Supplemental appendix

The saline infusion test for primary aldosteronism: implications of immunoassay inaccuracy

Graeme Eisenhofer,[1,2] Max Kurlbaum,[3] Mirko Peitzsch,[2] Georgiana Constantinescu,[1] Hanna Remde,[3] Manuel Schulze,[4] Denise Kaden, [2] Lisa Marie Müller,[5] Carmina Fuss,[3] Sonja Kunz,[5] Sylwia Kołodziejczyk-Kruk,[6] Sven Gruber,[7] Aleksander Prejbisz,[6] Felix Beuschlein,[5,7] Tracy Ann Williams,[5,8] Martin Reincke,[5] Jacques W.M. Lenders,[1,9] and Martin Bidlingmaier[5]

[1]Department of Medicine III, University Hospital Carl Gustav Carus, Technische Universität Dresden, Germany; [2]Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Carl Gustav Carus, Technische Universität Dresden, Germany; [3]Department of Internal Medicine I, Division of Endocrinology and Diabetes, University Hospital, University of Würzburg, Würzburg, Germany; [4]Center for Information Services and High Performance Computing, Technische Universität Dresden, Germany; [5]Department of Medicine IV, University Hospital, Ludwig Maximilian University of Munich, Munich, Germany; [6]Department of Hypertension, National Institute of Cardiology, Warsaw, Poland; [7]Department of Endocrinology, Diabetology and Clinical Nutrition, University Hospital Zurich (USZ) and University of Zurich (UZH), Zurich, Switzerland; [8]Division of Internal Medicine and Hypertension, Department of Medical Sciences, University of Turin, Turin, Italy; [9]Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands.

Contents	Page

Supplemental appendix

Supplemental Methods

Overview of the PROSALDO protocol

The overall objective of the PROSALDO study (**PRO**spective study on the diagnostic value of **S**teroid profiling in primary **ALDO**steronism) is to evaluate the combination of steroid profiling with machine learning for improved diagnosis and stratification of patients with suspected primary aldosteronism (PA). The study is a registered international multicenter trial (trial registration no: DRKS00017084) that began in October 2019 and aims to recruit 1500 patients tested for PA, among whom about one third will be expected to have the disease. As of the end of September 2021 total patient accrual stood at 550.

Inclusion of patients into the trial requires a diagnosis of hypertension, with exclusion of other forms of secondary hypertension and with suspicion of PA based on several criteria: 1. office blood pressure (BP) above 150/100 mmHg on two separate visits; 2. therapy resistant hypertension (office systolic BP $\geq$ 140 mmHg and/or diastolic BP $\geq$90 mmHg) with at least 3 different antihypertensives, including one diuretic; 3. hypertension and spontaneous (K+ $\leq$ 3.5 mmol/L) or diuretic-induced hypokalemia (K+ $\leq$ 3.0); 4. hypertension and incidentaloma; 5. hypertension with family history (<40 years age) of early onset hypertension or hemorrhagic stroke; 6. hypertension and obstructive sleep apnea; and 7. hypertension and a first degree relative with history of PA.

The trial follows procedures covered by Endocrine Society guidelines (1), with initial screening using the aldosterone to renin ratio (ARR) followed by confirmatory testing, in particular the saline infusion test (SIT), and in patients with a positive SIT progression to adrenal venous sampling (AVS). AVS is performed without corticotropin stimulation. Inclusion into AVS studies requires willingness to undergo adrenalectomy when AVS reveals lateralization of aldosterone secretion from a single adrenal (i.e., lateralization ratio > 4.0). Adrenalectomy without AVS or selective sampling is allowed for in occasional patients with a defined solitary adrenal mass and who satisfy other criteria (e.g., <35 years of age). The protocol calls for patients to undergo follow-up at 6-12 months after adrenalectomy. Follow-up observes the procedures outlined to assess post-operative outcomes of adrenalectomy, according to the Primary Aldosteronism Surgical Outcome (PASO) criteria (2). Similar procedures apply to patients who remain unoperated or in who PA is excluded by negative results of screening or the SIT, with follow at 6-12 months after the last diagnostic procedure.

Importantly, although the study depends on results of routine laboratory tests for diagnostic decision-making and patient flow through the protocol, it also provides for results of steroid profiling at initial screening and during AVS to define respective need for confirmatory tests and surgical intervention. Decision making based on steroid profiles from analyses of screening and baseline blood samples before the SIT, is according to machine learning algorithms established by the SPISCA (Steroid Profiling for Identification and Subtype Classification of primary Aldosteronism) study as published previously (3). These machine learning algorithms not only provide probabilities of the likelihood of PA versus primary hypertension, but also probabilities of unilateral disease due to somatic mutations of *KCNJ5*. The primary objective of the PROSALDO trial is to use this information to both diagnose PA and identify those patients with PA who have somatic *KCNJ5* mutations and who may undergo adrenalectomy without need for confirmatory testing or AVS providing imaging evidence of a solitary adrenal mass. Diagnostic decision-making based on machine learning involves three algorithms with initial cut-offs used in the trial dependent on mean probabilities indicating PA or PA subtypes above 50% and supported by at least two of the three algorithms.

Background to the present PROSALDO sub-study

Although patient accrual into the PROSALDO trial is incomplete (as of October 2021), the protocol does allow for sub-studies that do not depend on completion of the trial. The presently described sub-study, however, was not one of those originally planned as part of the PROSALDO trial and was only initiated after identification of severe interference in immunoassay measurements of aldosterone in one of the first patients enrolled into the protocol (4). Highly discordant results between immunoassay and LC-MS/MS measurements of plasma aldosterone in that index case led investigators of the PROSALDO trial to consider the possibility of a wider problem impacting both analytical accuracy of immunoassay measurements of aldosterone and subsequence inaccuracy of diagnostic procedures

dependent on those measurements. The presently described prospective diagnostic cohort sub-study was established to explore this possibility. The sub-study involved five European centers: 1. University Hospital Carl Gustav Carus Dresden, Dresden (DR), Germany; 2. University Hospital of Würzburg, Würzburg (WU), Germany; 3. Institute of Cardiology, Warsaw (WW), Poland; 4. University Hospital of Munich, Munich (MU), Germany; and 5. University Hospital of Zürich, Zurich (ZH), Switzerland.
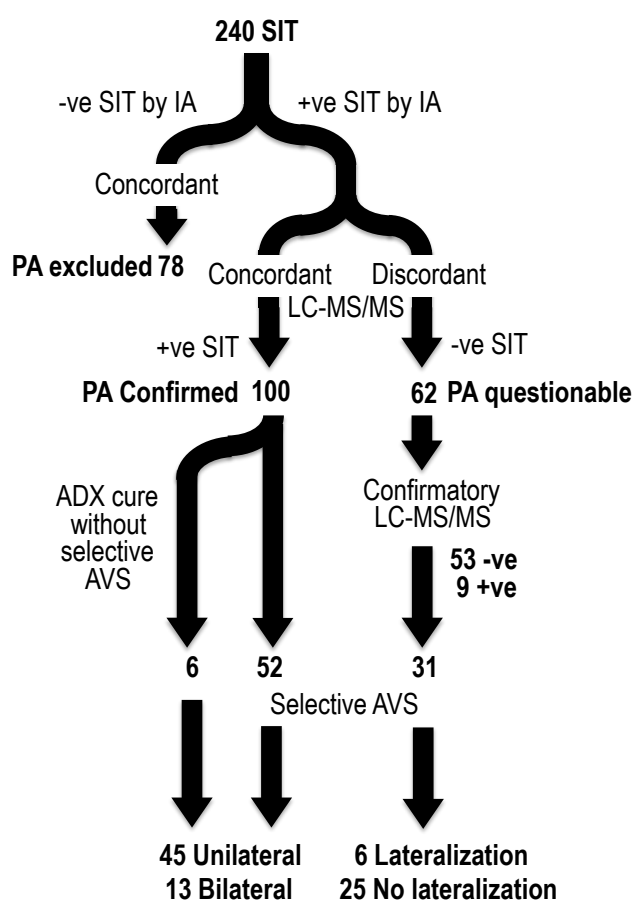
Patient study flow, data capture and reporting

Inclusion of patients into the present sub-study of the PROSALDO trial required a SIT and measurements of aldosterone by both the routine immunoassay used at the participating center and by LC-MS/MS at the center in DR responsible for steroid profiling. As early of October 2021 there were 240 patients tested for PA who satisfied this requirement (Supplemental figure 1).

