

Plasma steroid profiling in patients with adrenal incidentaloma

Kristina Berke,¹ Georgiana Constantinescu,¹ Jimmy Masjkur,¹ Otilia Kimpel,² Ulrich Dischinger,² Mirko Peitzsch,³ Aleksandra Kwapiszewska,⁴ Piotr Dobrowolski,⁴ Svenja Nölting,^{5,6} Martin Reincke,⁶ Felix Beuschlein,^{5,6} Stefan R. Bornstein,¹ Aleksander Prejbisz,⁴ Jacques W.M. Lenders,^{1,7} Martin Fassnacht² and Graeme Eisenhofer^{1,3}

¹Department of Medicine III, University Hospital Carl Gustav Carus, Technische Universität Dresden, Germany; ²Department of Internal Medicine I, Division of Endocrinology and Diabetes, University Hospital, University of Würzburg, Würzburg, Germany; ³Institute of Clinical Chemistry and Laboratory Medicine & University Hospital Carl Gustav Carus, Technische Universität Dresden, Germany; ⁴National Institute of Cardiology, Warsaw, Poland; ⁵ Department of Endocrinology, Diabetology and Clinical Nutrition, University Hospital Zurich and University of Zurich, Zurich, Switzerland; ⁶Department of Medicine IV, University Hospital, Ludwig Maximilian University of Munich, Munich, Germany; ⁷Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands

Contents	Page
Supplemental Methods.....	2
Inclusions according to Study Center and Protocol.....	2
Supplemental table 1.....	2
Population-based versus Pragmatic Patient Inclusion.....	2
Patient Follow-up.....	3
Normalizations for Discriminant Analyses	4
Supplemental table 2.....	5
Adjustments of PPV and NPV for Disease Prevalence.....	5
Supplemental Results.....	6
Individual Results for the 19 patients with ACC.....	6
Supplemental table 3.....	6
Plasma Steroids and Metanephrines in the Five Patient Groups.....	6
Supplemental table 4.....	7
Supplemental table 5.....	8
Logistic Regression	8
Supplemental figure 1.....	9
PPV and NPV According to Disease Prevalence.....	10
Supplemental figure 2.....	11
Supplemental figure 3.....	12
Supplemental Discussion of Study Limitations.....	13
Population Bias.....	13
Retrospective vs Prospective Study Designs and Patient Follow-up.....	13
Supplemental table 6.....	15
Future Perspectives and avoidance of Study Limitations	16
References.....	16

Supplemental Methods

Inclusions according to Study Center and Protocol

Inclusion of patients into this retrospective cross-sectional diagnostic study required imaging findings of an incidentally discovered abdominal mass according to entry criteria of one of three clinical protocols in place and approved by local Ethics committees at seven European centers: 1. University Hospital Carl Gustav Carus Dresden, Dresden, Germany; 2. University Hospital of Würzburg, Würzburg, Germany; 3. Institute of Cardiology, Warsaw, Poland; 4. University Hospital of Munich, Munich, Germany; 5. University Hospital of Zürich, Zurich Switzerland; 6. Radboud University Medical Center, Nijmegen, the Netherlands; 7. University Medical Center Schleswig-Holstein, Lübeck, Germany (Supplemental table 1).

Supplemental table 1. Numbers of patients included in the analysis according to study center and clinical protocol, before and after exclusions

	DR	WU	WW	MU	ZH	NI	LU
Numbers of patients initially included according to study protocol							
PMT	95	127	130	78	0	10	4
ENSAT	87	8	0	0	0	0	0
PROSALDO	56	3	0	0	16	0	0
TOTAL	238	138	130	78	16	10	4
Numbers of patients included according to study protocol after exclusions							
PMT	91	111	123	73	0	10	4
ENSAT	84	8	0	0	0	0	0
PROSALDO	55	3	0	0	15	0	0
TOTAL	230	122	123	73	15	10	4

Abbreviations: DR, Dresden; WU, Würzburg; WW, Warsaw; MU, Munich; ZH, Zurich; NI, Nijmegen; LU, Lübeck; PMT, Prospective Monoamine Producing Tumor study; ENSAT, European Network for the Study of Adrenal Tumors registry and biobanking protocol; PROSALDO, PROspective study on the diagnostic value of Steroid profiling in primary ALDOsteronism.

Population-based versus Pragmatic Patient Inclusion

As outlined in the manuscript that addressed the primary objective of the Prospective Monoamine-producing Tumor (PMT) study (1), inclusion of patients into that study was not population-based. As outlined in the clinical protocol, which is available at <https://pmt-study.pressor.org/>, this was because of both the expected referral nature of patients and the need to include 200 patients with pheochromocytoma and paraganglioma (PPGL) to reach sufficient numbers of patients with disease to appropriately assess and compare performance of diagnostic tests according to the power analysis. As previously clarified (1), at typical prevalences of PPGL among tested patients of between 0.8% to 1.6%, a population-based study would have required recruitment of 6- to 12-times more patients without PPGLs than the 2200 population patient target of the PMT study for this group. This was considered impractical, particularly with subsequent requirements of patient follow-up to further exclude or confirm disease in those patients in whom initial testing did not indicate PPGL. Thus, rather than a population-based approach for patient inclusion, the PMT study employed a pragmatic approach to ensure sufficient numbers of patients with disease to ensure accurate estimates of diagnostic sensitivity. The limitation of this approach was that estimates of positive predictive value (PPV) and negative predictive value (NPV) would not reflect usual prevalences of disease. To overcome this shortcoming and address variable pre-test prevalence of disease according to the different inclusion criteria (i.e., presence of signs and symptoms of presumed catecholamine excess,

findings of an incidentaloma, routine surveillance due to hereditary predisposition or previous history of a resected PPGL), PPV and NPV were calculated for different prevalence of disease.

The PROspective study on the diagnostic value of steroid profiling in primary ALDOsteronism (PROSALDO) similarly used a pragmatic rather than population-based approach to patient inclusion with a planned 1:2 ratio of patients with and without primary aldosteronism (PA). As in the PMT study, all patients were included based on clinical suspicion of disease according to several criteria and without knowledge or proof of the presence of disease at the time of study inclusion. A finding of an adrenal incidentaloma in a patient with hypertension was one of the several inclusion criteria for the PROSALDO trial.

Rather than focusing on patients with suspected PPGL or PA, the European Network for the Study of Adrenal Tumors (ENSAT) registry and biobanking protocol allows for inclusion of any patient with an adrenal mass. Although inclusive of all such patients, this protocol can nevertheless be subject to referral bias.

