Myriam Ott, Friederike Bauland, Vanessa Fischer, Daniel Köppl, Alexander Gaudl, Andrea Geistanger, Uta Ceglarek, Manfred Rauh and Judith Taibon*

Candidate reference measurement procedure based on isotope dilution-two dimensional-liquid chromatography–tandem mass spectrometry for the quantification of androstenedione in human serum and plasma

https://doi.org/10.1515/cclm-2024-1135
Received September 30, 2024; accepted May 4, 2025;
published online May 23, 2025

Abstract

Objectives: Androstenedione (ASD) is a biomarker used in the diagnosis of hyperandrogenism and adrenal hyperplasia. Therefore, accurate measurement of serum ASD is pivotal in clinical settings. We aimed to develop a novel highly selective reference measurement procedure (RMP) based on isotope dilution-two dimensional-liquid chromatography-tandem mass spectrometry (ID-2D-LC-MS/MS) for the quantification of ASD in human serum/plasma.

Methods: To achieve high selectivity and sensitivity, a two-dimensional heart-cut LC approach for LC-MS/MS and a supported liquid extraction sample preparation protocol were employed. Matrix effects were evaluated through a post-column infusion experiment and comparison of standard line slopes. Precision and accuracy were determined via a multi-day validation experiment. Reproducibility was assessed by comparing results from two independent laboratories, and measurement uncertainty (MU) was evaluated in compliance with current guidelines.

Results: Our novel RMP proved effective for measuring ASD concentrations ranging from 0.00800 ng/mL to 12.0 ng/mL (0.0279 nmol/L to 41.9 nmol/L) and demonstrated matrix-independence. The relative mean bias varied between 0.6 and 2.2 % across different matrices and concentration levels. Intermediate precision was observed to be between 1.2 and 1.9 %. The expanded measurement uncertainty for ASD target value assignment ranged from 1.7 to 2.5 %, irrespective of sample concentration. Equivalence to the JCTLM-listed RMP was evaluated through a method comparison study, revealing a high degree of agreement ($r \geq 0.997$). Additionally, by comparing results from two independent laboratories, the transferability and reproducibility of the RMP were confirmed.

Conclusions: This isotope dilution two-dimensional LC-MS/MS method can be used for the evaluation and standardization of routine assays and for measuring individual patient samples, ensuring traceability.

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

Introduction

Androstenedione (ASD) is a naturally occurring steroidal hormone synthesized in males and females by the adrenal glands and gonads. It serves as an intermediary in the biosynthesis of testosterone and estrone, playing a critical role in various physiological functions. Elevated levels of ASD may serve as an indicator of underlying disorders such as Cushing syndrome [1], congenital adrenal hyperplasia [2], and polycystic ovary syndrome [3]. These clinical presentations emphasize the importance of monitoring ASD levels for diagnostic purposes.

***Corresponding author: Judith Taibon**, PhD, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany,
E-mail: judith.taibon@roche.com
Myriam Ott, Vanessa Fischer and Andrea Geistanger, Roche Diagnostics GmbH, Penzberg, Germany
Friederike Bauland, Chrestos Concept GmbH & Co. KG, Essen, Germany
Daniel Köppl and Manfred Rauh, Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Erlangen, Germany
Alexander Gaudl and Uta Ceglarek, Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig Medical Center, Leipzig, Germany. https://orcid.org/0000-0002-7197-3868 (A. Gaudl)

A variety of published methods are available for quantifying ASD in clinical settings, from routine immunoassays [4, 5] to more accurate and selective methods based on liquid chromatography-mass spectrometry (LC-MS) [6–9]. In particular, one LC-MS/MS-based reference measurement procedure (RMP) developed by our group and published by Gradl et al. [10] has been recognized by the Joint Committee on Traceability in Laboratory Medicine (JCTLM) [11] as a higher-order RMP for the analysis of ASD [12]. In this RMP, the certified reference material from the National Measurement Institute in Australia (NMIA) was used as the primary standard. Furthermore, a quantitative nuclear magnetic resonance (qNMR) based procedure was described to value assign the primary reference material, ensuring direct traceability to SI units and independence from the availability of reference materials [10].

The aim of this project was to modernize the existing JCTLM-listed ID-LC-MS/MS based reference measurement procedure (RMP) for quantifying ASD in human serum and plasma. This encompassed updates to the preparation of calibrator materials, improvements in sample handling, optimization of the analytical setup, expansion of the measurement range and the reduction of measurement uncertainty. Moreover, the RMP must meet the predefined acceptance criteria for bias (B) and imprecision (coefficient of variation [CV]), as established by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) biological variation database.

We provide a thorough description of the methodology to ensure reproducibility of this candidate RMP in other laboratories. Detailed technical information regarding the ID-2D-LC-MS/MS protocol can be found in Supplementary Material 1, and the calculation of measurement uncertainty (MU) in Supplementary Material 2.

Materials and methods

A detailed operations procedure for the test procedure methodology as well as a list of laboratory equipment and relative requirements can be found in Supplementary Material 1.

Chemicals and reagents

LC-MS grade methanol (CAS 67-56-1), acetonitrile (CAS 75-05-8) and formic acid (99 % ULC/MS grade CAS 64-18-6) were purchased from Biosolve (Valkenswaard, The Netherlands). Ammonium fluoride (CAS 12125-01-8) was purchased from Sigma Aldrich (Taufkirchen, Germany). Isopropanol (CAS 67-63-0) and zinc sulfate heptahydrate (CAS 7446-20-0) were purchased from Honeywell (Charlotte, NC, USA).

Water was purified using a Millipore Milli-Q IQ 7000 system from Merck (Darmstadt, Germany). Steroid-free human serum (ID 11717863001) and albumin RPLA 1 (Surrogate matrix, Assay Quality, ID 11726536001) were obtained from Roche Diagnostics GmbH (Mannheim, Germany). Mixed-gender native plasma pools (Li-Heparin [CUST-BB-17022020-4C], dipotassium (K_2) ethylene diamine tetraacetic acid (EDTA) [CUST-BB-17022020-4E], and tripotassium (K_3) EDTA [CUST-BB-17022020-4D]) were obtained from Biotrend (Cologne, Germany). As surrogate matrix an albumin matrix with a concentration of 1 g/dL in phosphate-buffered saline was used. Native sample pool was composed of samples from 5 individual patients with endogenous ASD levels. All native sample pools were prepared in accordance with the Declaration of Helsinki. The primary reference material ASD (CAS 63-05-8) was purchased from NMIA (M955), and its isotope-labeled internal standard (ISTD), $^{13}C_3$-ASD (CAS 327048-86-2), was obtained from Supelco™ (Merck). 6-Dehydrotestosterone (CAS 2484-30-2) was obtained from the European Directorate for the Quality of Medicines (EDQM). 5β-Androstan-3,17-dione (CAS 1229-12-5), 4-androsten-3α-ol-17-one (CAS 2791-99-3), and 4-androsten-3β-ol-17-one (CAS 571-44-8) were purchased from Steraloids, Inc. (Newport, RI, USA). 1-Testosterone (CAS 65-06-5) was sourced from Cayman Chemical (Ann Arbor, MI, USA). Testosterone (CAS 58-22-0) and epitestosterone (CAS 481-30-1) were obtained from Supelco® (Merck) and dehydroepiandrosterone (DHEA, CAS 53-43-0) was obtained from Sigma-Aldrich (Merck).

