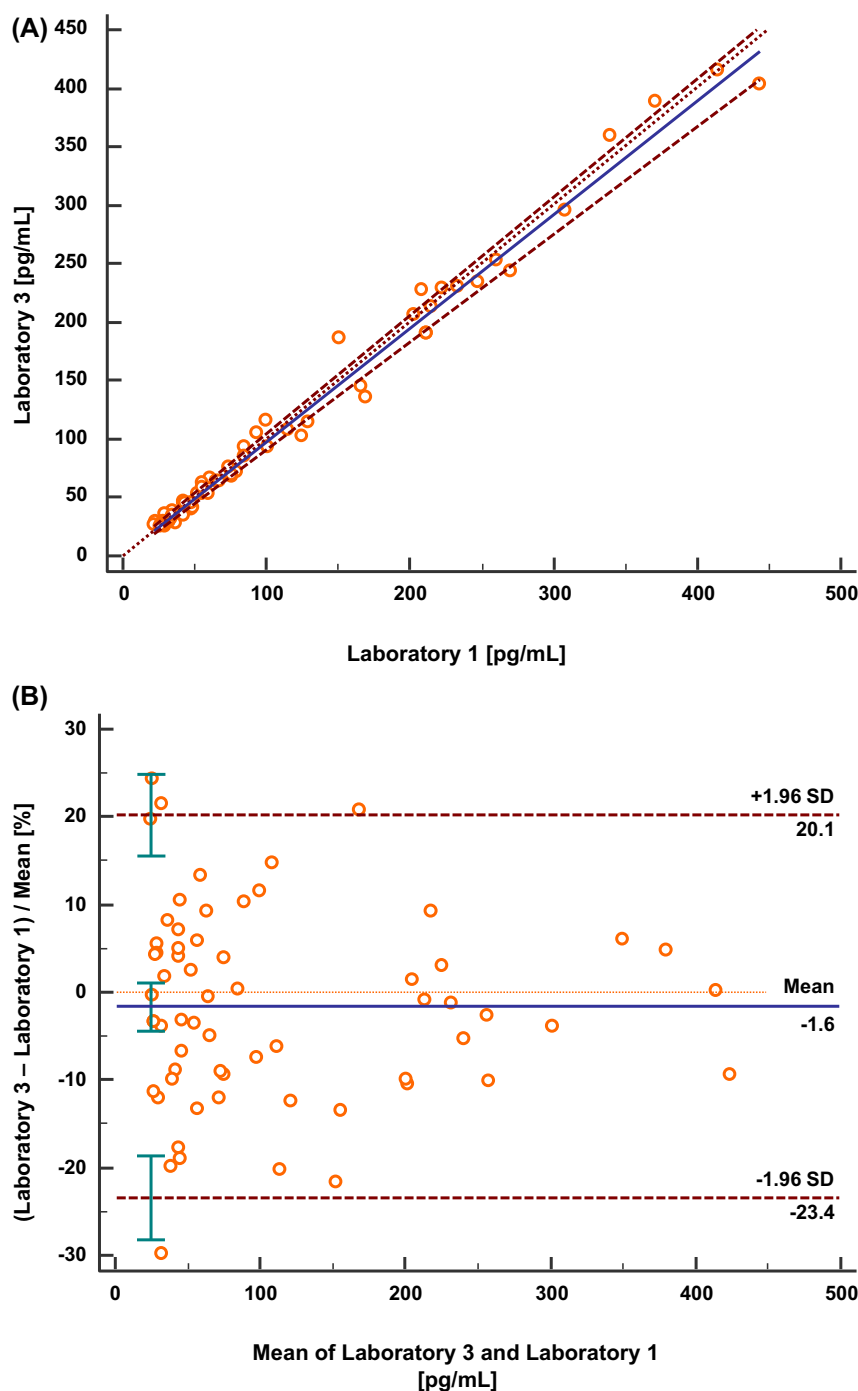


Figure 4: Results from the patient sample-based aldosterone method comparison study performed between two independent laboratories. (A) Passing-Bablok regression plot including the Pearson regression analysis for the method comparison study of the RMP ($n = 64$ patients, $n=1$ sample preparation) between the independent laboratories (Laboratory 1: Roche, Penzberg, and Laboratory 3: Institute of Laboratory Medicine at Leipzig University Hospital). Passing-Bablok regression plot (A); the regression analysis resulted in a regression equation $y = 0.97x + 0.65$, with a 95% confidence interval from 0.92 to 1.02 for the slope and from -2.2 to 3.4 for the intercept. The Pearson's correlation coefficient was ≥ 0.992 . (B) Bland-Altman plot for the method comparison study of the RMP ($n = 64$ patients, $n=1$ sample preparation) between two independent laboratories (Laboratory 1: Roche, Penzberg, and Laboratory 3: Institute of Laboratory Medicine at Leipzig University Hospital). The interlaboratory measurement bias was -1.6% (95% CI -4.4 to 1.1%) and the 2S interval of the relative difference was 21.7% (95% CI interval = 9.5%). CI, confidence interval; LC-MS/MS, liquid chromatography tandem mass spectrometry; RMP, reference measurement procedure; SD, standard deviation.

Table 4: Combined measurement uncertainty estimation for a single measurement.

	Level				
	60.0 pg/mL	150 pg/mL	1,000 pg/mL	Native patient sample 1 80 pg/mL	Native patient sample 2 200 pg/mL
Type B uncertainty	0.91	0.98	0.84	0.91	0.98
Calibrator preparation, CV %					
Characterization of reference material	0.2	0.2	0.2	0.2	0.2
Preparation of:					
Stock solution	0.22	0.22	0.22	0.22	0.22
Working solution	0.56	0.56	0.56	0.56	0.56
Spike solution	0.75	0.84	0.66	0.75	0.84
Matrix-based calibrator	0.91	0.98	0.84	0.91	0.99
Type A uncertainty	4.7	3.7	3.9	4.2	4.0
Intermediate precision, CV %					
Total measurement uncertainty	4.8	3.8	4.0	4.3	4.1
Single measurement, CV %					

Uncertainties are expressed as CVs for better interpretability. In all formulas, the terms “unc” corresponds to standard deviations. CV, coefficient of variation.