As a consequence of the aforementioned nature of patient inclusion into the three protocols, those included into the present study population according to findings of an adrenal incidentaloma cannot be expected to reflect proportions of different patient groups (i.e., for a population-based study. This study limitation was expected and, as subsequently clarified, is addressed according to similar methods employed for the PMT study (1).

Patient Follow-up

Follow-up of patients included by way of the PMT study initially followed the procedures outlined in the primary manuscript that arose from that study (1). However, that study, which was initiated in January 2011, was directed to diagnosis of PPGL. Thus, initial follow-up focused on identification of previously undiagnosed patients with PPGL and exclusion of PPGL in other patients. Nevertheless, exclusion of disease in patients with an incidentaloma was based in most patients on an alternative diagnosis (e.g., non-functional adenoma, adrenocortical carcinoma, aldosterone- or cortisol-producing adenoma) established after resection of the mass or based on routine laboratory tests employed to establish an alternative diagnosis. In other patients exclusion of PPGL was based on negative results of repeat biochemical testing at follow-up, including in some patients negative results of confirmatory testing according to the clonidine suppression test. Imaging characteristics, such as unenhanced attenuation Hounsfield unit (HU) values on computed tomography (CT), also provided evidence to exclude a PPGL in isolated patients. Among the patients with incidentaloma included via the PMT study, an alternative diagnosis was thereby initially reached in 241 patients by January 2018.

Further follow-up of patients without a diagnosis of PPGL, who were enrolled into the PMT study, was initiated in January 2020. This involved comprehensive review of patient records within hospital information systems at Munich, Würzburg and Dresden by a single investigator (KB). Two other investigators searched and reviewed medical records for patients enrolled into the PMT study at Warsaw (AK) and Nijmegen (JWML). This second follow-up of patients enrolled into the PMT study was aimed at confirming, refining or correcting initial diagnoses and included establishing a diagnosis in those patients in whom PPGL was excluded on the basis of negative biochemical testing for PPGL. That second follow-up was completed in April 2021 and was then further supplemented by a final check of patient classifications by study investigators at each of the aforementioned centers, which was completed at the end of May 2021. Through this process, among the 444 patients with incidentaloma included via the PMT study, a final diagnosis was reached in 437 patients. However, among those 437 patients a further 25 patients were excluded from the study on the basis of measurable plasma concentrations of dexamethasone or due to findings that the incidentally discovered mass had an extra-adrenal location. Thus, paragangliomas were excluded.

Importantly, among the patients with a corrected diagnosis on follow-up there was one 45-yr old male patient with a 1.4x1.5 cm right adrenal mass discovered incidentally by CT in 2010. That patient was included into the PMT study in May 2016 at which time the adrenal mass remained stable in size on imaging. The patient did not have hypertension and apart from reporting periods of nausea was otherwise asymptomatic. Biochemical testing indicated a plasma concentration of normetanephrine of 102 pg/mL (0.56 nmol/L), which was under the upper cut-off of age specific reference intervals (148 pg/mL, 0.81 nmol/L), and a plasma concentration of metanephrine of 93 pg/mL (0.47 nmol/L) that was just above the upper cut-off of reference intervals (88 pg/mL, 0.45 nmol/L). With follow-up under the PMT protocol, PPGL was excluded on the basis of imaging characteristics that suggested an adenoma. In November 2020 the patient was referred back to the endocrine outpatient clinic because of development of palpitations, tremor and hyperhidrosis. Plasma concentrations of metanephrine were increased to 166 pg/mL (0.84 nmol/L) and a CT revealed a slightly enlarged adrenal mass (1.8x1.5 cm) with CT unenhanced attenuation HU values of 20. A mass measuring 2.2 cm was resected in February 2021 and pathologically confirmed to be a pheochromocytoma.

For the 75 patients with adrenal incidentaloma initially included by way of the PROSALDO trial, the focus of that study was on confirmation and exclusion of primary aldosteronism. Among the 26 patients with PA that were included into the analysis by way of this protocol, that diagnosis was ultimately based on final confirmation by a saline infusion test (SIT) that required both plasma aldosterone concentrations above (58 ng/L) 162 pmol/L by liquid chromatography tandem mass spectrometry (LCMS/MS) and above 61 ng/L (170 pmol/L) by routine immunoassay measurements. Exclusion of PA was based on negative results for LC-MS/MS measurements for the SIT, which when positive by immunoassay had to be confirmed to be negative by an independent LC-MS/MS method. In this way PA was excluded in 49 patients. Among these 49 patients, a plasma concentration of cortisol above 1.8 µg/dL (50 nmol/L) after the dexamethasone suppression test (DST) was used to define autonomous cortisol secretion (ACS) in 14 patients. Pheochromocytoma was excluded based on negative biochemical test results for measurements of plasma free metanephrines and/or an alternative diagnosis. For two patients a final diagnosis under PROSALDO could not be reached and these patients were excluded from the final study population. The remaining patients with a negative DST were defined to have non-functional adrenal incidentaloma (NFAI).

The remaining 95 patients from Dresden and Würzburg were included into the study by way of the ENSAT registry and biobanking protocol. These patients were enrolled into that protocol between April 2013 and May 2021. Diagnosis of adrenocortical carcinoma (ACC), ACS, NFAI, PA or pheochromocytoma was based on established routine diagnostic procedures supplemented by review of medical records undertaken at Dresden by three investigators (K.B., G.C., J.M.) and at Würzburg by one investigator (O.K.). Three patients were excluded, two based on lack of sufficient information for a diagnosis and the other based on diagnosis of an ACTH-secreting adrenal pheochromocytoma.

Normalizations for Discriminant Analyses

Plasma concentrations of many adrenal steroids show marked differences according to sex and age, including lowered concentrations with advanced age (2). For some steroids, such as corticosterone and DHEA, upper cut-offs of reference intervals are between 2.5- and 9.7-fold higher in 25 compared to 75 year old patients and vary according to sex. In the present series of patients with incidentaloma, although age for half of all patients was between 50 and 67 years, the range varied widely from 17 to 86 years and necessitated consideration in models. However, rather than using age-specific reference intervals for normalization we chose to use the geometric means of age-related distributions (supplemental table 2). These were calculated similarly to age and sex-specific reference intervals described in the original report and using the data from the reference population and according to the curve-fitting procedures described in that report (1). For aldosterone and cortisol, where influences of

age were significant but relatively small, linear models were used for correction based on regression equations. Similarly, linear models were used for calculating age-specific mean concentrations of normetanephrine according to large populations of patients described elsewhere and where the data are also available in excel datasets of the supplement (3). For analytes that showed no or negligible relationships with age (i.e., 18-oxocortisol, 18-hydroxycortisol and metanephrine), sex-specific geometric mean concentrations were calculated. Age and sex specific geometric means were then used as denominators to normalize plasma concentrations of analytes, with a further base 10 logarithmic transformation applied before discriminant analyses.