Measurements of aldosterone by LC-MS/MS invariably follow those by immunoassay, with some delay particularly for the four centers outside of DR from which patient specimens must be shipped on dry ice to DR. Specimens received at the DR laboratory undergo LC-MS/MS measurements within 10 days of receipt. Results of these measurements, including machine learning interpretations of steroid profiles, are uploaded into a RedCap (Research Electronic Data Capture) system immediately after final analysis of mass spectra. After expert review and validation of reports, investigators at study centers are notified by automated email of the availability of validated reports for download as PDF files.

The RedCap system is a secure, web-based software platform designed to support data capture for research studies by way of several features: 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources (5,6).

For initial immunoassays of aldosterone, three centers (MU, ZH and WW) employ the Diasorin Liaison chemiluminescence immunoassay and two centers (DR, WU)

Supplemental figure 1. Patient flow through the study. Abbreviations: SIT, saline infusion test; IA, immunoassay; LC-MS/MS liquid chromatography with tandem mass spectrometry; AVS adrenal venous sampling; ADX, adrenalectomy

employ the Immuno Diagnostic Systems iSYS chemiluminescence immunoassay. The cut-offs for positive (pathologic) versus negative (non-pathologic) results for plasma aldosterone concentrations after the SIT for immunoassays were set at 170 pmol/L (61 ng/L), as defined by Endocrine Society guidelines (1), while the cut-off for LC-MS/MS measurements was defined at 162 pmol/L (58 ng/L), as established previously for mass spectrometry-based measurements (7). Concordance of immunoassay and LC-MS/MS results was defined by results for both measurements falling below or above those cut-offs. Thereby among the 240 patients with SIT results for both measurements, concordantly negative or non-pathologic results were returned in 78 patients while concordantly positive or pathologic results were returned in 100 patients (Supplemental figure 1). The remaining 62

patients (26%) all returned discordant positive results of the SIT by immunoassay measurements and negative results by LC-MS/MS measurements. For those 62 patients with discordant results, additional confirmatory LC-MS/MS measurements were performed in all patients, among whom 53 (85%) had repeatedly negative results for the SIT.

As of the end of early October 2021, 52 of the 100 patients with concordant positive SIT results had undergone AVS with selective sampling compared to 31 of the 62 patients with discordant results and selective AVS (Supplemental figure 1). An additional six patients with concordant results of the SIT, but in who AVS was not selective or was not carried out, underwent adrenalectomy on the basis of a solitary adrenal mass indicated by computed tomography, highly pathogenic biochemical test results and in one case also young age. Since all six patients showed complete biochemical cure at 6-12 months of follow-up, these patients were included with the 52 who had concordant positive results and selective AVS. Among those 58 patients, 45 were determined to have lateralized aldosterone secretion or unilateral PA, while 13 had no AVS-based evidence of lateralized aldosterone secretion. This was in contrast to the 31 patients with discordant results in whom the diagnosis of PA was questionable, six of whom showed lateralized aldosterone secretion and 25 no evidence of lateralized aldosterone secretion according to AVS.

Assay precision and limits of quantification
As also outlined earlier, initial measurements of plasma aldosterone involved an in-house LC-MS/MS method developed at DR for measurements of multiple steroids and either the DiaSorin Liaison immunoassay or the Immunodiagnostic Systems iSYS immunoassay at all centers including DR. Thereafter, samples from the 62 patients with discordant results for the SIT underwent further measurements of aldosterone, first using an independent Chromsystems LC-MS/MS method at either WU or MU and by an additional immunoassay according to which ever method was not used for initial measurements. Lower limits of quantification (LOQ), which reflect the concentration of aldosterone above which intra-assay coefficients of variation are less than 20%, varied from as low as 22 pmol/L for the in-house LC-MS/MS method at DR to 103 pmol/L for the iSYS immunoassay at WU and MU (Supplemental table 1).
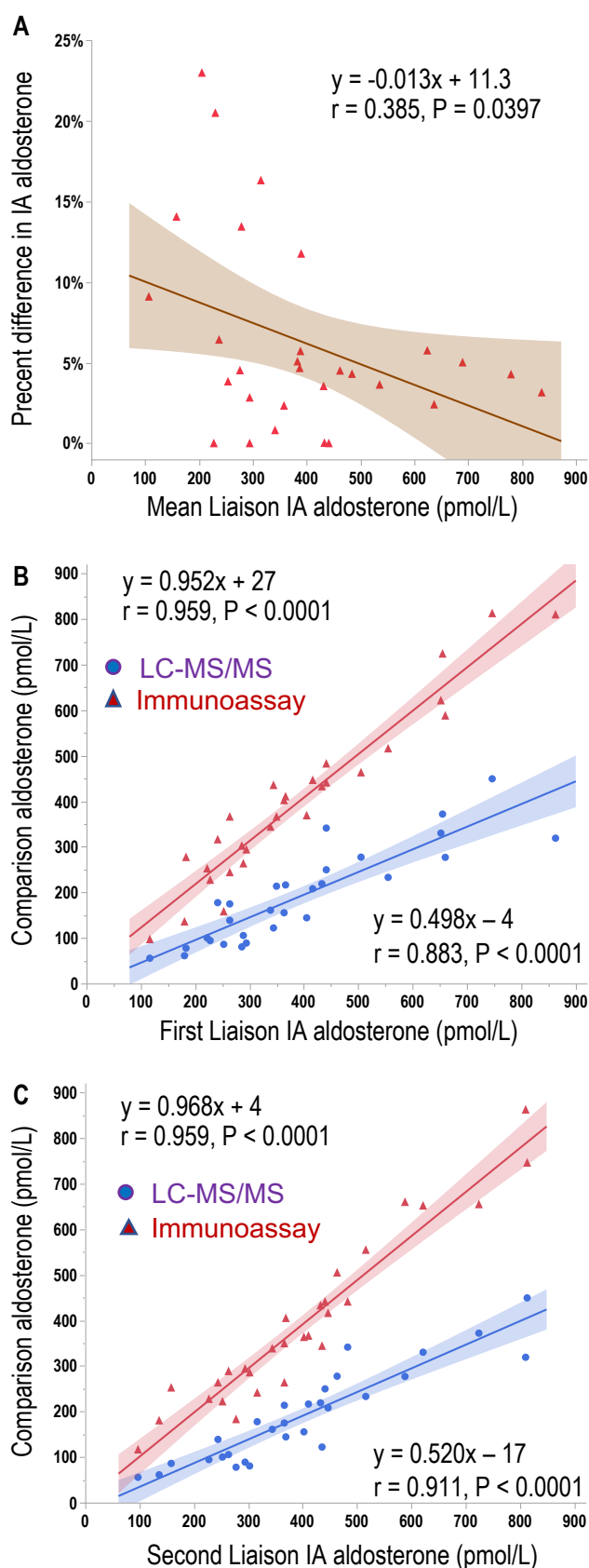
Supplemental table 1. Lower limits of quantification (LOQ) and interassay coefficients of variation (CV) for plasma measurements of aldosterone according to assay method and concentration (Conc)