Table 5: Combined measurement uncertainty estimation for target value assignment (n=6).

	Level				
	60.0 pg/mL	150 pg/mL	1,000 pg/mL	Native patient sample 1 80 pg/mL	Native patient sample 2 200 pg/mL
Type B uncertainty	0.91	0.98	0.84	0.91	0.98
Calibrator preparation, CV %					
Characterization of reference material	0.2	0.2	0.2	0.2	0.2
Preparation of:					
Stock solution	0.22	0.22	0.22	0.22	0.22
Working solution	0.56	0.56	0.56	0.56	0.56
Spike solution	0.75	0.84	0.66	0.75	0.84
Matrix-based calibrator	0.91	0.98	0.84	0.91	0.98
Type A uncertainty	0.8	1.4	1.2	2.1	1.5
Mean of measurement results, CV %					
Total measurement uncertainty	1.2	1.7	1.5	2.3	1.8
Target value assignment, CV %					

Uncertainties are expressed as CVs for better interpretability. In all formulas, the terms “unc” correspond to standard deviations. CV, coefficient of variation.

Conclusions

This paper presents an analytical protocol based on ID-LC-MS/MS for the quantification of aldosterone in human serum and plasma. In contrast to the already published candidate RMP [2] this method uses a two-dimensional heart-cut LC approach in combination with two orthogonal stationary and mobile phases which increased peak capacity, minimized matrix effects and allowed the separation of relevant interferences. The optimized sample preparation protocol is

based on supported liquid extraction (SLE) and was proven to be easy transferable from SLE cartridges to a less labor-intensive automated 96-well plate format.

Special attention was paid to the characterization of the reference material as well as the preparation of calibrator material and the calculation of the measurement uncertainty. The use of the qNMR-characterized reference material for direct calibration of our RMP ensures traceability to the SI-unit kilogram, and thus the determination of the true value of aldosterone. In addition, an optimized multiple

point calibrator preparation scheme including all relevant information on materials used (analytical balance, volumetric flasks, and pipettes) as well as a detailed description on the estimation of uncertainty is provided.