Supplemental Table 2. Factors for normalization of plasma concentrations of selected analytes for influences of age and sex.

Analytes	Males	Females
Steroids (ng/mL)		
Aldosterone	$-0.000468 \times \text{age} + 0.07587$	$-0.001079 \times \text{age} + 0.11737$
18-Oxocortisol*	0.0128	0.0128
18-Hydroxycortisol*	0.8489	0.5956
11-Deoxycorticosterone	$0.33046 \times 10^{-0.0081356 \times \text{age} - 0.669511}$	$0.33046 \times 10^{-0.0096353 \times \text{age} - 0.673573}$
Corticosterone	$14.83 \times \text{age}^{-0.522}$	$92.6 \times \text{age}^{-0.8867} - 081928$
11-Deoxycortisol	$0.34646 \times 10^{-0.0018946 \times \text{age} - 0.187305}$	$0.34646 \times 10^{-0.0025784 \times \text{age} - 0.310487}$
Cortisol	$-0.72786 \times \text{age} + 141.01$	$-1.2210547 \times \text{age} + 156.74$
17-Hydroxyprogesterone†	$2.502 \times \text{age}^{-0.3158}$	0.750 / 0.224
Androstenedione	$-0.2887 \times \text{age}^{-0.2771} + 1.7814$	$2.2552 \times e^{[-0.02023 \times \text{age}]}$
DHEA	$11.868 \times e^{[-0.03028 \times \text{age}]}$	$8.2805 \times e^{[-0.02637 \times \text{age}]}$
DHEA-Sulfate	$0.3685 \times 10^{-0.0077304 \times \text{age} + 4.0167611}$	$0.3685 \times 10^{-0.0051228 \times \text{age} + 3.7204205}$
Catecholamine O-methylated metabolites (pg/mL)		
Normetanephrine	$0.54895 \times \text{age} + 43.09$	$0.54895 \times \text{age} + 43.09$
Metanephrine*	33.9	25.7

*For 18-oxocortisol, 18-hydroxycortisol and metanephrine, impacts of age were not significant or minimal and normalizations were based on mean concentrations of the reference population. †for 17-hydroxyprogesterone there is a pronounced impact of menopause on distributions of this steroid and normalizations were based on mean concentrations for females before and after the age of 50.

Adjustments of PPV and NPV for Disease Prevalence

Results for PPV and NPV reported in table 2 were derived from the study population at proportions of 3.3% (19/577) for ACC, 18.0% (104/577) for ACS, 54% (312/577) for NFAI, 11.2% (65/577) for PA and 13.3% (77/577) for pheochromocytoma. As outlined in the earlier section (Population-based versus Pragmatic Patient Inclusion) inclusion of patients was not population based. Therefore, estimates of PPV and NPV were calculated according to differences in prevalence based on the following equations,

$$\text{PPV} = \frac{\text{Sens} \times P}{(\text{Sens} \times P) + [(1-\text{Spec}) \times (1-P)]} \quad \text{NPV} = \frac{\text{Spec} \times (1-P)}{[(1-\text{Sens} \times P) \times P] + [(\text{Spec}) \times (1-P)]}$$

where P represents prevalence and Sens and Spec respectively represent diagnostic sensitivity and specificity.

Supplemental results

Individual Results for the 19 Patients with ACC

Mean tumor diameters in patients with ACC ranged from 3.5 to 12.5 cm (Supplemental table 3). There was no clear relationship between tumor diameter and elevations of six key steroids above upper cut-offs of reference intervals.

Supplemental table 3. Tumor diameters and plasma concentrations of six selected steroids in patients with adrenocortical carcinoma

Patient information				Plasma concentrations (PC) of steroids (PC) and upper cut-offs (UC)												
ID	Age (yr)	Sex (M/F)	BMI kg/m ²	TD (cm)	11-DOC		DOC		AED		17OHP		P		DHEAS	
					PC	UC	PC	UC	PC	UC	PC	UC	PC	UC	PC	UC
1	50	M	27.8	7.6	35.00	0.89	0.273	0.117	9.61	1.92	3.13	1.39	0.12	0.22	2327	3270
2	75	F	22.9	5.9	3.74	0.59	0.127	0.094	0.76	1.10	34.81	0.82	0.56	1.29	1427	1387
3	47	M	38.7	10.0	14.92	0.89	0.688	0.121	1.99	1.96	1.49	1.39	0.18	0.22	1739	3354
4	61	F	27.3	5.1	10.38	0.59	1.028	0.106	2.30	1.42	1.01	0.82	0.38	1.29	2077	2219
5	81	M	24.9	7.6	6.34	0.89	0.130	0.090	3.49	1.65	2.08	1.39	0.18	0.22	617	2475
6	56	F	24.1	8.0	5.60	0.59	0.202	0.110	5.14	1.53	2.48	0.82	0.15	1.29	3913	2395
7	65	M	27.7	12.5	10.10	0.89	0.071	0.102	13.91	1.76	0.83	1.39	0.01	0.22	15182	2896
8	49	M	30.9	9.1	9.68	0.89	1.753	0.119	0.92	1.94	0.40	1.39	0.66	0.22	940	3317
9	86	F	27.1	8.5	4.03	0.59	0.108	0.088	1.07	0.93	0.47	0.82	0.05	1.29	303	617
10	46	M	19.8	12.1	2.92	0.89	0.123	0.122	2.13	1.98	3.14	1.39	0.24	0.22	234	3377
11	71	M	24.9	5.3	0.45	0.89	0.243	0.097	0.54	1.72	1.02	1.39	0.27	0.22	942	2742
12	47	F	26.5	3.5	0.22	0.59	0.085	0.121	0.78	1.79	3.07	2.26	11.04	1.29	734	2692
13	51	F	28.4	10.2	0.54	0.59	0.039	0.116	1.52	1.67	0.37	0.82	0.08	1.29	6144	2582
14	47	M	23.1	6.6	0.56	0.89	0.067	0.121	1.15	1.97	0.60	1.39	0.06	0.22	3395	3366
15	47	M	32.5	5.5	0.55	0.89	0.057	0.121	0.89	1.97	0.78	1.39	0.19	0.22	2312	3361
16	48	M	27.8	9.3	0.55	0.89	0.061	0.119	0.76	1.95	1.35	1.39	0.22	0.22	2638	3326
17	66	M	26.2	6.6	0.51	0.89	0.079	0.101	0.98	1.75	0.70	1.39	0.17	0.22	1812	2864
18	60	M	35.2	10.0	0.25	0.89	0.028	0.107	0.53	1.81	0.60	1.39	0.09	0.22	751	3037
19	70	F	24.9	5.8	0.09	0.59	0.010	0.098	0.40	1.21	0.13	0.82	0.01	1.29	515	1735