| Assay method | LOQ* | Low | | Mid | | High | |
| Location and method | Conc | Conc | CV | Conc | CV | Conc | CV |
	pmol/L	pmol/L	%	pmol/L	%	pmol/L	%
LC-MS/MS							
DR - In-house	22	252	9.6	655	6.7	16091	5.9
WU – Chromsystems	78	272	3.2	660	2.5	2586	3.0
MU – Chromsystems	78	272	8.2	660	3.8	2586	4.5
Immunoassay							
DR – iSYS	64	360	5.7	665	3.8	1706	4.5
WU – iSYS	103	332	10.6	693	6.4	1632	6.4
MU – iSYS	103	164	12.0	ND	ND	1967	3.0
MU – Liaison	96	111	11.4	ND	ND	1568	2.4
ZU – Liaison	53	150	8.6	654	6.6	ND	ND
WW – Liason	53	164	10.5	513	6.9	2292	6.3

*LOQ defined as the minimal plasma concentration for which the intra-assay coefficient of variation remains above 20%. Abbreviations: DR, Dresden; WU, Würzburg; MU, Munich; ZU, Zurich; WW, Warsaw; ND, no data.

Inter-assay coefficients of variation, as expected, were generally highest in the lower concentration range compared to middle and high concentration ranges for aldosterone (Supplemental table 1).

A

y = -0.013x + 11.3
r = 0.385, P = 0.0397

B

y = 0.952x + 27
r = 0.959, P < 0.0001

● LC-MS/MS
▲ Immunoassay

y = 0.498x – 4
r = 0.883, P < 0.0001

C

y = 0.968x + 4
r = 0.959, P < 0.0001

● LC-MS/MS
▲ Immunoassay

y = 0.520x – 17
r = 0.911, P < 0.0001

Intra-assay coefficients of variation were as expected invariably lower than inter-assay coefficients (data not shown).

Measurements of plasma aldosterone in 29 samples from 10 patients with discordant results between the Liaison immunoassay and the DR LC-MS/MS method were repeated using the Liaison immunoassay. These constituted samples at screening and before and after the SIT. As expected, precision of repeat measurements, as assessed from the percent difference of the two immunoassay measurements from the mean, was related to mean concentrations such that higher percent differences (lower precision) were observed at lower concentrations (Supplemental figure 2A). The mean percent difference (6.4%) was similar to the intra-assay coefficients of variation reported by all three centers using this immunoassay for quality control samples in the mid to low range (Supplemental table 1).

Comparisons of the concentrations of plasma aldosterone for the first and second measurements by the Liaison immunoassay showed close to one-to-one agreement, further confirming adequate reproducibility of these measurements (Supplemental figure 2B&C). In contrast to the relationships for the two immunoassay measurements, relationships of both first and second immunoassay measurements with LC-MS/MS measurements of aldosterone indicated consistently and considerably lower plasma concentrations of aldosterone by LC-MS/MS than by the Liaison immunoassay. The differences in slopes indicated approximately 2-fold higher concentrations measured by the Liaison immunoassay than by LC-MS/MS. Furthermore, the similarities in the differences in these slopes for the first immunoassay compared to the second immunoassay clarified that the extent of discordance was reproducible according to the two separate immunoassay measurements.

Supplemental figure 2. Precision and reproducibility of the Liaison immunoassay (A) and reproducibility of discordance with LC-MS/MS measurements according to repeated measurements of plasma aldosterone by the Liaison immunoassay (B & C). Panel A illustrates the relationship of mean aldosterone concentrations with the percent differences from the mean. Panels B and C illustrate relationships of the first (A) and second (B) immunoassay measurements with LC-MS/MS measurements of aldosterone (●) and respective second and first immunoassay measurements of aldosterone (▲).

Supplemental appendix

Investigation of immunoassay interference

Twenty-one plasma samples (three separate samples from seven patients) — including samples with discordant result for the SIT by immunoassay and LC-MS/MS measurements in each of the seven patients — were identified in biobanks with sufficient remaining specimen for investigation of immunoassay interference. This investigation followed a previously published procedure in which plasma was processed using solid phase purification (8). This procedure is the same as that used to prepare plasma specimens for LC-MS/MS measurements of aldosterone and other steroids. While allowing measurements of steroids and other low molecular weight molecules, the procedure removes proteins and other macromolecules, including potential heterophile antibody interferents.

After SpeedVac-facilitated evaporative dry-down of purified eluants, lyophilized specimens were shipped to the laboratory of MU partners where samples were reconstituted in assay buffer and subject to immunoassay by both Liaison and iSYS methods. Corrections for recovery of extracted samples were achieved by spiked additions of aldosterone to other aliquots of the same sample. Immunoassay measurements of purified samples were then compared to measurements of the original plasma specimens by the same immunoassay and by LC-MS/MS. As a negative control, three sets of low, mid and high range quality control samples were also extracted with subsequent concentrations measured by the Liaison immunoassay (311, 1245, 12895 pmol/L) matching to those measured by LC-MS/MS (261, 1068 and 13201 pmol/L).

Patient follow-up

Among the 78 patients with concordant negative test results of the SIT, six patients underwent follow-up that was compliant with protocol requirements and after at least 6 months following negative test results of the SIT (Supplemental table 2). All six patients were confirmed not to have PA at that follow-up on the basis of a negative ARR.

Among the 100 patients with concordant positive test results for the SIT, as of 28/10/21 there were 35 who underwent adrenalectomy, including 29 with lateralized aldosterone secretion and another six as outlined above in whom adrenalectomy was based on other considerations. Among those 35 patients who underwent adrenalectomy, follow-up of more than 6 months after surgical intervention was carried out in 19 patients according to procedures that were compliant with PASO criteria and protocol requirements (Supplemental table 2). In 17 of these patients, follow-up confirmed biochemical cure. For the other two patients there remains need for a SIT to clarify biochemical cure. Among the other patients without lateralization, follow-up has been carried out in five patients, but not in a manner compliant with protocol requirements that would enable confirmation or exclusion of PA beyond initial routine diagnostic tests, in particular the SIT.

Among the 62 patients with discordant negative and positive test results for the SIT by respective LC-MS/MS and immunoassay measurements, there were 32 patients who underwent AVS among whom sampling was selective in 31 and indicated lateralization in six, one of whom underwent adrenalectomy (Supplemental table 2). Follow-up in that single patient showed no improvement in blood pressure; also measurements of aldosterone by the Liaison immunoassay showed positive test results for the ARR and SIT, the latter yielding a post-SIT value of aldosterone of 305 pmol/L, not lower than that at confirmation 15 months earlier (264 pmol/L). LC-MS/MS measurements are not, however, available. For the other patients with discordant results and no evidence of lateralization, follow-up that was compliant with protocol requirements has been possible in one patient, who showed a negative test result of the SIT by LC-MS/MS measurements. For the other patients, follow-up has either not yet been possible or has not been compliant with protocol requirements.