Preparation of calibrators and quality control (QC) samples

Three calibration stock solutions were independently prepared by weighing 1 mg (± 5 %) of the reference material NMIA (M955 Batch 04-S-02, 99.6 ± 0.3 % [k=2]) using an ultramicrobalance and dissolving it in acetonitrile to 100 mL, resulting in a 10 µg/mL (34.9 µmol/L) solution. The exact concentration was based on the material's purity and weight. Two stock solutions were further diluted with methanol/water (1 + 1, v + v) to prepare 100 ng/mL (349 nmol/L) working solutions. These were used to prepare eight calibrator spike solutions of varying concentrations in methanol/water (1 + 1, v + v). Each calibrator spike solution was diluted in an albumin surrogate matrix (1 + 49, v + v), resulting in a range of 0.00800–12.0 ng/mL (0.0279–41.9 nmol/L).

Quality control samples (QCs) were prepared using the third independent primary stock solution (10 µg/mL, ± 5 %). This solution was diluted with methanol to create a working solution (0.1 µg/mL) used to prepare respective spike solutions. Final matrix based QC concentrations were 0.0240, 2.00, and 10.0 ng/mL (0.0838, 6.98, and 34.9 nmol/L). Additionally, two native patient samples were used as QCs to establish a control chart.

ISTD solution

The commercially available stock solution, $^{13}C_3$-ASD 100 µg/mL (15 µL), was transferred into a volumetric flask (250 mL) and diluted with methanol/water (1 + 1, v + v) up to the calibration mark, to obtain an ISTD solution with a concentration of 6 ng/mL (20.7 nmol/L).

Sample preparation

Native serum, plasma (lithium-heparin, K_2EDTA, K_3EDTA), and albumin surrogate matrix (1 g/dL) can be used as sample matrices. A 200 µL sample was transferred into a 1.5 mL screw cap, and 40 µL of internal standard working solution was added. Samples were incubated on an overhead shaker for 15 min at room temperature. Then, 600 µL of precipitation reagent was added, followed by shaking for 10 min on a thermomixer (Eppendorf 5,382 thermomixer C) at 1,400 rpm and centrifuging for 20 min at 15,000 rcf (centrifuge 5,430 R, Eppendorf). The supernatant (600 µL) was loaded onto 1 cc Support Liquid Extraction (SLE) Waters HLB Oasis Prime cartridges. Cartridges were washed with 1,000 µL Milli-Q water. Extraction was performed by adding 200 µL acetonitrile (twice), evaporating the extract, and reconstituting the residue in 100 µL methanol/water (1 + 1, v + v). The solution was transferred into a high-performance liquid chromatography (HPLC) vial with a 100 µL insert.

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

A two-dimensional heart-cut LC approach was used for the chromatographic purification. This approach involves the use of two orthogonal stationary phases: a Waters Acquity Ethylene Bridged Hybrid (BEH) C18 column (50 × 2.1 mm, 1.7 µm, Waters Corporation, Milford Massachusetts, USA) and a Raptor Biphenyl column (150 × 2.1 mm, 2.7 µm, Restek, Bellefonte, Pennsylvania, USA). The mobile phases used in this method consist of water (A1) and acetonitrile (B1) in the first dimension, and 0.2 mM ammonium fluoride in water (A2) and methanol/acetonitrile (7 + 3, v + v) (B2) in the second dimension. ASD was detected in multiple reaction monitoring (MRM) mode using an AB Sciex QTrap 6,500 + mass spectrometer operating in electrospray positive ionization mode. The quantifier ion transition (ASD m/z 287.1 → 97.0) and the corresponding internal standard transition ($^{13}C_3$-ASD m/z 290.1 → 100.0) served as the basis for ASD quantitation. Supplementary Material 1 includes detailed LC and MS methods.

System suitability test (SST)

To ensure long-term traceability, an SST was established to check sensitivity and chromatographic resolution before each sequence. Two levels (SST1 and SST2) in methanol/water (1 + 1, v + v) contained ASD at 0.00800 ng/mL and 12.0 ng/mL (0.0279 nmol/L and 41.9 nmol/L), respectively.

SST1 was injected to check sensitivity, with the signal-to-noise (S/N) ratio calculated using Analyst software (AB Sciex). An S/N ratio of ≥10 was required. SST2 was injected to check for carryover, followed by a double blank injection. The peak observed in the first blank was required to be≤20 % of the peak area of calibrator 1 to pass the SST. A retention time of 7.9 ± 0.5 min for ASD was required for both SST1 and SST2.

Calibration data processing and structure of analytical series

Calibrator levels were measured at increasing concentrations at the beginning and end of every sequence, using 8 levels to generate the final calibration function through linear regression (y=ax + b) of area ratios (analyte/internal standard).

For processing the raw data file, the Analyst software (v1.6.2 or higher) was used with Intelli Quan as Quantitation Integration Algorithm. ASD and its internal standard showed a retention time of 7.9 min and were integrated with a 30 s window. Raw data were processed with a smoothing factor of 3. The calibration curve was linear with a 1/x weighting.

The sequence setup varied based on measurement requirements. For reference value assignments, sample preparations (n=x) depended on the desired measurement uncertainty, with samples measured on at least two different days. For method comparison studies or complaint sample measurements, samples were prepared with n=1.

Method validation

Assay validation and determination of MU were performed following established guidelines for validation. These guidelines include the *Clinical & Laboratory Standards Institute's C62A Liquid Chromatography-Mass Spectrometry Methods* [13], and the *Guide to the expression of uncertainty in measurement* (GUM) [14].

Analytical performance specifications (APS) have been calculated based on the European Federation of Clinical Chemistry and Laboratory Medicine biological variation (BV) database [15]. For ASD, several entries can be found in the database for the desirable specifications. Based on these entries the recommended specification for imprecision (CV), bias (B), maximum expanded allowable measurement uncertainty (MAU), and total error (TE) range between 2.9 and 5.8 % CV_{RMP}, 2.7–5.0 % $B_{RMP,}$ 3.8–7.7 % MAU_{RMP} and 9.1–14.5 % TE_{RMP} [15, 16]. Therefore, these criteria are established as the acceptance criteria for the candidate RMP.

Selectivity

Selectivity was assessed by spiking the following potential interferences into a neat solution (methanol/water [1 + 1, v + v]) (concentration 10 ng/mL; 34.7–34.9 nmol/L) to assess the selectivity to ASD (concentration 5 ng/mL; 17.5 nmol/L) and the internal standard $^{13}C_3$-ASD: 6-dehydrotestosterone, testosterone, dehydroepiandrosterone (DEHA), 4-androsten-3β-ol-17one, 1-testosterone, epitestosterone, 4-androsten-3α-ol-17-one and 5β-androstan-3,17-dione. Baseline separation of the analytes was evaluated (chromatographic resolution (FWHM) R≥1.5).

To determine if there were any signals from the matrix that could interfere with the quantifier and qualifier transitions of the analytes, both the analyte-free matrix and the pool of native human matrix were examined at the expected retention time.

Furthermore, the analyte-free matrix was spiked with the labeled internal standard to assess any residual unlabeled analyte present within the stable isotope labeled internal standard.

Matrix effects (ME)

In a qualitative post-column infusion setup, a solution containing 50.0 ng/mL (173 nmol/L) of $^{13}C_3$-ASD in methanol/water (1 + 1, v + v) was infused into the effluent of the HPLC column using a T-piece at a flow rate of 10 µL/min. Various processed matrix samples were then injected, including a neat solution (methanol/water [1 + 1, v + v]), the native human serum pool, an analyte-free surrogate matrix, and three plasma matrices (Li-Heparin, K_3EDTA, K_2EDTA). The purpose of this experiment was to observe any changes in the MS signals at the expected retention time, which would indicate an effect on ionization caused by components in the matrix.