In contrast to the candidate RMP [2], which uses a bracketing calibration that requires a pre-estimation of aldosterone concentration in native samples and a relatively high sample volume, our approach in combination with the SLE protocol allows us to measure low sample volume complaint samples as well as large method comparison studies in a reasonable time frame in addition to the intended use of target value assignment. Furthermore, we showed that all requirements for an aldosterone RMP on accuracy and precision were fully met. The transferability of the method to the second independent laboratory was shown without significant increase in bias between laboratories. This proves the method to be easy transferable and robust.

The candidate RMP is also in good agreement with the overall RELA reference laboratory network. While in 2016 only three laboratories participated, in 2018 there were already five and in 2021 already seven laboratories, one of which is listed as a JCTLM service laboratory. In combination with the qNMR approach, this highly selective 2D-LC-MS/MS method, provides a traceable and reliable platform for the standardization of routine tests and for the evaluation of aldosterone in clinical samples.

Acknowledgments: We would like to thank Aline Hoffmeister, Monika Kriner, Alexandra Herbig, Marion Deuster and Michael Dedio for their support in selecting and providing samples. Third-party medical writing assistance, under the direction of the authors, was provided by Erin Slobodian and Anna King, PhD, of Ashfield MedComms, an Inizio company, and was funded by Roche Diagnostics GmbH (Penzberg, Germany).

Research funding: Manfred Rauh and Uta Ceglarek with team are funded cooperation partners of Roche Diagnostics GmbH. Roche employees holding Roche non-voting equity securities (Genussscheine): Judith Taibon, Sara Cheikh Ibrahim, Christian Geletneky, Andrea Geistanger.

Author contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 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; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that

questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Competing interests: Judith Taibon, Neeraj Singh, Sara Cheikh Ibrahim, Christian Geletneky, Galina Babitzki and Andrea Geistanger are all employees of Roche Diagnostics GmbH. Tobias Santner was an employee of Roche Diagnostics GmbH at the time of the study. Manfred Rauh and Daniel Köppl are employees of the Erlangen Clinic for Children and Adolescents, University Hospital Erlangen. Uta Ceglarek and Alexander Gaudl are employees of the Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital Leipzig.

Informed consent: Not applicable.

Ethical approval: All procedures were in accordance with the Helsinki Declaration. All samples used were exclusively anonymized, residual samples.

References

1. Rainey WE, White P. Functional adrenal zonation and regulation of aldosterone biosynthesis. *Curr Opin Endocrinol Diabetes Obes* 1998;5: 175–82.
2. Zhang Q, Han L, Zheng S, Ouyang F, Wu X, Yan J, et al. An isotope dilution liquid chromatography-tandem mass spectrometry candidate reference measurement procedure for aldosterone measurement in human plasma. *Anal Bioanal Chem* 2021;413:4471–81.
3. Katsu Y, Baker ME. Subchapter 123E – aldosterone, handbook of hormones, 2nd ed. Academic Press; 2021:951–3 pp. Available from: <https://ms.b-ok.cc/book/18832028/529801?dsource=recommend> [Accessed 31 August 2022].
4. Scott JH, Menouar MA, Dunn RJ. Physiology, aldosterone. [Updated 2022 May 4]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK470339/> [Accessed 30 Aug 2022].
5. Panteghini M. Traceability as a unique tool to improve standardization in laboratory medicine. *Clin Biochem* 2009;42:236–40.
6. Joint Committee on Traceability in Laboratory Medicine (JCTLM). Traceability – the key to accurate laboratory results for patients. Available from: <https://www.jctlm.org/> [Accessed 31 August 2022].
7. Kessler A, Siekmann L, Weykamp C, Geilenkeuser WJ, Dreazen O, Middle J, et al. External Quality Assessment Scheme for reference laboratories – review of 8 years' experience. *Clin Chem Lab Med* 2013;51:997–1005.
8. JCTLM and IFCC. External quality assessment scheme for reference laboratories in laboratory medicine (RELA). Available from: <http://dgkl-rfb.de:81/> [Accessed 31 August 2022].
9. Siekmann L, Spiegelhalder B, Breuer H. Bestimmung von Aldosteron im Plasma des Menschen unter Verwendung der Kombination Gas-Chromatographie/Massenspektrometrie. *Fresenius' Z für Anal Chem* 1972;261:377–81.
10. Stöckl D, Reinauer H, Thienpont LM, De Leenheer AP. Determination of aldosterone in human serum by isotope dilution gas chromatography/mass spectrometry using a new heptafluorobutyl derivative. *Biol Mass Spectrom* 1991;20:657–64.
11. Thienpont LM, Van Nieuwenhove B, Stöckl D, Reinauer H, De Leenheer AP. Determination of reference method values by isotope dilution-gas chromatography/mass spectrometry: a five years'