Among all patients with concordant positive results for the SIT and lateralized aldosterone secretion, 90% had undergone adrenalectomy by 28/10/21, an expectedly higher proportion than for the patients with discordant results (Supplemental table 2). Follow-up, when completed according to protocol requirements, has indicated biochemical cure in all patients with concordantly positive SIT results; according to PASO criteria this assures a final diagnosis of unilateral PA.

Supplemental table 2. Patient flow and follow-up status according to PROSALDO protocol requirements in relation to study center and negative, positive and discordant results of immunoassay and LC-MS/MS measurements of aldosterone after the saline infusion test (SIT)

Study center	DR	MU	WU	WW	ZH	Total
SIT IA & MS	47	60	46	28	59	240
SIT Negative	21	16	21	6	14	78
FU possible (>6 mths)	16	13	7	6	10	52
FU done	7	6	1	0	4	18
FU compliant	3	1	1	NA	1	6
FU outcomes						
PA excluded	3	1	1	0	1	6
PA confirmed	0	0	0	0	0	0
PA exclusion unconfimed	13	10	6	6	9	44
SIT Positive	21	16	15	13	35	100
AVS	16	12	12	10	15	65
AVS selective	11	12	11	7	11	52
AVS Lateralization	9	10	9	6	5	39
Adrenalectomy*	15	6	8	1	5	35
FU possible (>6 mths)	12	4	1	0	3	20
FU done	12	4	0	0	3	19
FU compliant	12	4	NA	NA	3	19
Adrenalectomy FU Outcomes						
Unilateral / cure	11	4	0	0	2	17
Bilateral / no cure	0	0	0	0	1	1
PA excluded	0	0	0	0	0	0
PA unconfirmed	1	0	0	0	1	2
AVS No lateralization	2	2	2	1	6	13
FU possible (>6 mths)	2	1	0	0	4	7
FU done	2	1	0	0	2	5
FU compliant	0	0	NA	NA	0	0
AVS no lateralization outcomes						
Bilateral PA confirmed	0	0	0	0	0	0
Bilateral PA unconfirmed	2	1	0	0	4	7
PA excluded	0	0	0	0	0	0
SIT Discordant	5	28	10	9	10	62
AVS	3	13	7	3	6	32
AVS selective	2	13	7	3	6	31
AVS Lateralization	1	3	2	0	0	6
Adrenalectomy	0	1	0	0	0	1
FU posssible (>6 mths)	0	1	0	0	0	1
FU done	0	1	0	0	0	1
FU compliant	NA	1	NA	NA	NA	1
Adrenalectomy FU outcomes						
Unilateral / cure	0	0	0	0	0	0
Bilateral / no cure	0	0	0	0	0	0
PA excluded	0	0	0	0	0	0
PA unconfirmed	0	1	0	0	0	1
AVS No lateralization	1	10	5	3	6	25
FU possible (>6 mths)	1	8	2	0	2	13
FU done	1	3	0	0	0	4
FU compliant	1	1	0	NA	NA	2
AVS no lateralization outcomes						
Bilateral PA confirmed	0	0	0	0	0	0
Bilateral PA unclear	1	9	5	3	6	24
PA excluded	0	1	0	0	0	1

*Adrenalectomy performed in 5 DR patients based on imaging data without selective AVS results and 1 other DR patient based on imaging results and young age. Abbreviations: AVS adrenal venous sampling; FU, follow-up; NA, not applicable due to lack of FU; PA, primary aldosteronism.

Aldosterone measurements in patients with discordant results
Among the 62 patients with discordant results of initial LC-MS/MS and immunoassay measurements for the SIT there was considerable variance in the extent of discordance (Supplemental table 3). Although for some patients the discordance was minimal (e.g., ZH1) and could simply reflect measurement imprecision, for almost all others the discordance was clearly delineated.

Supplemental table 3. Plasma concentrations (pmol/L) of aldosterone measured by the four assay methods in the 62 patients with discordant results for the initial immunoassay and LC-MS/MS methods
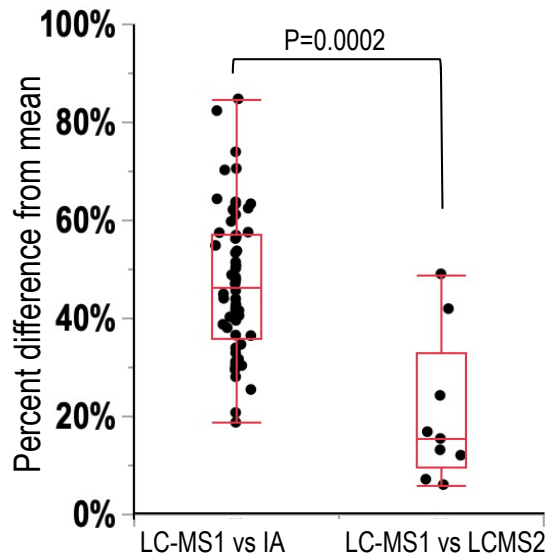
ID	LCMS1	LCMS2	Liaison	iSYS	ID	LCMS1	LCMS2	Liaison	iSYS
DR1	50	89	222	473	MU26	75	78	200	172
DR2	92	81	168	470	MU27	105	123	258	211
DR3	25	<78	133	332	WU1	147	155	502	378
DR4	141	97	430	183	WU2	119	85	139	211
DR5	144	103	267	208	WU3	119	130	289	214
MU1	119	<78	222	153	WU4	144	183	268	186
MU2	94	92	227	141	WU6	89	144	394	186
MU3	69	111	297	180	WU7	89	132	319	189
MU4	92	80	274	189	WU8	133	114	263	241
MU5	58	141	178	<103	WU9	141	183	NM	324
MU6	75	<78	214	122	WU10	69	74	NM	191
MU7	139	126	264	133	WU11	36	56	NM	239
MU8	80	77	286	203	WW1	141	120	286	161
MU8	39	<78	172	<103	WW2	83	135	380	144
MU9	67	125	175	<103	WW3	136	191	338	181
MU10	128	145	289	211	WW4	92	138	300	126
MU11	78	155	183	136	WW5	122	166	205	168
MU12	67	120	300	266	WW6	55	70	236	104
MU13	89	110	294	119	WW7	31	40	175	<65
MU14	100	163	222	180	WW8	150	172	544	173
MU15	72	<78	206	<103	WW9	50	74	185	<65
MU16	69	<78	178	<103	ZH1	155	155	226	155
MU17	42	<78	239	NM	ZH2	122	161*	344	189
MU18	47	<78	195	<103	ZH3	136	135	206	158
MU19	86	209	244	<103	ZH4	89	87	208	69
MU20	111	<78	228	139	ZH5	128	78	246	112
MU21	97	<78	187	316	ZH6	150	169	363	271
MU22	117	108	219	189	ZH7	92	78	172	87
MU23	64	<78	252	<103	ZH8	119	101	305	159
MU24	58	81	180	<103	ZH9	80	78	170	139
MU25	67	195	247	205	ZH10	153	155	280	242

Grey highlighted fields serve to illustrate positive results that were discordant with the initial LC-MS/MS results. Patient IDs are indicated for each of the five centers (DR, MU, WU, WW and ZH) according to patients at each centers. NM, not measured. *For this patient a different LC-MS/MS method was used for the measurement.