Abbreviations: PC, plasma concentration; UC, upper cut-off; BMI, body mass index; TD, tumor diameter; 11-DOC, 11-deoxycortisol; DOC, 11-deoxycorticosterone, AED, androstenedione; 17-OHP, 17-hydroxyprogesterone; P, progesterone; DHEAS, dehydroepiandrosterone sulfate. Plasma concentrations of the six steroids are shown according to age- and/or sex-specific upper cut-offs of reference intervals and where increased above those cut-offs are shown in yellow highlight. Tumor diameters are shown as mean diameters.

Plasma Steroids and Metanephrines in the Five Patient Groups

As detailed in Supplemental table 4 and the following text, there were highly variable significant differences in 13 of 16 steroids of the five patient groups. Plasma concentrations of corticosterone, cortisone and 21-deoxycortisol were the only steroids not to show significant differences.

Supplemental table 4. Plasma concentrations of 16 adrenal steroids and metanephrines among the 577 patients with adrenal incidentaloma

	ACC	ACS	NFAI	PA	PHEO
N	19	104	312	65	77
Aldosterone (ng/L)	36 [18-61]	36 [19-64]	42 [22-76]	156 [83-313]***	44 [22-70]
18-Oxocortisol (ng/L)	16 [10-42]	13 [6-26]	13 [5-28]	73 [31-324]***	10 [5-23]
18-Hydroxycortisol (ng/mL)	0.71 [0.391-0.96]	0.71 [0.45-0.94]	0.63 [0.44-0.94]	1.17 [0.71-2.577]**	0.52 [0.32-0.76]*†
18-Hydroxycorticosterone (ng/mL)	0.31 [0.20-1.46]§	0.43 [0.20-0.77]	0.42 [0.29-0.60]	0.83 [0.40-1.61]*†	0.52 [0.37-0.92]*
11-Deoxycorticosterone (ng/L)	108 [61-243]**	50 [29-75]*	36 [24-56]	51 [33-103]*	41 [23-65]
Corticosterone (ng/mL)	2.10 [1.01-4.05]	1.95 [1.10-3.66]	1.94 [1.23-3.62]	1.97 [1.34-3.29]	2.48 [1.20-3.99]
11-Deoxycortisol (ng/mL)	2.92 [0.51-9.68]**	0.38 [0.25-0.58]*	0.26 [0.16-0.40]	0.29 [0.19-0.47]	0.26 [0.15-0.46]†
Cortisol (ng/mL)	117 [79-184]	123 [93-153]	119 [92-150]	106 [83-131]	134 [96-170]§
Cortisone (ng/mL)	17.7 [14.2-22.9]	19.2 [15.5-22.6]	19.4 [16.0-22.4]	18.3 [14.0-21.6]	19.8 [16.9-23.5]
11-Dehydrocorticosterone (ng/mL)	0.59 [0.40-0.80]*	0.81 [0.45-1.03]*	0.94 [0.67-1.32]	0.80 [0.54-1.18]	1.26 [0.91-1.74]*†¶
21-Deoxycortisol (ng/L)	11 [2-41]	17 [9-33]	19 [9-39]	24 [12-49]	20 [9-33]
Progesterone (ng/L)	180 [83-265]**	35 [17-84]	39 [19-89]	52 [26-106]	45 [20-71]
17-Hydroxyprogesterone (ng/mL)	1.01 [0.60-2.48]**	0.44 [0.22-0.77]	0.46 [0.21-0.79]	0.44 [0.24-0.73]	0.42 [0.21-0.72]
Androstenedione (ng/mL)	1.07 [0.76-2.30]**	0.43 [0.30-0.74]	0.56 [0.37-0.83]	0.61 [0.49-0.91]†	0.60 [0.46-0.83]†
DHEA (ng/mL)	2.33 [1.19-5.13]	1.24 [0.71-2.13]*	1.74 [1.07-3.35]	2.58 [1.48-4.53]†	2.73 [1.51-3.97]†
DHEAS (ng/mL)	1739 [734-2637]*	393 [212-648]****	655 [341-1220]	867 [494-1482]	929 [506-1505]
Normetanephrine (ng/L)	91 [57-105]†	92 [67-128]	79 [64-111]	71 [48-96]	627 [276-1882]***
Metanephrine (ng/L)	27 [17-32]	31 [21-45]	28 [18-41]	27 [21-42]	170 [40-482]***

Plasma concentrations are shown in ng/mL or ng/L. ***P<0.001, **<0.02 higher than all groups; ****P<0.0001 lower than all groups; *P<0.02 different from NFAI; †P<0.01 different from ACS; §P<0.05 different from PA; ¶P<0.02 different from ACC. Table excludes data for testosterone, dihydrotestosterone and methoxytyramine. For conversion of ng/mL to nmol/mL or ng/L to nmol/L divide by molecular weight (Supplemental table 5).

Supplemental table 5. Formulae and molecular weights (MW) of the 19 adrenal steroids measured by LC-MS/MS

Analytes	Formula	MW
Steroids		
Aldosterone	C ₂₁ H ₂₈ O ₅	360.44
18-Oxocortisol	C ₂₁ H ₂₈ O ₆	376.45
18-Hydroxycortisol	C ₂₁ H ₃₀ O ₆	378.46
18-Hydroxycorticosterone	C ₂₁ H ₃₀ O ₅	362.46
11-Deoxycorticosterone	C ₂₁ H ₃₀ O ₃	330.46
Corticosterone	C ₂₁ H ₃₀ O ₄	346.46
11-Deoxycortisol	C ₂₁ H ₃₀ O ₄	346.46
Cortisol	C ₂₁ H ₃₀ O ₅	362.46
Cortisone	C ₂₁ H ₂₈ O ₅	360.44
11-Dehydrocorticosterone	C ₂₁ H ₂₈ O ₄	344.46
21-Deoxycortisol	C ₂₁ H ₃₀ O ₄	346.46
Progesterone	C ₂₁ H ₃₀ O ₂	314.46
17-Hydroxyprogesterone	C ₂₁ H ₃₀ O ₃	330.46
Androstenedione	C ₁₉ H ₂₆ O ₂	286.41
DHEA	C ₁₉ H ₂₈ O ₂	288.42
DHEA-Sulfate	C ₁₉ H ₂₈ O ₅ S	368.49
Testosterone	C ₁₉ H ₂₈ O ₂	288.42
Dihydrotestosterone	C ₁₉ H ₃₀ O ₂	290.44
Dexamethasone	C ₂₂ H ₂₉ FO ₅	392.46
Catecholamine O-methylated metabolites		
Normetanephrine	C ₉ H ₁₃ NO ₃	183.20
Metanephrine	C ₁₀ H ₁₅ NO ₃	197.23
Methoxytyramine	C ₉ H ₁₃ NO ₂	167.21