In addition to the T-piece experiment, a comparison of standard line slopes was conducted using the following matrices: neat solution, surrogate matrix, native human serum pool, and Li-Heparin plasma matrix. The endogenous content of the native matrices was considered, and depending on the content, a calibration curve with a limited measuring range was created. Neat samples were diluted to match the final concentration of processed calibrator levels, while matrix calibrator levels were prepared using solid-phase extraction (SPE) (Supplementary Material 1). The slopes and coefficients of determination were compared between these different matrices and confidence intervals required to overlap. Furthermore, the calibrator samples in the surrogate matrix and native serum were evaluated as controls by using neat calibration as the standard. The recoveries were then reported as the percentage of the measured concentration recovered relative to the nominal concentration.

Linearity

To evaluate linearity, three sets of calibrators were prepared in a surrogate matrix, with each set involving individual weighings. The calibration range (0.00800 ng/mL to 12.0 ng/mL; 0.0279–41.9 nmol/L) was expanded by ± 20 % by preparing two additional spike solutions, resulting in spiked samples with final concentrations of 0.00640 ng/mL and 14.4 ng/mL of ASD (0.0223 nmol/L to 50.3 nmol/L). The correlation coefficient (r) and residuals were calculated for each curve, with the correlation coefficient required to be≥0.999. Additionally, the linearity of the method was confirmed by evaluating the recovery of serially diluted samples using the preferred regression model for calculation. Pool 11 was a native serum sample spiked to a concentration of ~12 ng/mL (41.9 nmol/L), while for pool 1 a steroid-free human serum sample was used. Using these two sample pools, pools 2 to 10 were serially diluted according to the ratios 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). The measurement results

should demonstrate a linear relationship. Additionally, the recovery is reported as the percentage of measured concentration relative to the nominal concentration of the sample pool.

Precision, trueness and accuracy

Precision was assessed over multiple days using ANOVA-based variance-components analysis, considering variability from injections, preparations, calibrations, and days. The experiment involved measurements conducted over five individual days, with two individual calibrator preparations for two measurement sequences (Part A and Part B) each day. For each part, three spiked samples (at concentrations of ~0.0240 ng/mL [0.0838 nmol/L] in surrogate matrix, and 5.00 ng/mL [17.5 nmol/L] and 10.0 ng/mL [34.9 nmol/L] in steroid-free human serum) and two native patient samples (~0.850 ng/mL and 1.50 ng/mL, 2.97 nmol/L and 5.24 nmol/L) were prepared in triplicate and injected twice.

Data evaluation was conducted using Biowarp, an internal statistical program based on the Variance Component Analysis Roche Open Source software package in R [17]. As independent production of calibrators (including preparation of stock and working solutions, spike solutions, and matrix-based calibrator levels) does not account for this variability, it was evaluated as type B uncertainty, as explained in more detail in Supplementary Material 2.

Trueness and accuracy were evaluated using the same spiked samples from the precision experiment. Additionally, Li-Heparin plasma samples were fortified with concentrations of 1.00 ng/mL, 5.00 ng/mL, and 10.0 ng/mL (3.49 nmol/L, 17.5 nmol/L and 34.9 nmol/L). The validity of dilution was assessed using high concentrated surrogate matrix samples at concentration levels of 13.0 ng/mL and 20.0 ng/mL (45.4 nmol/L and 69.8 nmol/L). For both serum and plasma, samples were prepared in triplicate for each part, A and B, on a single day. Additionally, for diluted samples, samples were prepared in five replicates for each part. Based on the entries in the EFLM database the defined acceptance criteria for the RMP range between 2.9 and 5.8 % CV_{RMP} and 2.7–5.0 % B_{RMP}.

Lower limit of the measuring interval (LLMI)

The precision, trueness and accuracy at the LLMI were determined by measuring spiked surrogate matrix samples in the expected concentration range of the quantitation limit (0.00800 ng/mL, 0.0279 nmol/L). Sample preparation was repeated six times. Bias and precision were determined and had to be within the range defined in the accuracy and precision experiment.

Sample stability

The stability of processed samples was evaluated at 8 °C. Samples were measured after 1, 6, and 15 days of storage, and concentrations were determined using a freshly prepared calibration. Recoveries were calculated by comparing the measured value with the nominal concentration (t=0 days). Stability of spike solutions (1.00 ng/mL, 6.00 ng/mL and 12.0 ng/mL; 3.49 nmol/L, 20.9 nmol/L and 41.9 nmol/L) and control material (0.0240 ng/mL, 2.00 ng/mL and 10.0 ng/mL, 0.0838 nmol/L, 6.98 nmol/L and 34.9 nmol/L) stored at −80 °C was evaluated for 13 weeks (n=1 per each time point). The recoveries were calculated by comparing the measured value with freshly prepared samples. All individual measurements had to fall within the predetermined Total Error criterion as described above. Stability can be guaranteed for a measurement interval of 2–28 days for y–1 day, and for a measurement interval of >4 weeks for y–1 week.

Equivalence of results between independent laboratories

The agreement of our novel RMP between two independent laboratories (Site 1: Roche Diagnostics GmbH, Penzberg and Site 2: Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen) was assessed by a method comparison study including 76 samples. Of these samples, 36 were anonymized residual patient serum samples and 20 were native serum pools. In addition, 10 spiked serum samples and 10 diluted serum samples were needed, to cover the entire working range. The samples were measured over a period of 3 days or more, with a random allocation and single preparation for each sample.

Furthermore, a three-day precision experiment was performed at Site 2 based on the experimental design described above. Spiked samples were provided by Roche Diagnostics GmbH Penzberg (Site 1). Both laboratories prepared their own calibrator levels using the NMIA standard as primary reference material. Adaptations to the sample preparation protocol were required at Site 2 in order to use the 96-well plate format. All steps before the SPE were done in one single 96-well deep well plate from polypropylene. In brief, 200 µL sample and 40 µL internal standard were incubated before adding 600 µL of precipitation reagent. The sample was mixed, centrifuged, and 600 µL of supernatant was loaded onto a prepared 96-well plate for SPE (Waters

Oasis HLB). After washing, the samples were eluted using 2 × 200 µL acetonitrile. The eluate was collected in one set of glass inserts supplied by Hirschmann (Eberstadt, Germany). After evaporation, the samples were reconstituted (100 µL), mixed, sealed and measured.

A second method comparison study was conducted between Site 1, performing the RMP, and a third laboratory performing the analysis using the SPE LC-MS/MS-based method outlined by Gaudl et al. [9] (Site 3: Institute of Laboratory Medicine at Leipzig University Hospital).

Equivalence of results between the candidate and the JCTLM listed RMP

To assess the agreement between our novel candidate RMP and the JCTLM-listed RMP (C16RMP6), a total of 20 anonymized leftover patient serum samples were prepared and measured (n=1) using the respective methods referred to in the study. There are differences in the protocols regarding the preparation of calibrators, the volume of samples used for sample preparation, and the method of chromatographic separation. Specifically, for our novel candidate RMP, the two-dimensional chromatographic separation approach is utilized.

Estimation of measurement uncertainty (MU)

MU was determined according to the GUM [14] and considered the following steps: (i) the uncertainty of the primary reference material based on its certificate, (ii) the estimation of uncertainty in the preparation of calibrator materials, and (iii) the estimation of uncertainty in the LC-MS/MS method. Detailed information is reported in Supplementary Material 2.

The overall uncertainty in the measurement process for a single measurement was determined by considering two main factors: the uncertainty in preparing the calibrator (unc_{cal}) and the calculated uncertainty in the precision experiment (unc_{prec}). To establish reference or target values, multiple sample preparations were conducted for each sample on at least two different days, and the average (n=x) of these results was calculated. The measurement uncertainty was then determined by combining the uncertainty in calibrator preparation (unc_{cal}) with the uncertainty (SD) of the mean of the measurement results (unc_{mean}). To avoid underestimating the error, the calibrator level with the highest uncertainty (unc_{cal}) for each individual sample concentration level was selected when combining uncertainties. The resulting 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. The maximum expanded allowable measurement uncertainty (MAU) was determined to range between 3.8 and 7.7 %.