There were nine patients (MU14, MU19, MU25, WU4, WU9, WW3, WW5, WW8 and ZH6) among the 62 patients with discordant results for initial LC-MS/MS and immunoassay methods who upon second LC-MS/MS measurements had positive results concordant with the initial immunoassays (Supplemental table 3). For seven of those patients the second measurement by LC-MS/MS was

within 20% of the cut-offs for the SIT (162 pmol/L) for LC-MS/MS Also, for two of those nine patients (MU19 and WW5) the second immunoassay yielded negative results that were concordant with the first but not the second LC-MS/MS method. In others (e.g., WW3, WW8, ZH6) although there were discordant negative and positive tests results between the two LC-MS/MS methods, there remained large differences between measurements between both LC-MS/MS methods and the initially discordant immunoassay measurement.

Supplemental figure 3. Percent differences from the mean of two measurements for the 62 discordant results between LC-MS/MS and immunoassay measurements and the 9 discordant results between the two LC-MS/MS measurements.

Three of the nine patients (33%) with discordant results for the SIT by LC-MS/MS measurements showed differences in measurements from the mean of more than 20% (supplemental figure 3). However, one of these three patients with a difference of 42% (MU19) had negative results for the SIT by the second immunoassay. In contrast, all except one of the 62 patients (98%) with discordant results for the SIT according to initial LC-MS/MS and immunoassay measurements showed differences in measurements from the mean of more than 20%. More than 50% (32/62) showed differences of more than 45% compared to only one of the nine patients with discordant results according to LC-MS/MS measurements.

Lateralization ratios, contralateral suppression & machine learning interpretations

The 11 patients with selective AVS who showed lateralized aldosterone secretion and who were diagnosed with unilateral PA on the basis of biochemical cure on follow-up had 2.7-fold higher lateralization ratios, less than a tenth the contralateral suppression indexes and 4.3-fold higher machine learning probabilities of PA than the six patients with discordant test results for the SIT and AVS evidence of aldosterone lateralization (Supplemental figure 4).
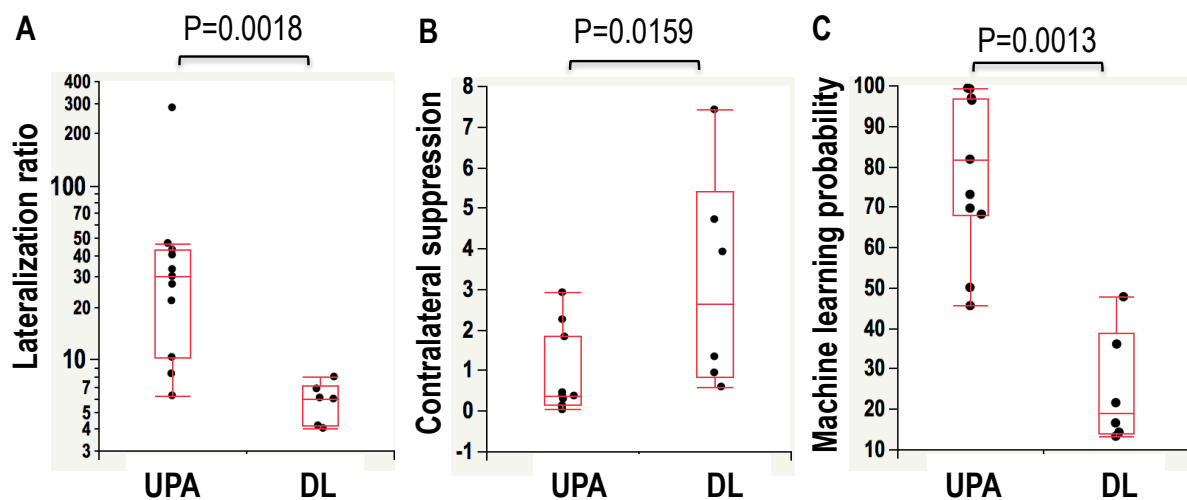
Supplemental figure 4. Lateralization ratios (A), contralateral suppression indexes (B) and machine learning probabilities of PA in 11 patients with concordant positive test results for the SIT and unilateral PA (UPA) compared to six patients with discordant test results for the SIT who also showed AVS evidence of lateralization (DL)

Supplemental appendix

Immunoassay interference
Plasma concentrations of aldosterone in samples from patients with discordant SIT results were consistently higher before than after removal of interferents for all seven patients in who repeat immunoassays were performed (Supplemental table 4). Concentrations for many samples measured by immunoassays after processing were below the limits of quantification (LOQ) compared to concentrations in plasma specimens before processing. Patients 1, 2, 3, 5 and 6 all had positive SIT results (>170 pmol/L) for the Liaison immunoassay that were discordant with the negative results for both LC-MS/MS methods and the iSYS immunoassay. After processing to remove interferents, all five patients showed non-pathologic SIT results (<170 pmol/L) for the Liaison immunoassay that were concordant with the non-pathologic results for the other assays. Patient no 4 had positive SIT results (>170 pmol/L) for the iSYS immunoassay that were discordant with negative results for both LC-MS/MS methods and the Liaison immunoassay, while patient 7 showed positive SIT results for both the Liaison and iSYS immunoassays that were discordant with the negative results of both LC-MS/MS methods. In all above cases, discordant positive and negative results became concordantly negative after removal of interferents.

Supplemental table 4. Plasma aldosterone concentrations (pmol/L) for six patients measured by two independent LC-MS/MS methods and by Liaison and iSYS immunoassays before and after processing of plasma to remove macromolecular interferents

| Patient	Sample	LC-MS/MS		Liaison		iSYS	
		LCMS-1	LCMS2	Before	After	Before	After
1	BL1	61	<78	200	<96	114	<124
	BL2	97	<78	299	<96	275	<124
	SIT*	47	<78	195	<96	<103	<124
2	BL1	125	153	255	116	183	<124
	BL2	164	154	329	145	264	<124
	SIT*	72	<78	206	<96	<103	<124
3	BL1	222	188	414	197	297	126
	BL2	311	263	505	247	416	244
	SIT*	64	<78	252	<96	<103	<124
4	BL1	333	169	480	301	696	383
	BL2	197	93	302	209	627	176
	SIT**	25	<78	133	<96	332	<124
5	BL1	172	133	227	102	216	<124
	BL2	222	279	422	166	399	233
	SIT*	39	<78	172	<96	<103	<124
6	BL1	130	133	291	118	180	<124
	BL2	75	109	241	<96	158	<124
	SIT*	58	81	180	<96	<103	<124
7	BL1	150	103	269	138	372	<124
	BL2	286	275	403	188	638	280
	SIT***	97	<78	187	<96	316	<124

Measurements for each of the 7 patients were from screening samples (BL1) and from samples taken before (BL2) and at the end of saline infusion tests (SIT). *Discordant results for the SIT restricted to the Liaison immunoassay; **Discordant result for the SIT restricted to the iSYS immunoassay; ***Discordant results for the SIT involving both Liaison and iSYS immunoassays

Supplemental appendix

ROC curves for the SIT

ROC curves were established for the SIT based on two different criteria for classification of the presence of disease. A classification of PA was based either strictly on post-surgical biochemical cure in 17 patients who had undergone adrenalectomy or more widely on combinations of immunoassay and LC-MS/MS post-SIT pathologic results involving plasma concentrations above respective upper cut-offs of 170 and 162 pmol/L. For the latter criteria of presumed PA, there were nine patients with results for a second LC-MS/MS that were pathogenic (Supplemental table 3), according to measured concentrations between 163 and 209 pmol/L. However, for two of those nine patients with respective measured concentrations of 122 and 86 pmol/L for the first LC-MS/MS and 168 and 209 pmol/L for the second LC-MS/MS, the results of the second immunoassay yielded respective non-pathogenic results of 168 and <103 pmol/L. Those two patients were therefore excluded from ROC curve analyses. On this basis 107 patients were classified with presumed PA.