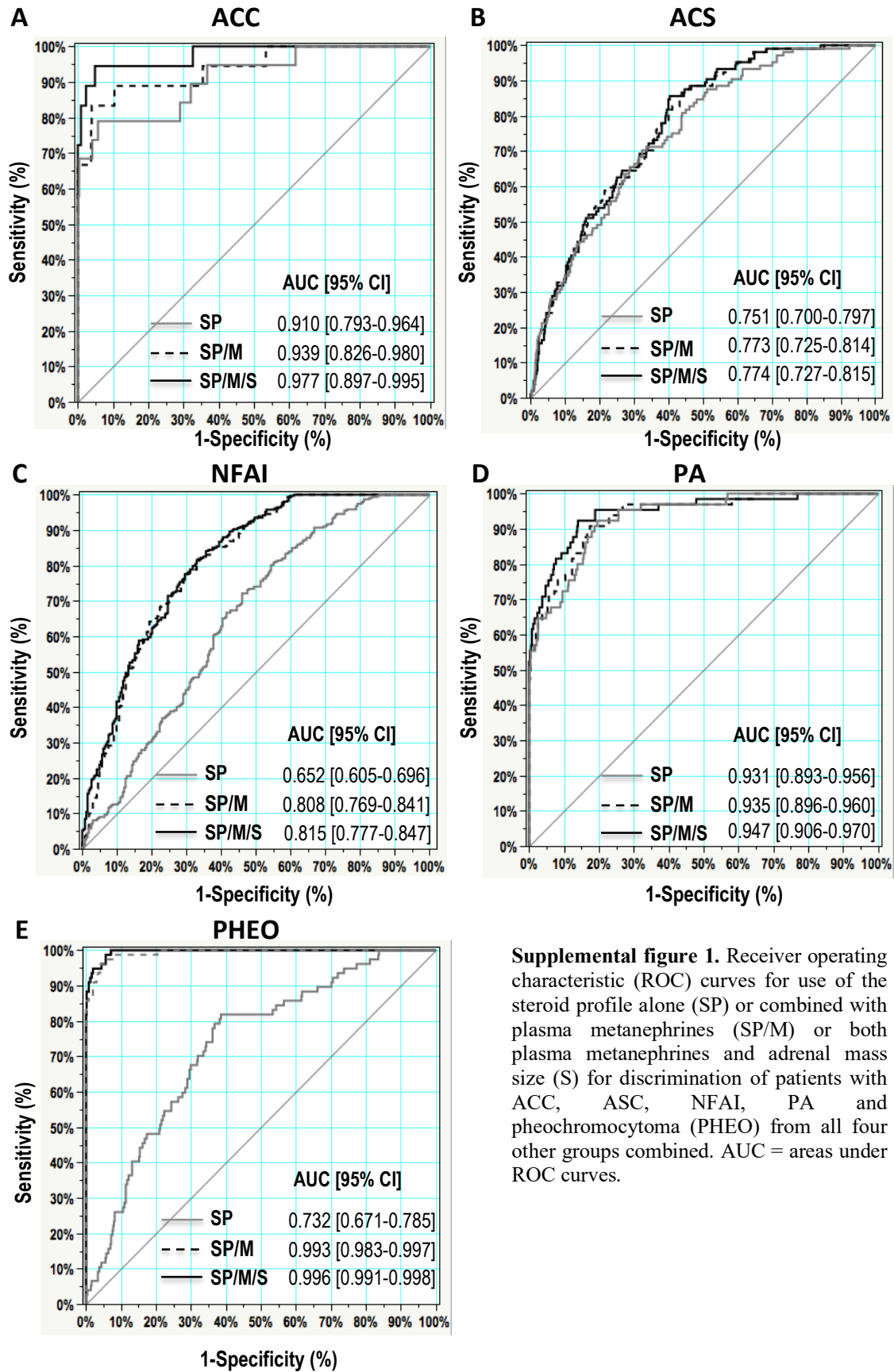
Without adjustment for differences in age and sex plasma concentrations of aldosterone and 18-oxocortisol were higher ($P < 0.001$) in patients with PA than all other groups (Supplemental table 4). Plasma concentrations of 18-hydroxycortisol were also higher ($P < 0.02$) in patients with PA than other groups, and this steroid showed additional lower ($P < 0.02$) concentrations in patients with pheochromocytoma than in those with ACS and NFAI. Plasma concentrations of 18-hydroxycorticosterone and 11-deoxycorticosterone were also higher ($P < 0.02$) in patients with PA than those with NFAI; for the latter steroid, 11-deoxycorticosterone, plasma concentrations were also higher ($P < 0.02$) in patients with ACS than NFAI, but this steroid showed additional higher ($P < 0.02$) concentrations in patients with ACC than all other groups. Plasma 11-deoxycortisol was considerably higher ($P < 0.001$) in patients with ACC than all other groups and showed additional higher concentrations in patients with ACS than in those with NFAI and pheochromocytoma. Plasma concentrations of cortisol were higher ($P < 0.05$) in patients with pheochromocytoma than those with PA. In

contrast, plasma concentrations of 11-dehydrocorticosterone in patients with pheochromocytoma were higher ($P < 0.02$) than in those with ACC, ACS and NFAI, but not compared to PA, and showed additionally lower ($P < 0.02$) concentrations in patients with ACC and ACS than those with NFAI.

Progesterone, 17-hydroxyprogesterone and androstenedione all showed higher ($P < 0.02$) plasma concentrations in patients with ACC than other groups. The latter steroid, androstenedione, also showed higher ($P < 0.05$) plasma concentrations in patients with pheochromocytoma and PA than in those with ACS. Plasma concentrations of DHEA were also higher ($P < 0.05$) in patients with pheochromocytoma than in those with ACS, and showed additionally higher ($P < 0.05$) concentrations in patients with NFAI and PA than in patients with ACS. Furthermore, while plasma concentrations of DHEAS were higher in all groups than in patients with ACS, this steroid also showed higher ($P < 0.05$) plasma concentrations in patients with ACC than in NFAI. In contrast to the steroids, plasma concentrations of normetanephrine and metanephrine were higher ($P < 0.001$) in patients with pheochromocytoma than all other groups.