- experience of two European Reference Laboratories. *Eur J Clin Chem Clin Biochem: J Forum Eur Clin Chem Soc* 1996;34:853–60.
12. Singh N, Taibon J, Pongratz S, Geletneky C. Absolute content determination by quantitative NMR (qNMR) spectroscopy: a curious case of aldosterone. *RSC Adv* 2021;11:23627–30.
 13. Nelson MA, Waters JF, Toman B, Lang BE, Rück A, Breitruck K, et al. A new realization of SI for organic chemical measurement: NIST PS1 primary standard for quantitative NMR (benzoic acid). *Anal Chem*. 2018;90:10510–7.
 14. Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG1). Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM), (2008; GUM 1995 with minor corrections). Available from: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6 [Accessed 15 September 2022].
 15. European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). EFLM biological variation database – aldosterone. Available from: <https://biologicalvariation.eu/search?query=Aldosterone> [Accessed 30 August 2022].
 16. Stöckl D, Sluss PM, Thienpont LM. Specifications for trueness and precision of a reference measurement system for serum/plasma 25-hydroxyvitamin D analysis. *Clin Chim Acta Int J Clin Chem* 2009;408:8–13.
 17. C62-A C. Liquid chromatography-mass spectrometry methods; approved guidelines. Available from: https://clsi.org/media/1346/c62a_sample.pdf [Accessed 29 September 2022]. 2014.
 18. European Medicines Agency (EMA). ICH harmonised tripartite Guidelines, Validation of analytical procedures: Text and methodology Q 2 (R1); 1995. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-q-2-r1-validation-analytical-procedures-text-methodology-step-5_en.pdf [Accessed 28 September 2022].
 19. Matuszewski BK. Standard line slopes as a measure of a relative matrix effect in quantitative HPLC–MS bioanalysis. *J Chromatogr B* 2006;830:293–300.
 20. Botelho JC, Shacklady C, Cooper HC, Tai SS, Van Uytvanghe K, Thienpont LM, et al. Isotope-dilution liquid chromatography-tandem mass spectrometry candidate reference method for total testosterone in human serum. *Clin Chem* 2013;59:372–80.
 21. Botelho JC, Ribera A, Cooper HC, Vesper HW. Evaluation of an isotope dilution HPLC tandem mass spectrometry candidate reference measurement procedure for total 17- β estradiol in human serum. *Anal Chem* 2016;88:11123–9.
 22. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev* 2008;29:S49–52.
 23. Schuetzenmeister A, DufeyVCA F. Variance component analysis. Available from: <https://cran.r-project.org/web/packages/VCA/index.html> [Accessed 15 September 2022].
 24. Analyse-it contributors. Analyse-it. Available from: <https://analyse-it.com/> [Accessed 22 August 2022].
 25. Gaudl A, Kratzsch J, Bae YJ, Kiess W, Thiery J, Ceglarek U. Liquid chromatography quadrupole linear ion trap mass spectrometry for quantitative steroid hormone analysis in plasma, urine, saliva and hair. *J Chromatogr A* 2016;1464:64–71.
 26. Schiffer L, Shaheen F, Gilligan LC, Storbeck KH, Hawley JM, Keevil BG, et al. Multi-steroid profiling by UHPLC-MS/MS with post-column infusion of ammonium fluoride. *J Chromatogr B Analyt Technol Biomed Life Sci* 2022;1209:123413.
-
- Supplementary Material:** This article contains supplementary material (<https://doi.org/10.1515/cclm-2022-0996>).