Absence of disease was defined according to concordantly non-pathologic test results of the SIT by both LC-MS/MS and immunoassays. For patients with discordant results, disease was considered excluded when results for two LC-MS/MS methods showed concordantly negative results. In this way 131 patients were classified to have non-pathogenic results.

For LC-MS/MS measurements, the 17 patients with strictly defined criteria for disease confirmation had post-SIT plasma aldosterone concentrations ranging from 180 pmol/L in each of two patients to a highest value of 1820 pmol/L. Using ROC curve-derived Youden indexes, and according to the strictly defined classification of PA, the optimal cut-offs for the SIT were defined as 180 pmol/L (65 ng/L) and 337 pmol/L (121 ng/L) for respective LC-MS/MS and immunoassay measurements of aldosterone (Supplemental figure 5A). These cut-offs provided 100% specificity and sensitivity for LC-MS/MS measurements compared to 97.6% specificity and 94.1% sensitivity for immunoassay measurements. For immunoassays, diagnostic sensitivity could be increased to 100% with a lower cut-off of 272 pmol/L, but this was associated with an unacceptable drop in specificity to 88.8%.
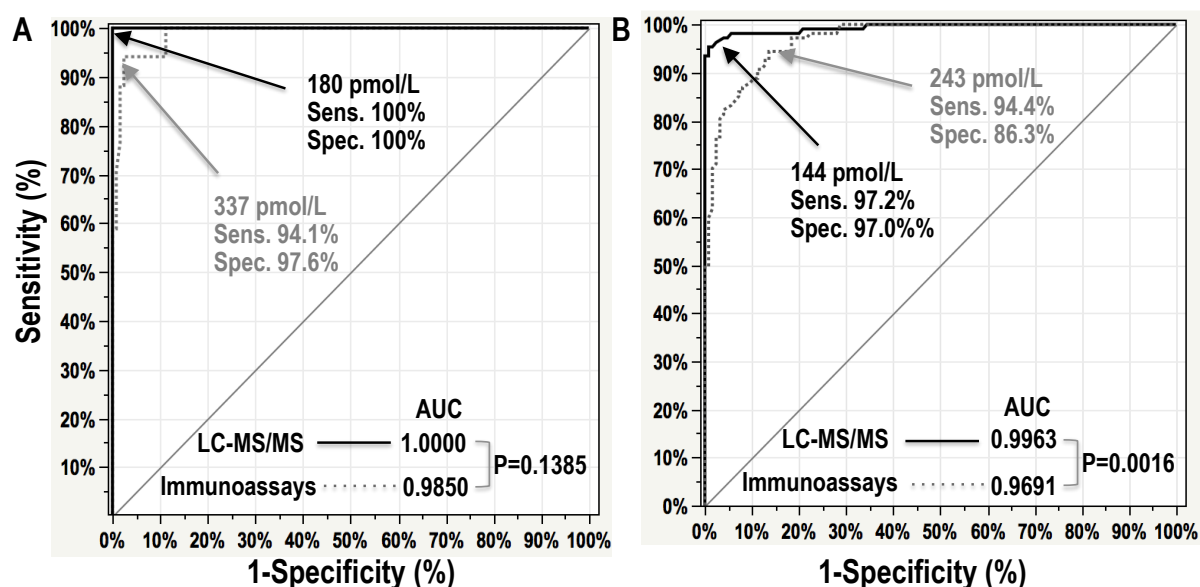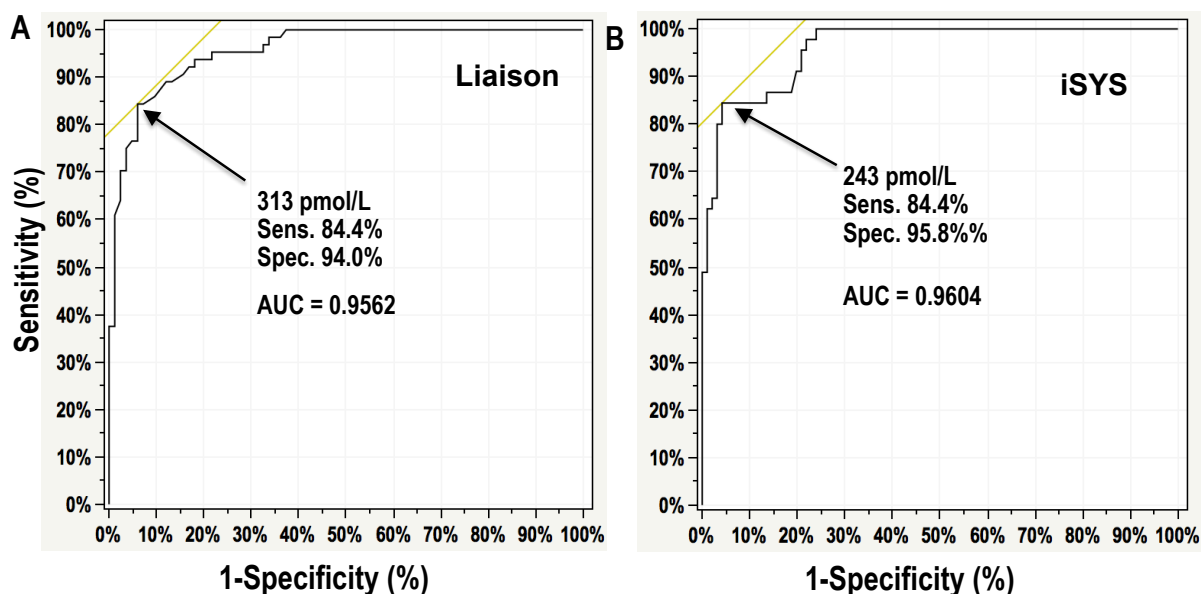

Supplemental figure 5. ROC curves for the SIT shown for measurements by LC-MS/MS and immunoassays (IA) and according to classification of PA based on post-operative cure after adrenalectomy (A) or based on concordantly positive test results for the SIT by both LC-MS/MS methods. Areas under the curves (AUC) are provided and significant of differences between AUC are shown according to analyses of paired LC-MS/MS and immunoassay data. Optimized cut-offs were assessed according to Youden indexes with associated diagnostic sensitivities (Sens.) and specificities (Spec.).

According to the criteria of presumed PA (i.e., concordant pathogenic results of LC-MS/MS), the optimal cut-offs were defined as 144 pmol/L (52 ng/L) and 243 pmol/L (88 ng/L) for respective LC-MS/MS and immunoassay measurements of aldosterone (Supplemental figure 5B). These cut-offs provided 97.0% specificity and 97.2% sensitivity for LC-MS/MS measurements compared to only 86.3% specificity and 94.4% sensitivity for immunoassay measurements. More acceptable specificity

of 96.0% could be achieved for immunoassay measurements with a cut-off of 303 pmol/L (109 ng/L), but this was associated with a diagnostic sensitivity of only 82.2%. The area under the ROC curve (AUC) was higher (P=0.0016) according to measurements of aldosterone by LC-MS/MS than measurements by both immunoassays considered together.