Logistic Regression

Logistic regression analysis using nine steroids (11-deoxycortisol, 18-oxocortisol, androstenedione, 18-hydroxycortisol, corticosterone, DHEAS, cortisol, 17-hydroxyprogesterone and aldosterone) that were selected based on stepwise variable selection and construction of receiver operating characteristic (ROC) curves indicated best discrimination for the steroid panel alone was for patients with PA versus all others (Supplemental figure 1).



Supplemental figure 1. Receiver operating characteristic (ROC) curves for use of the steroid profile alone (SP) or combined with plasma metanephrines (SP/M) or both plasma metanephrines and adrenal mass size (S) for discrimination of patients with ACC, ACS, NFAI, PA and pheochromocytoma (PHEO) from all four other groups combined. AUC = areas under ROC curves.

The area under the ROC curve, with 95% confidence intervals (CIs), for discrimination of patients with PA from all other patients was 0.931 [0.893-0.956] and was not significantly increased by inclusion in models of measurements of plasma metanephrines or plasma metanephrines and adrenal

mass size (panel D). Next best discrimination using the steroid panel along was observed for patients with ACC (panel A) where the AUC was 0.910 [0.793-0.964]. Although the AUC increased to 0.977 [0.897-0.955], this increase failed to reach significance ($P=0.0574$) according to 95% CIs of the difference in AUCs [-0.001-0.0779].

AUCs for steroid panels used alone were 0.751 [0.700-0.797], 0.732 [0.674-0.785] and 0.652 [0.605-0.696] for the respective identification of patients with ACS (panel B), pheochromocytoma (panel E) and NFAI (panel C) among all other groups. As expected, addition of plasma metanephrines to models considerably improved ($P<0.0001$) identification of patients with pheochromocytoma according to an AUC of 0.993 [0.928-0.997] and 95% CIs of the difference in AUCs [0.205-0.317]. The AUC was little affected thereafter by additional inclusion of tumor size to the model (panel E).

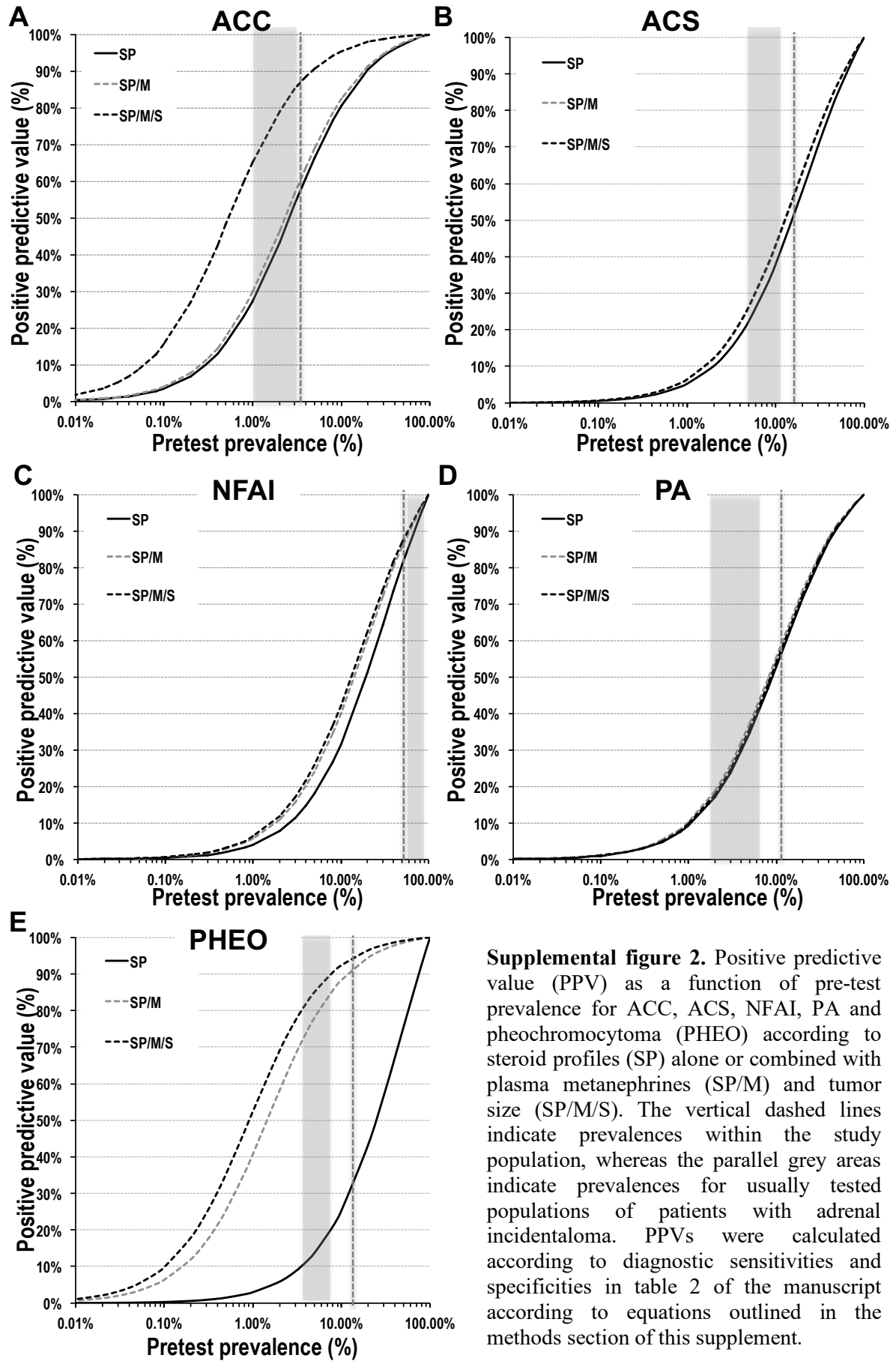
Addition of plasma metanephrines to models also considerably improved ($P<0.0001$) diagnostic performance for identification of patients with NFAI compared to all others (panel C) according to an AUC of 0.808 [0.769-0.841] and 95% CIs of the difference in AUCs [0.115-0.197]. Thereafter, addition of tumor size to the model had negligible influence on diagnostic performance.