Results

Selectivity

Effective separation of ASD from 6-dehydrotestosterone, 4-androsten-3α-ol-17-one and 5β-androstan-3,17-dione was achieved in the first dimension (Figure 2A). ASD, testosterone, DHEA, 4-androsten-3β-ol-17-one, 1-testosterone and epitestosterone were transferred to the second dimension and successfully separated from ASD with a chromatographic resolution R (FWHM) ≥ 3.2 (Figure 2B).

During the analysis, no interfering signals were observed for the quantifier transition (*m*/*z*: 287.1 -> 97.0) and the corresponding internal standard transition (*m*/*z*: 290.1-> 100.0) within the expected retention time window (7.9 min ± 0.5 min) (Figure 1). This observation was consistent in both the analyte-free surrogate matrix and the native serum pool.

The internal standard used in this method was evaluated for any residual unlabeled analyte, and no significant analyte signal was observed. Therefore, it can be confidently used as the internal standard for our novel RMP.

Matrix effects (ME)

The results of the post column infusion experiment indicated that none of the tested matrices exhibited ion suppression or enhancement during the retention time of the analyte of interest.

The slopes of the standard lines were determined to be 0.770 (95 % CI 0.768–0.773) for the neat solution, 0.762 (95 % CI 0.758–0.766) for the surrogate matrix, 0.761 (95 % CI 0.748–0.774) for the native matrix, and 0.773 (95 % CI 0.759–0.778) for the Li-Heparin plasma. The comparable slopes and overlapping corresponding confidence intervals suggest that they are not significantly different from each other, confirming the absence of a matrix effect. Additionally, excellent correlation was observed (r≥0.999), regardless of the matrix used. Calibrator samples in the surrogate matrix, native serum, and Li-Heparin plasma were evaluated as controls using the neat calibration as the standard. The average recovery for the surrogate matrix was 99 % (95 % CI 98–101), for the native serum matrix was 99 % (95 % CI 98–100), and for the Li-Heparin plasma was 98 % (95 % CI 97–100), thereby further proving matrix independence.

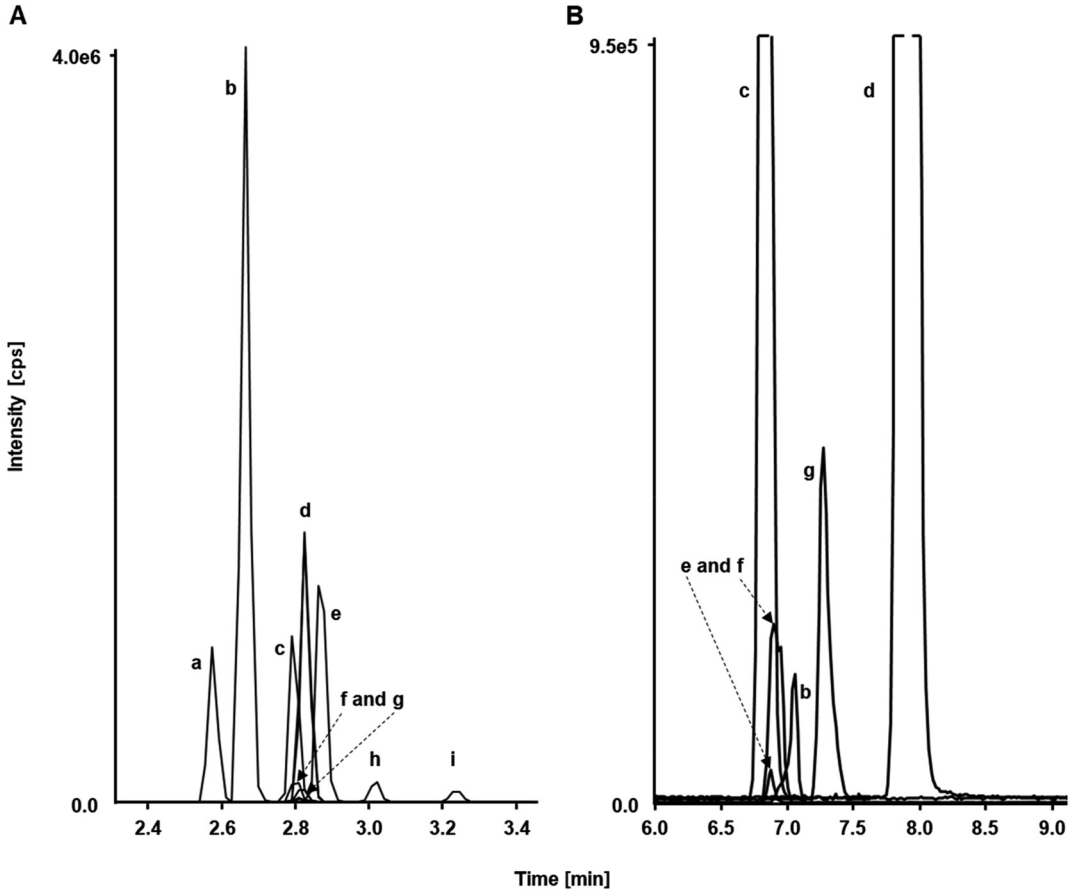

Figure 1: Chromatographic separation of androstenedione from potential interfering compounds. Concentration ASD in neat solution (methanol/water [1 + 1, v + v]) 5.00 ng/mL (17.5 nmol/L), concentration of all potential interfering compounds 10.0 ng/mL (34.7–34.9 nmol/L). (A) Chromatogram (overlay) of the 1st dimension. (B) Chromatogram (overlay) for the 2nd dimension. Interferences are separated from ASD with a chromatographic resolution (FWHM) ≥ 5.4.

Linearity

The residuals of the linear regression model, which utilized a 1/x weighting, were randomly and uniformly distributed. Excellent correlation (r≥0.999) was achieved for all individual calibration curves and acceptance criteria were fulfilled.

Additionally, the linearity of the method was confirmed by analyzing serially diluted samples. The measurement results exhibited a linear relationship with an r≥0.999. Recoveries for the serial dilutions ranged from 99 to 102 %.

Lower limit of the measuring interval (LLMI)

The LLMI was determined using a spiked matrix sample with a concentration of 0.00800 ng/mL (0.0279 nmol/L), which corresponds to the concentration of the lowest calibrator level. The relative deviation (n=6) was 3.5 % and CV was determined at 3.1 %, thus fulfilling the defined APS.

Precision, trueness and accuracy

The intermediate precision, which includes variances related to between-day calibration, preparation, and injection, ranged between 1.2 and 1.9 %. The repeatability CV ranged between 0.9 and 1.6 % across all concentration levels (Table 1).

Trueness and accuracy were assessed by analyzing spiked surrogate matrix, steroid-free human serum, and native Li-Heparin plasma pools. The relative mean bias ranged from 0.6 to 2.2 %, regardless of the matrix and concentration level.