Since the two immunoassays differed in their discordance with LC-MS/MS measurements, ROC curves were also established independently for each immunoassay (Supplemental figure 6A & B). However, due to limited numbers of patients with post-operative cure for each immunoassay method, this was only possible for the relaxed wider criteria of presumed PA based on concordant positive results. From these analyses optimal cut-offs for immunoassays were defined as 313 and 243 pmol/L (113 and 88 ng/L) for respective Liaison and iSYS immunoassays; these cut-offs yielded respective specificities and sensitivities of 94.0% and 84.4% for the Liaison immunoassay and 95.8% and 84.4% for the iSYS immunoassay.

Supplemental figure 6. ROC curves for the SIT according to measurements by the Liaison (A) and iSYS (B) immunoassays. Areas under the curves (AUC) are provided and optimized cut-offs were assessed according to Youden indexes with associated diagnostic sensitivities (Sens.) and specificities (Spec.).

Expanded discussion: study strengths and limitations
Although this study has a number of strengths, there are also limitations related to lack of "gold standards" to conclusively exclude disease beyond use of diagnostic tests themselves. Similarly, for patients with positive test results who do not undergo surgical resection of an adrenal (i.e., those classified with bilateral or idiopathic PA) there is no "gold standard" to conclusively confirm disease beyond use of diagnostic tests. Thus, for both groups of patients the diagnoses are only presumed and are dependent on classifications of pathogenic versus non-pathogenic results according to the cut-offs for the SIT, as reported elsewhere (1,9). As detailed further below such cut-offs only provide a guide and the grey zone around those cut-offs should be considered in relation to not only the continuum of the disease in PA, but also variations in accuracy and precision of different methods for measurement of aldosterone.

The aforementioned limitations and considerations are relevant to almost all studies involving diagnosis of PA. Nevertheless, the focus of the present report was not to define new cut-offs for the SIT or establish whether LC-MS/MS provides a superior method compared to immunoassay measurements for diagnosis of PA. Rather the focus of the present report was to characterize inaccuracies in immunoassay measurements of aldosterone in relation to LC-MS/MS measurements, which were assumed to provide superior analytical specificity and accuracy compared to the former. This assumption was tested in two ways: 1. repeating measurements using a second independent LC-MS/MS method and an additional immunoassay in patients in who there was discordance according to

results of the SIT; 2. repeating discordant immunoassays after solid phase removal of potential macromolecular interferents.

The first approach of repeated measurements by independent methods provides a strength of the present study compared to previous method comparison studies of LC-MS/MS and immunoassay measurements of aldosterone during the SIT (10,11). This approach was made possible by the prospective nature of the study with collection and banking of additional samples that could be tested by further measurements using independent analytical methods. This approach in particular enabled identification of patient samples for which there was either clear analytical inaccuracy of one immunoassay compared to all other three methods of analysis or where both immunoassays were discordant with the two LC-MS/MS methods.

The second approach employing solid-phase sample clean-up was an additional novel study strength that enabled identification of analytical interference as a cause of immunoassay inaccuracy and discordant results for the SIT. However, due to limited remaining specimen in biobanks for most of the patients with discordant results, these analyses were limited to seven patients; thus, a study limitation was that these repeated analyses could not conclusively establish that similar interferences from circulating macromolecules also caused discordant results for other patients. Nevertheless, discordance was eliminated by sample processing in all seven patients and for both immunoassays; thus, it is likely that similar interferences account for the observed inaccuracies of immunoassays for most if not all other patient samples.

Expanded discussion: ROC curves and diagnostic cut-offs for the SIT
As clarified above, establishment of revised cut-offs for the SIT or whether LC-MS/MS provides a superior method compared to immunoassay measurements for diagnosis of PA were not primary objectives of this study. Furthermore, the suggested ranges in cut-offs, from 144 to 180 pmol/L for LC-MS/MS-based measurements and the much higher ranges in cut-offs for immunoassays can only be regarded as preliminary until more robust classifications of disease status are enabled. While confirmation of unilateral disease by PASO compliant follow-up to establish biochemical cure provides a fully robust method for disease confirmation, this was only possible for 17 patients of the present analysis. Moreover, use of this criterion for disease confirmation meant that 83 other patients with concordant positive test results for the SIT had to be excluded. As a consequence of this, the 100% values for diagnostic sensitivities and specificities are illusionary. Nevertheless, the defined cut-off of 180 pmol/L does provide a preliminary upper cut-off to achieve 100% sensitivity since any higher value would have resulted in two patients with confirmed unilateral PA being missed.

By defining cut-offs for the SIT using patients with follow-up confirmed unilateral disease, it might be argued that patients with idiopathic, bilateral or milder forms of the disease may be missed (12). On the other hand, such cut-offs are more appropriate than lower cut-offs for identifying patients most suitable for AVS in whom lateralized aldosterone secretion can be determined and cure achieved by adrenalectomy. This would also minimize numbers of patients without lateralized aldosterone secretion undergoing a procedure that has limited clinical benefit to those particular patients, thereby prioritizing this labor-intensive, difficult and costly procedure to patients most likely to benefit.

Preliminary revised cut-offs for the SIT for both the Chromsystems LC-MS/MS kit method and two immunoassays (a radioimmunoassay and the iSYS immunoassay) have been recently described elsewhere (11). In that report optimal cut-offs were defined at 150 pmol/L (54 ng/L) for the immunoassays and 191 pmol/L for LC-MS/MS. However, disease classification was based on routine diagnostics, including an immunoassay-based cut-off for the SIT of only 139 pmol/L (50 ng/L) and without follow-up confirmatory studies. As we now show here the analytical interference-associated inaccuracies of immunoassay measurements compromises their use for disease classification. Moreover, if we had applied the cut-offs of 150 and 191 pmol/L for respective immunoassays and LC-MS/MS measurements (as opposed to those of 170 pmol/L and 162 pmol/L in the present study) the proportion of discordant positive and negative results for the SIT would have been considerably larger than the 26% outlined in the present report. Nevertheless, the cut-off of 191 pmol/L for the Chromsystems LC-MS/MS kit method might be appropriate given that this method appears to measure plasma aldosterone about 15% higher than the LC-MS/MS method at DR.

As outlined previously by Thuzar et al. (7), and in keeping with the higher concentrations of aldosterone measured by immunoassays than by LC-MS/MS, optimal cuts-off for the SIT are also higher and not lower by immunoassay than by LC-MS/MS. Our analyses are in agreement with that conclusion. Importantly, both the study of Thuzar et al. (7), and an earlier report that established the SIT cut-off of 162 pmol/L for aldosterone (9), used the fludrocortisone suppression test as an alternative method to exclude PA. This provides at least a partial solution to satisfy requirements for an independent method to the negative results for the SIT to exclude disease, even if that test may not represent a true "gold standard".