Addition of plasma metanephrines to the models resulted in a small but significant improvement ($P=0.0100$) also for discrimination of patients with ACS from all other groups (panel B), as indicated by an increase in the AUC to 0.773 [0.725-0.814] and 95% CIs of the difference in AUCs [0.006-0.041]. Again thereafter, addition of tumor size to the model had negligible influence on diagnostic performance.

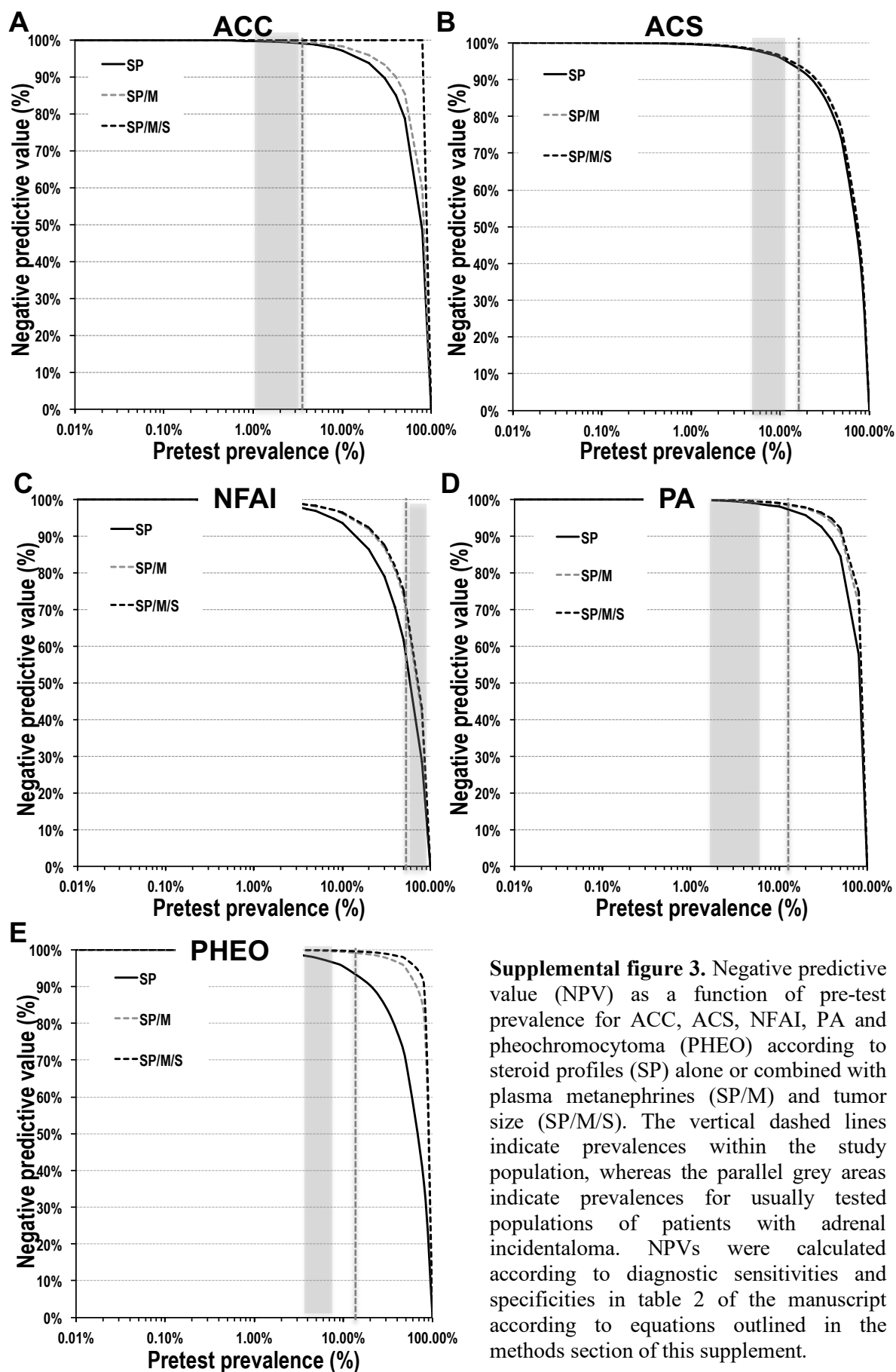
PPV and NPV According to Disease Prevalence

As expected, positive PPV increased with increasing prevalence of the different types of tumors among patients with incidentaloma, whereas NPV decreased with prevalence (Supplemental figures 2&3). As also indicated in table 2 of the manuscript, the PPV for use of the plasma steroid profile to identify patients with ACC was significantly impacted by inclusion of tumor size in the model (Supplemental figure 1A); furthermore, this impact was apparent at all pretest prevalences of disease, including those below 3% expected in population-based inclusion of patients. Also, in line with results of logistic regression (supplemental figure 1) and table 2 of the manuscript, only inclusion of measurements of plasma metanephrines had any impact on the PPV for steroid profiling in patients with ACS and NFAI; however, for ACS the impact was minimal (Supplemental figure 2B&C). For use of steroid profiles to identify patients with PA (Supplemental figure 2D), there was no additional impact of measurements of plasma metanephrines or tumor size on PPV beyond that derived from steroid profiling; PPVs were lower (14-46%) at prevalences of PA expected with population-based sampling (e.g., 2-6%) than according to the proportion of the study population (11.3%) for which PPVs were 56-59%. For pheochromocytoma (Supplemental figure 2E), the PPV for the combination of all three sets of variables in the model was also lower (80-88%) at expected prevalences (4-6%) of the tumor among patients with adrenal incidentaloma than at proportion (13.3%) of the present study where the PPV was 95% (Table 2).

In contrast to PPVs, NPVs were little impacted by differences in prevalence of the different types of tumors up until prevalences were above 3%, with impacts most apparent above 20% (Supplemental figure 3). For ACC and functional adrenal tumors, and in contrast to PPVs, NPVs were calculated to be higher at prevalences of these tumors expected according to population-based inclusion of patients with incidentaloma compared to the proportions of the study population. On the other hand for NFAI, NPVs were calculated to be lower at expected prevalences among patients with adrenal incidentaloma than those for the present study population. At prevalences expected among patients with adrenal incidentaloma, the addition of metanephrines to steroid profiles had most impact on NPVs for patients with NFAI and pheochromocytoma (Supplemental figure 3, panels C&E).