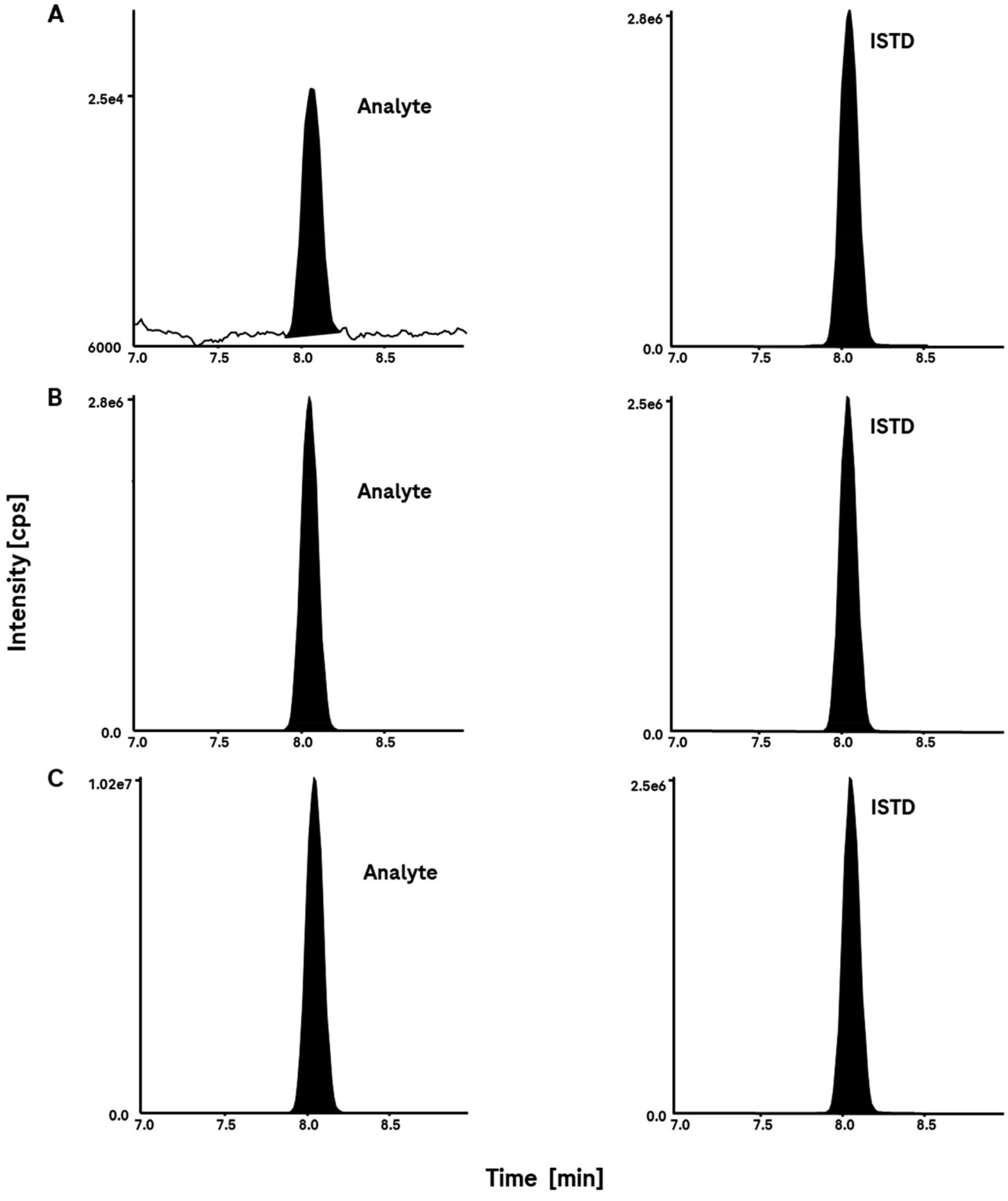

Figure 2: Androstenedione LC-MS/MS derived analytical readouts. (A) Chromatogram of calibrator level 1 with a concentration of 0.00800 ng/mL (0.0279 nmol/L) spiked in surrogate matrix (left) and ISTD (right); (B) pooled native patient human serum with a concentration of 1.48 ng/mL (5.17 nmol/L), analyte (left) and ISTD (right); (C) fortified Li-heparin patient sample with a concentration of 5.52 ng/mL (19.3 nmol/L), analyte (left) and ISTD (right). ISTD, internal standard; SRM, standard reference material.

Table 1: Precision performance parameters for androstenedione quantification using the candidate RMP (n=60 measurements).

Variance source	CV%				
	Level 1 0.024 ng/mL (0.0838 nmol/L)	**Level 2 5.00 ng/mL (17.5 nmol/L)**	**Level 3 10.0 ng/mL (34.9 nmol/L)**	**Patient sample 1 0.874 ng/mL (3.05 nmol/L)**	**Patient sample 2 1.48 ng/mL (5.17 nmol/L)**
Intermediate	**1.9**	**1.2**	**1.8**	**1.9**	**1.5**
Between-day	0.9	0.8	0.9	0.3	0.4
Between-calibration	0.4	0.2	0.5	1.5	1.0
Repeatability	**1.6**	**0.9**	**1.4**	**1.2**	**1.1**
Between-preparation	0.9	0.4	1.2	0.7	0.6
Between–injection	1.3	0.8	0.8	1.0	0.9

Conversion factor ng/mL to nmol/L: 3.49. CV, coefficient of variation; RMP, reference measurement procedure. The coefficients of variation for repeatability and intermediate precision, which were determined from the individual variances, are printed in bold.

Table 2: Bias and 95 % CI of surrogate matrix, serum and Li-heparin plasma samples and dilutions.

Concentration	Surrogate matrix and serum samples		Concentration	Plasma	
	Mean bias, %	**95 % CI, %**		**Mean bias, %**	**95 % CI, %**
Level 1 0.024 ng/mL (0.0838 nmol)	0.9	−0.9 to 2.7	Level 1 1.00 ng/mL (3.49 nmol/L)	0.6	−0.9 to 2.1
Level 2 5.00 n/mL (17.5 nmol/L)	2.1	1.4 to 2.8	Level 2 5.00 n/mL (17.5 nmol/L)	2.2	0.8 to 3.7
Level 3 10.0 ng/mL (34.9 nmol/L)	2.2	1.4 to 2.9	Level 3 10.0 ng/mL (34.9 nmol/L)	1.5	0.9 to 2.1
Dilution 1 13.0 ng/mL (45.4 nmol/L)	−0.5	−1.5 to 0.5		–	–
Dilution 2 20.0 ng/mL (69.8 nmol/L)	−1.0	−1.5 to −0.6		–	–

The mean bias and corresponding confidence intervals were calculated using the individual sample biases of n=6 preparations for serum and plasma samples and n=10 preparations for sample dilutions. Conversion factor ng/mL to nmol/L: 3.49. CI, confidence interval.

Dilution integrity was evaluated using two spiked samples at concentration levels of approximately 13.0 ng/mL and 20.0 ng/mL (45.4 nmol/L and 69.8 nmol/L). The mean deviation was found to be within −0.5 % and −1.0 % (Table 2). Overall, the method demonstrated no statistically significant bias and showed independence from the matrix used. The data presented fulfill the APS as described above.

Sample stability

Processed samples remained stable for 14 days at 8 °C, with average recoveries between 98 and 101 %. Furthermore, the stability of spike solutions and spiked frozen control materials stored at −80 °C was assessed over a 12-week period. Average recoveries for spike solutions ranged from 107 to 109 %, while for spiked matrix based materials, they ranged from 96 to 106 %.

Equivalence of results between independent laboratories

The method comparison study involved 76 samples that were measured at two separate laboratories (Site 1: Roche Diagnostics GmbH, Penzberg and Site 2: Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen). These included 36 unaltered native leftover patient samples, 20 pools, 10 spiked native samples and 10 diluted samples. One pooled sample was identified as an outlier using the LORELIA outlier test [18] and was therefore excluded from the analysis.