Expanded discussion: Assay precision

As mentioned above, it should be appreciated that any stipulated cut-offs for the SIT only provide a guide; there is a significant grey zone around those cut-offs, dependent not only on the continuum of the disease in PA, but also variations in accuracy and precision of different methods for measurement of aldosterone. Precision of measurements can be particularly problematic at lower concentrations, such as after the SIT. As apparent from the inter-assay coefficients of variation in supplemental table 1, the degree of imprecision applies near equally to immunoassay and LC-MS/MS measurements of aldosterone. Moreover, at the LOQ of most assays the imprecision is usually set to 20% and as detailed in supplemental table 1 inter-assay coefficients of variation around the cut-offs of the SIT can be expected to range from 10% and 20%. This implies that at a measured value of 160 pmol/L (just under the cut-off of 162 pmol/L for LC-MS/MS post-SIT measurements) plasma aldosterone may be measured again by the same method as high as 192 pmol/L or as low as 148 pmol/L. With a different method, even if similarly accurate (i.e., minimal bias), there may be some further imprecision.

With the aforementioned considerations in mind, it can be determined from the data presented in supplemental table 3 and figure 3 that the discordant post-SIT LC-MS/MS results for six of the nine patients could simply reflect assay imprecision around the cut-offs of 162 pmol/L. Moreover, for one of the remaining three patients, the second immunoassay measurement of aldosterone after the SIT was concordantly negative and in agreement with measurements of the first but not the second LC-MS/MS method. This leaves only two patients with unexplained discordant results of more than 20% from the mean, though for one of these the difference was only 24%. This compares with all but one of the 62 patients with discordant results between initial immunoassay and LC-MS/MS measurements of aldosterone of beyond 20%. Thus, for these 98% of patients, assay imprecision is highly unlikely to account for the discordant immunoassay and LC-MS/MS results. Rather the data indicate that analytical interference-related inaccuracy and a resulting bias towards variably higher than true concentrations of aldosterone measured by immunoassays accounts for the differences in results.

Results closely around the grey zone nevertheless must be interpreted with some caution, including those that are concordant, and as we clarify in this report particularly when there is only a single measurement. For such results it may be useful to follow-up with additional measurements or confirmatory tests.

Expanded discussion: perspective

In contrast to screening tests, which should offer high diagnostic sensitivity, the SIT is a confirmatory test and therefore should offer high specificity and thereby high positive predictive value. As shown by Thuzar et al. (7) and also suggested here, current cut-offs for the SIT for immunoassay measurements of aldosterone result in suboptimal diagnostic specificity; this does not confer the optimal positive predictive value required for a confirmatory test. While the optimal cut-offs for the SIT involving immunoassay-based measurements of aldosterone suggested here are considerably higher than by LC-MS/MS measurements or those outlined previously (1,9), it can now be appreciated that those other higher cut-offs are largely a result of interference-associated inaccuracies of immunoassay measurements. Through employment of procedures to eliminate interferents it should be possible to better harmonize cut-offs for the SIT between immunoassay and LC-MS/MS methods. The solid phase purification of plasma used in this report represents one method to eliminate interferents; there are, however, numerous alternative procedures, including solvent-based sample clean up procedures, that may also be employed to improve accuracy of immunoassay-based measurements of aldosterone (13).

Supplemental appendix

Although the findings of this report imply that there may be over-diagnosis of PA among those patients who are tested using immunoassay measurements of aldosterone, it must also be appreciated that current difficulties of the diagnostic process mean that many patients whom might have PA are never tested; thus, the overall consensus is that PA remains grossly under-diagnosed among the general population of patients with hypertension (14-18). Through the aforementioned solutions and by way of other advances it can be expected that improvements will be realized in both identification and management of patients with PA. Hopefully such improvements should make the processes for diagnosis of PA less onerous and more readily accessible to clinicians who might then be willing to consider this diagnosis. We might then be able to realize the true extent and nature of the disease and subsequently be better prepared to manage and treat those patients in whom the diagnosis is clear. Until then, previous findings and conclusions that have depended on immunoassay-based diagnostics may need to be re-evaluated in light of the data reported here.

References

1. Funder JW, Carey RM, Mantero F, et al. The management of primary aldosteronism: Case detection, diagnosis, and treatment: An endocrine society clinical practice guideline. J Clin Endocrinol Metab 2016; 101:1889-1916
2. Williams TA, Lenders JWM, Mulatero P, et al. Outcomes after adrenalectomy for unilateral primary aldosteronism: An international consensus on outcome measures and analysis of remission rates in an international cohort. Lancet Diabetes Endocrinol 2017; 5:689-699
3. Eisenhofer G, Duran C, Cannistraci CV, et al. Use of steroid profiling combined with machine learning for identification and subtype classification in primary aldosteronism. JAMA Netw Open 2020; 3:e2016209
4. Constantinescu G, Bidlingmaier M, Gruber M, et al. Mass spectrometry reveals misdiagnosis of primary aldosteronism with scheduling for adrenalectomy due to immunoassay interference. Clin Chim Acta 2020; 507:98-103
5. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (redcap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42:377-381
6. Harris PA, Taylor R, Minor BL, et al. The redcap consortium: Building an international community of software platform partners. J Biomed Inform 2019; 95:103208
7. Thuzar M, Young K, Ahmed AH, et al. Diagnosis of primary aldosteronism by seated saline suppression test-variability between immunoassay and HPLC-MS/MS. J Clin Endocrinol Metab 2020; 105
8. Peitzsch M, Dekkers T, Haase M, et al. An lc-ms/ms method for steroid profiling during adrenal venous sampling for investigation of primary aldosteronism. J Steroid Biochem Mol Biol 2015; 145:75-84
9. Stowasser M, Ahmed AH, Cowley D, et al. Comparison of seated with recumbent saline suppression testing for the diagnosis of primary aldosteronism. J Clin Endocrinol Metab 2018; 103:4113-4124
10. Fries CM, Bae YJ, Rayes N, et al. Prospective evaluation of aldosterone lc-ms/ms-specific cutoffs for the saline infusion test. Eur J Endocrinol 2020; 183:191-201
11. Fuss CT, Brohm K, Kurlbaum M, et al. Confirmatory testing of primary aldosteronism with saline infusion test and LC-MS/MS. Eur J Endocrinol 2021; 184:167-178
12. Brown JM, Robinson-Cohen C, Luque-Fernandez MA, et al. The spectrum of subclinical primary aldosteronism and incident hypertension: A cohort study. Ann Intern Med 2017; 167:630-641
13. Ward G, Simpson A, Boscato L, Hickman PE. The investigation of interferences in immunoassay. Clin Biochem 2017; 50:1306-1311
14. Mulatero P, Monticone S, Burrello J, Veglio F, Williams TA, Funder J. Guidelines for primary aldosteronism: Uptake by primary care physicians in europe. J Hypertens 2016; 34:2253-2257
15. Funder J. Primary aldosteronism: Treatment of the disease, and new therapeutic approaches. Best Pract Res Clin Endocrinol Metab 2019: 34(2):101368. doi: 10.1016/j.beem.2019.101368
16. Libianto R, Fuller PJ, Young MJ, Yang J. Primary aldosteronism is a public health issue: Challenges and opportunities. J Hum Hypertens 2020; 34:478-486
17. Yozamp N, Hundemer GL, Moussa M, et al. Intraindividual variability of aldosterone concentrations in primary aldosteronism: Implications for case detection. Hypertension 2021; 77:891-899
18. Conroy PC, Hernandez S, Graves CE, et al. Screening for primary aldosteronism is underutilized in patients with obstructive sleep apnea. Am J Med 2021; Sep 9:S0002-9343(21)00559-3. doi: 10.1016/j.amjmed.2021.07.041.