The regression analysis resulted in a regression equation with a slope of 0.98 (95 % Cl 0.97 to 0.99) and an intercept of 0.000 (95 % Cl −0.003 to 0.006). The Pearson correlation value was r≥0.999. The interlaboratory measurement bias was −0.9 % (95 % CI interval from −2.8 to 1.0 %) and the 2S interval of the relative difference

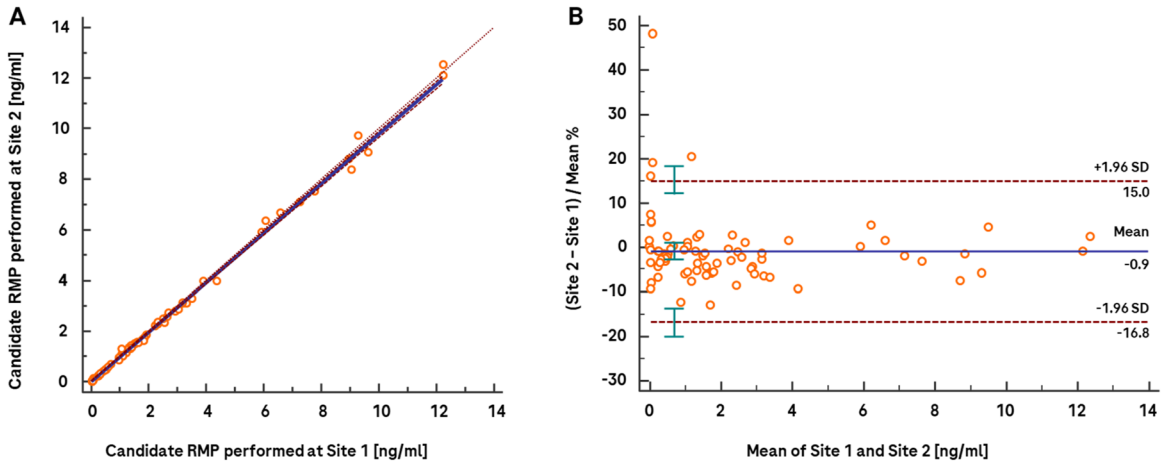

Figure 3: Results from the patient sample-based androstenedione method comparison study performed between two independent laboratories. (A) Passing–Bablok regression plot including the Pearson regression analysis for the method comparison study of the RMP (n=75) samples between the independent laboratories (Laboratory 1: Roche Diagnostics GmbH, Penzberg and Laboratory 2: Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen). The regression analysis resulted in a regression equation with a slope of 0.98 (95 % CI 0.97 to 0.99) and an intercept of 0.000 (95 % CI −0.003 to 0.006). The Pearson correlation value was≥0.999. (B) Bland–Altman plot for the method comparison study of the RMP (n=75) between two independent laboratories (Laboratory 1 and Laboratory 2). The interlaboratory measurement bias was −0.9 % (95 % CI interval from −2.8 to 1.0 %) and the 2S interval of the relative difference was ± 15.9 % (lower limit CI interval from −20.0 to −13.6 %, upper limit CI interval from 11.8 to 18.2 %). CI, confidence interval; RMP, reference measurement procedure; SD, standard deviation.

was ± 15.9 % (lower limit CI interval from −20.0 to −13.6 %, upper limit CI interval from 11.8 to 18.2 %) (Figure 3). The 3-day precision experiment at Site 2 showed comparable coefficient of variations (CVs) to those observed at Site 1, with intermediate precision of 1.8–3.8 % and repeatability of 0.5–3.7 %. Two results had to be excluded due to measurement issues for the highest level. Both the data scatter and data bias prove that the proposed RMP is transferable between independent laboratories.

A second method comparison study was conducted between the RMP (performed at Site 1) and an in-house developed LC-MS/MS assay performed at Site 3 (Institute of Laboratory Medicine at Leipzig University Hospital) [9]. In total, 76 samples were measured, whereas 13 samples were found below the LOQ of the in-house developed LC-MS/MS assay. Therefore, the analysis included 63 samples.

Performing a Passing Bablok regression, the method yielded a regression equation with a slope of 0.92 (95 % CI 0.90 to 0.98) and an intercept of 0.002 (95 % CI −0.065 to 0.035). The correlation coefficient was r≥0.993. Furthermore, the Bland-Altman analysis demonstrated a bias of −6.9 % (95 % CI −9.6 to −4.2 %), which was found to be statistically different from zero (Figure 4). This indicates the presence of systematic error or bias in the measurements, contributing to the overall imprecision of the study. Additionally, the 2S interval of the relative difference, with a range of ± 21.2 % (95 % CI −32.7 to −23.4 % and 9.6–18.9 %).

Equivalence of results between the candidate and the JCTLM listed RMP

Equivalence between the candidate RMP and the JCTLM-listed RMP was shown by measuring 20 anonymized native patient samples covering the measuring range. During the analysis, one sample was found to be below the quantification limit of the JCTLM-listed RMP and was subsequently excluded from the dataset. The subsequent data analysis using Passing-Bablok revealed a high level of agreement, as evidenced by the regression equation with a slope of 0.98 (95 % CI 0.93 to 1.05) and an intercept of −0.01 (95 % CI −0.06 to 0.03), along with a strong correlation coefficient of r≥0.997. The measurement bias was −2.8 % (95 % CI interval from −5.2 to −0.4 %) and the 2S interval of the relative difference was ± 9.8 % (lower limit CI interval from −16.7 to −8.4 %, upper limit CI interval from 2.7 to 11.0 %) (Figure 5).

Naturally, both measurement procedures include inherent measurement errors, and the error in the calculated difference between the two procedures is consequently larger. Each RMP must meet the stricter performance specifications of CV_{ref}≤2.9 % and B_{ref}≤2.7 %. According to the total error concept, the maximum allowable difference is calculated as $2B_{ref} + 1.96 * sqrt(CV_{ref}^2 + CV_{ref}^2)$, resulting in 13.4 %. The presented results align with the pre-determined acceptance criteria and the accuracy and intermediate precision data from individual methods.

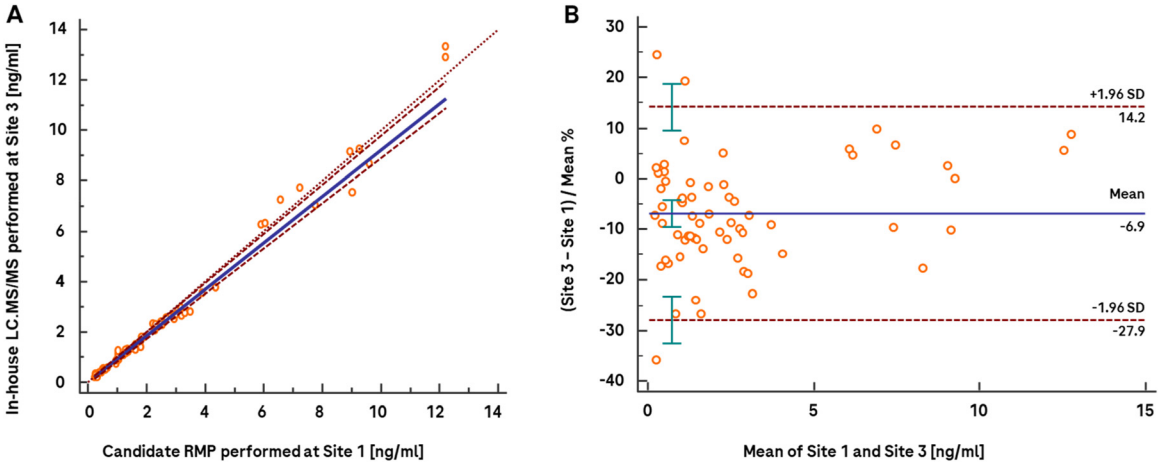

Figure 4: Results from the patient sample-based androstenedione method comparison study performed between the candidate RMP and a well-established routine RMP. (A) Passing–Bablok regression plot including the Pearson regression analysis for the method comparison study of the RMP (n=63) samples between the independent laboratories (Laboratory 1: Roche Diagnostics GmbH, Penzberg and Laboratory 3: Site 3: Institute of Laboratory Medicine at Leipzig University Hospital). The regression analysis resulted in a slope of 0.92 (95 % CI 0.90 to 0.98) and an intercept of 0.002 (95 % CI −0.065 to 0.035). The correlation coefficient was r≥0.993. Bland-Altman analysis demonstrated a bias of −6.9 % (95 % CI −9.6 to −4.2 %) and the 2S interval of the relative difference was ± 21.2 % (95 % CI −32.7 to −23.4 % and 9.6 to 18.9 %

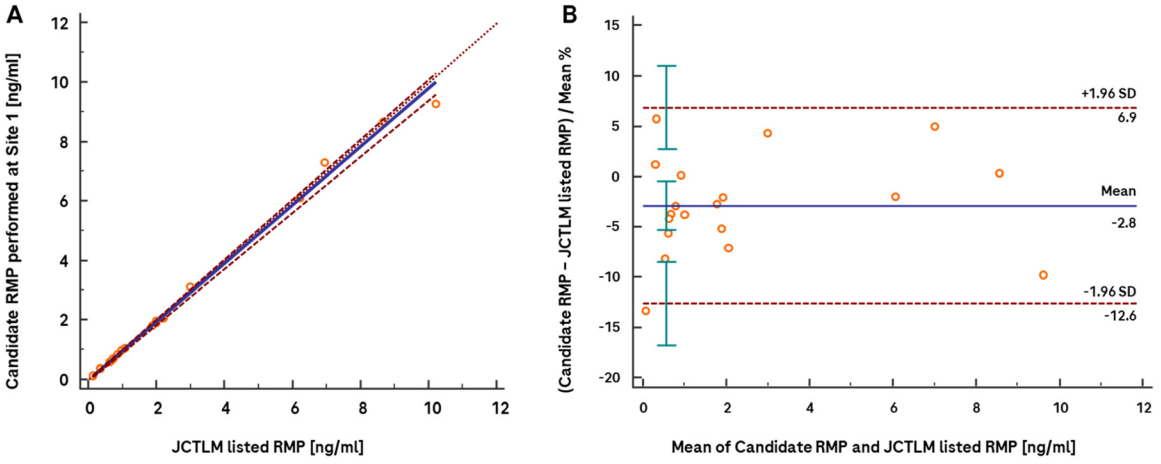

Figure 5: Results from equivalence testing between the JCTLM listed and the candidate RMP. (A) Passing–Bablok regression plot including the Pearson regression analysis for the method comparison study of the RMP (n=19) samples between the JCTLM listed and the candidate RMP. The regression analysis resulted in a regression equation with a slope of 0.98 (95 % CI 0.93 to 1.05) and an intercept of −0.01 (95 % CI −0.06 to 0.03). The correlation coefficient was r≥0.997. (B) Bland–Altman plot for the method comparison study of the RMP (n=19) between the JCTLM listed and the candidate RMP. The measurement bias was −2.8 % (95 % CI interval from −5.2 to −0.4 %) and the 2S interval of the relative difference was ± 9.8 % (lower limit CI interval from −16.7 to −8.4 %, upper limit CI interval from 2.7 to 11.0 %). CI, confidence interval; JCTLM, joint committee on traceability in laboratory medicine RMP, reference measurement procedure; SD, standard deviation.

Estimation of measurement uncertainty (MU)

The measurement process for single measurements exhibited an overall uncertainty ranging from 1.4 to 2.1 % regardless of the sample concentration, as indicated in Table 3. To account for a confidence level of ~95 % assuming a normal distribution, the resulting overall uncertainty was multiplied by a coverage factor of k=2. As a result, the expanded measurement uncertainties fall within the range of 2.7–4.1 %.

For the establishment of reference or target values, multiple sample preparations were carried out for each sample on at least two different days. The average of these results, calculated using n=6, was used. The overall uncertainties were found to range from 0.8 to 1.6 %, with expanded uncertainties ranging from 1.7 to 3.1 % (k=2), as shown in Table 4.

Table 3: Exemplary overview of measurement uncertainty for androstenedione quantification with the candidate RMP in serum samples for single measurements.

	Level				
	Level 1 0.024 ng/mL (0.0838 nmol/L)	**Level 2 5.00 ng/mL (17.5 nmol/L)**	**Level 3 10.0 ng/mL (34.9 nmol/L)**	**Patient sample 1 0.874 ng/mL (3.05 nmol/L)**	**Patient sample 2 1.48 ng/mL (5.17 nmol/L)**
Type B uncertainty	**0.8**	**0.7**	**0.6**	**0.7**	**0.7**
Calibrator preparation, CV, %					
Characterization of reference material	0.15	0.15	0.15	0.15	0.15
Preparation of					
Stock solution	0.16	0.16	0.16	0.16	0.16
Working solution	0.34	–	–	–	–
Spike solution	0.61	0.45	0.36	0.53	0.53
Matrix based calibrator	0.80	0.68	0.62	0.74	0.74
Type A uncertainty	**1.9**	**1.2**	**1.8**	**1.9**	**1.5**
Intermediate precision, CV, %					
Total measurement uncertainty (k=1), CV, %	**2.1**	**1.4**	**1.9**	**2.1**	**1.7**
Expanded measurement uncertainty (k=2), CV, %	**4.1**	**2.7**	**3.7**	**4.1**	**3.4**

Conversion factor ng/mL to nmol/L: 3.49. CV, coefficient of variation. The measurement uncertainty of the whole approach for a single measurement estimated as a combination of the uncertainty of calibrator preparation (type B uncertainty) and uncertainty of the precision experiment (type A uncertainty) are given in bold.

Table 4: Exemplary overview of measurement uncertainty for androstenedione target value assignment (n=6) with the candidate RMP in serum samples.

	Level				
	Level 1 0.024 ng/mL (0.0838 nmol/L)	**Level 2 5.00 ng/mL (17.5 nmol/L)**	**Level 3 10.0 ng/mL (34.9 nmol/L)**	**Patient sample 1 0.874 ng/mL (3.05 nmol/L)**	**Patient sample 2 1.48 ng/mL (5.17 nmol/L)**
Type B uncertainty	**0.8**	**0.7**	**0.6**	**0.7**	**0.7**
Calibrator preparation, CV, %					
Characterization of reference material	0.15	0.15	0.15	0.15	0.15
Preparation of					
Stock solution	0.16	0.16	0.16	0.16	0.16
Working solution	0.34	–	–	–	–
Spike solution	0.61	0.45	0.36	0.53	0.53
Matrix based calibrator	0.80	0.68	0.62	0.74	0.74
Type A uncertainty	**1.0**	**0.7**	**0.6**	**1.4**	**0.7**
Intermediate precision, CV, %					
Total measurement uncertainty (k=1), CV, %	**1.3**	**1.0**	**0.8**	**1.6**	**1.0**
Expanded measurement uncertainty (k=2), CV, %	**2.5**	**1.9**	**1.7**	**3.1**	**2.1**

Conversion factor ng/mL to nmol/L: 3.49. CV, coefficient of variation. The measurement uncertainty of the whole approach for a target value assignmet estimated as a combination of the uncertainty of calibrator preparation (type B uncertainty) and uncertainty of the precision experiment (type A uncertainty) are given in bold.

Discussion

We have developed a high-quality candidate RMP for the quantification of ASD that meets all the predefined acceptance criteria required for a higher-order RMP.

The candidate Reference Measurement Procedure (RMP) leverages two-dimensional heart-cut chromatography combined with orthogonal stationary and mobile phases, which enhances selectivity and specificity. In contrast, the JCTLM-listed RMP, while demonstrating good selectivity, does not

utilize this advanced chromatographic technique. Additionally, the candidate RMP offers significant enhancements in sensitivity compared to the previous method. The candidate RMP enables the quantification of ASD over a broader measurement range (0.00800–12.0 ng/mL [0.0279–41.9 nmol/L]) compared to the previous method's range of 0.05–12 ng/mL [0.175–41.9 nmol/L]. This improvement in the lower measurement range was achieved through refined sample preparation protocol, allowing for more accurate and sensitive detection of ASD in patient samples. Furthermore, the new candidate RMP was refined to further decrease measurement uncertainty. As a result, significant effort was devoted to establishing an optimal calibration and control scheme, along with reducing method variability. The expanded measurement uncertainties (k=2) for individual measurements ranged from 3.4 to 4.1 %, whereas the measurement uncertainties of the JCTLM-listed RMP varied from 6.6 to 12.9 % (k=2.23) [10]. For both RMPs these uncertainties can be further reduced by performing multiple measurements.

Equivalence to the JCLTM-listed ID-LC-MS/MS RMP was demonstrated by measuring anonymized native patient samples covering the measuring range and showed a good agreement between the listed RMP and the new candidate RMP. In addition, a method comparison study and a precision experiment confirmed the transferability of the RMP to a second independent laboratory, demonstrating its robustness and ease of implementation. Additionally, adapting the sample preparation protocol to a 96-well plate format enhances sample throughput, enabling large method comparison studies to be conducted within a reasonable time frame.

The method comparison study between the RMP and the in-house LC-MS/MS-based assay demonstrated a good correlation with an expected scattering, despite observing a negative deviation. The assay was calibrated using a commercially available kit that lacked information on the absolute content of the analyte and traceability to SI units at the time the study was conducted. This lack of traceability could introduce bias and imprecision in the calibration process, potentially affecting measurement accuracy and contributing to the observed discrepancies in the regression analyses.

Conclusions

In this paper, a modernized LC-MS/MS-based candidate RMP for the quantification of ASD in human serum and plasma is presented, with a higher sensitivity and an expanded measuring range from 0.00800 ng/mL to 12.0 ng/mL (0.0279 nmol/L to 41.9 nmol/L). This RMP serves as a traceable and reliable platform for the standardization of routine assays and the evaluation of clinically relevant samples.

Acknowledgments: We would like to thank Aline Hoffmeister, Monika Kriner, and Michael Dedio for their support in selecting and providing samples. Editorial support, under the direction of the authors, was provided by Graziella Greco and Jade Drummond of inScience Communications, Springer Healthcare Ltd, UK, and was funded by Roche Diagnostics GmbH. (Penzberg, Germany).

Research ethics: All procedures were in accordance with the Helsinki Declaration. All samples used were exclusively anonymized samples.

Informed consent: Not applicable.

Author contributions: All authors have contributed to the manuscript conception and design; acquisition, or analysis and interpretation of data; drafting or revision; and final approval of the published article.

Use of Large Language Models, AI and Machine Learning Tools: Roche Chat, Roche's artificial intelligence (AI) Technology, was used to improve the language of the manuscript.

Conflict of interest: Myriam Ott, Andrea Geistaner, Judith Taibon, are all employees of Roche Diagnostics GmbH. Vanessa Fischer was an employee of Roche Diagnostics GmbH at the time the study was conducted. Friederike Bauland is an employee of Chrestos Concept GmbH & Co. KG, Essen, Germany. Manfred Rauh and Daniel Köppl are employees of the Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen. Uta Ceglarek and Alexander Gaudl are employees of the Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig Medical Center. Roche employees holding Roche non-voting equity securities (Genussscheine): Judith Taibon, Myriam Ott, Andrea Geistanger.

Research funding: This research was funded by Roche Diagnostics GmbH. Manfred Rauh and Uta Ceglarek with team are funded cooperation partners of Roche Diagnostics GmbH

Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Zhang S, Lo JC, Jaffe MG, Arzumanyan H. Elevated serum androstenedione level in a patient with ectopic adrenocorticotropic hormone syndrome. AACE Clin Case Rep 2023;9:142–5.
2. Al-Kofahi M, Ahmed MA, Jaber MM, Tran TN, Willis BA, Zimmerman CL, et al. An integrated PK-PD model for cortisol and the 17-hydroxyprogesterone and androstenedione biomarkers in children with congenital adrenal hyperplasia. Br J Clin Pharmacol 2021;87:1098–110.

3. Lerchbaum E, Schwetz V, Rabe T, Giuliani A, Obermayer-Pietsch B. Hyperandrogenemia in polycystic ovary syndrome: exploration of the role of free testosterone and androstenedione in metabolic phenotype. PLoS One 2014;9:e108263.
4. Wei R, Bowers K, Kroner GM, Payto D, Colón-Franco JM. The Androstenedione Roche Elecsys immunoassay has superior comparability to the LC-MS/MS assay than the Siemens Immulite immunoassay. Pract Lab Med 2022;31:e00279.
5. Owen WE, Roberts WL. Performance characteristics of the IMMULITE 2000 androstenedione assay. Clin Chim Acta 2007;383:168–71.
6. Xu W, Li H, Guan Q, Shen Y, Cheng L. A rapid and simple liquid chromatography-tandem mass spectrometry method for the measurement of testosterone, androstenedione, and dehydroepiandrosterone in human serum. J Clin Lab Anal 2017;31. https://doi.org/10.1002/jcla.22102.
7. Kushnir MM, Blamires T, Rockwood AL, Roberts WL, Yue B, Erdogan E, et al. Liquid chromatography–tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone, and testosterone with pediatric and adult reference intervals. Clin Chem 2010;56:1138–47.
8. Tournier M, Pouech C, Quignot N, Lafay F, Wiest L, Lemazurier E, et al. Determination of endocrine disruptors and endogenic androgens and estrogens in rat serum by high-performance liquid chromatography–tandem mass spectrometry. Steroids 2015;104:252–62.
9. 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.
10. Gradl K, Taibon J, Singh N, Albrecht E, Geistanger A, Pongratz S, et al. An isotope dilution LC-MS/MS-based candidate reference method for the quantification of androstenedione in human serum and plasma. Clin Mass Spectrom 2020;16:1–10.
11. Joint committee on traceability in laboratory medicine (JCTLM). https://www.jctlm.org/ [Accessed 15 Jun 2024].
12. Joint committee on traceability in laboratory medicine (JCTLM). Androstenedione RMP. https://www.jctlmdb.org/#/app/search/search-results?keywords=androstenedione&code_status=P [Accessed 17 Jun 2024].
13. Clinical and Laboratory Standards Institute (CLSI). C62-A. Liquid chromatography-mass spectrometry methods; approved guideline. In: CLSI document C62-A. Wayne, PA: Clinical and Laboratory Standard Institute; 2014.
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). https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6 [Accessed 15 Jun 2024].
15. European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). EFLM biological variation database – androstenedione. https://biologicalvariation.eu/search?query=Androstenedione [Accessed 19 Nov 2024].
16. Braga F, Panteghini M. Performance specifications for measurement uncertainty of common biochemical measurands according to Milan models. Clin Chem Lab Med 2021;59:1362–8.
17. VCA. Variance component analysis. https://cran.r-project.org/web/packages/VCA/VCA.pdf [Accessed 15 Jun 2024].
18. Rauch G, Geistanger A, Timm J. A new outlier identification test for method comparison studies based on robust regression. J Biopharm Stat 2011;21:151–69.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2024-1135).