Supplemental figure 2. Positive predictive value (PPV) as a function of pre-test prevalence for ACC, ACS, NFAI, PA and pheochromocytoma (PHEO) according to steroid profiles (SP) alone or combined with plasma metanephrines (SP/M) and tumor size (SP/M/S). The vertical dashed lines indicate prevalences within the study population, whereas the parallel grey areas indicate prevalences for usually tested populations of patients with adrenal incidentaloma. PPVs were calculated according to diagnostic sensitivities and specificities in table 2 of the manuscript according to equations outlined in the methods section of this supplement.



Supplemental figure 3. Negative predictive value (NPV) as a function of pre-test prevalence for ACC, ACS, NFAI, PA and pheochromocytoma (PHEO) according to steroid profiles (SP) alone or combined with plasma metanephrines (SP/M) and tumor size (SP/M/S). The vertical dashed lines indicate prevalences within the study population, whereas the parallel grey areas indicate prevalences for usually tested populations of patients with adrenal incidentaloma. NPVs were calculated according to diagnostic sensitivities and specificities in table 2 of the manuscript according to equations outlined in the methods section of this supplement.

Supplemental discussion of study limitations

Population Bias

As outlined in the earlier section entitled “Population-based versus Pragmatic Patient Inclusion”, the present study was not population based. The study instead took advantage of patients with adrenal incidentaloma recruited into three clinical protocols, two of which had a specific focus on either pheochromocytoma and paraganglioma or PA. The consequence of this was that relative numbers of recruited patients with the various types of adrenal lesions (ACC, ACS, NFAI, PA and pheochromocytoma) did not match those expected with unbiased population-based inclusion. In particular, the percentages of patients with certain functional tumors, such as pheochromocytoma (13.3%) and PA (11.3%), were higher than expected in unselected patients with adrenal incidentaloma. As a reciprocal consequence of this, the proportion of patients with NFAI (54.1%) was lower than expected in unselected patients with adrenal incidentaloma.

Such population imbalances are true of almost all diagnostic studies of adrenal tumors and particularly those involving retrospective analyses of rare tumors such as pheochromocytoma and ACC. For example, as outlined in the European Society of Endocrinology clinical practice guideline on management of adrenal incidentaloma (4), the reported median 8% frequency of ACC among incidentaloma was likely impacted by selection bias of the available retrospective series. Even for large prospective studies with a focus on diagnostics in rare adrenal tumors it can be difficult to achieve population-based sampling free from selection bias of referral centers. For example in multicenter study of urine steroidomics in ACC by Bancos et al (5), seven of the 21 participating centers recruited patients with ACC at a median proportion of 35% of included patients. Although those cases of clear selection bias were excluded, that report does illustrate the difficulty of avoiding population bias even in prospective series involving rare adrenal tumors.

As clarified in the preceding section entitled “PPV and NPV according to Disease Prevalence”, the key limitation associated with not following population-based recruitment of patients is that PPVs are overestimated and NPV are underestimated for over-represented populations; in contrast, for under-represented populations, PPVs are underestimated and NPV are overestimated. This shortcoming, however, does not impact estimates of diagnostic performance according to sensitivity, specificity and areas under ROC curves. In fact population-based recruitment of patients can be a study limitation in itself for studies of rare diseases where it may not be possible to recruit sufficient patients for accurate estimates of diagnostic sensitivity. For example, as outlined in the PMT study (1), a multicenter prospective study of patients tested for PPGLs, recruitment of the required 200 patients with disease would have required an impossible total of about 20,000 patients according to a population-based approach; alternatively for more realistic population-based recruitment of 2,000 patients those with disease would have been limited to a population of only about 20, too low for reliable estimates of diagnostic sensitivity and from this assessments of any differences in diagnostic performance among the various tests. As outlined for the PMT study (1), the solution for non-population based approaches is to calculate PPVs and NPVs at different prevalences of disease, as was also done in the present study.

To clarify the extent population bias is likely to have impacted estimates of PPV and NPV, the data display in supplemental figures 2 and 3 include grey area vertical bands to show likely population-based prevalences of adrenal lesions in relation to the proportions of the present study in dashed vertical lines (Supplemental figures 2&3). Those likely prevalences were based on two previous studies of adrenal incidentaloma involving a little over 1,000 patients each (6,7). Those two studies indicated prevalences of between 1.0-4.5% for ACC, 5.0-9.2% for ACS, 83.3-85.0% for NFAI, 1.6-6.1% for PA and 4.2-6.0% for pheochromocytoma.

Retrospective vs Prospective Study Designs and Patient Follow-up

As outlined in several sections of the manuscript discussion section, the more significant and serious limitation of this study did not involve population bias, but rather over-reliance on routine diagnostic testing to classify disease; also for most adrenal lesions there remains a lack of “gold standards” or alternative criteria to routine diagnostic testing to further confirm or exclude disease. For some rare adrenal lesions, such as ACC and pheochromocytoma, histopathology offers an acceptable “gold-standard” for disease confirmation; however, such gold standards are not yet available for other more

common adrenal tumors and cannot be applied to the vast majority of patients in whom ACC and pheochromocytoma have been excluded by initial diagnostic testing. That testing, even for pheochromocytoma, is not 100% sensitive and there will always be false-negatives who can only be identified by follow-up. Without that, diagnostic sensitivity and PPV indicated according to routine diagnostic approaches will always be over-estimated and NPV underestimated. This also applies to new experimental procedures, such as steroid profiling, when there is no follow-up to confirm or exclude disease. For such new procedures applied in these situations diagnostic performance can only be expected to approach and not exceed performance of routine tests, and particularly so when the performance of the routine tests that we rely in for disease classification is suboptimal.

As we have also documented here, even with follow-up, positive tests can be misinterpreted as false-positives and it may only be after prolonged follow-up that a patient is eventually identified to have disease and a true-positive rather than a false-positive test result is then revealed at initial testing over many years previously. In such instances specificity is underestimated, with additional inaccuracies in estimates of PPV and NPV for previous testing.

The aforementioned issues are even more problematic for the diagnostic classifications of patients with PA and ACS. For patients with these lesions, even when resected there are currently no routinely available histopathological procedures to reliably confirm disease. For patients with PA, when an adrenal or adrenal nodule is resected the only currently accepted method to confirm unilateral disease involves follow-up biochemical testing to establish biochemical cure (8). This is a relatively recently outlined approach and for the current series is only being routinely applied to patients enrolled by way of the PROSALDO trial. That trial, which is still underway, has already revealed diagnostic inaccuracies of routine diagnostic tests involving immunoassay measurements of aldosterone (9); these measurements can be impacted by interferences causing false positive test results for confirmatory tests that, without histopathology or alternative methods, remain the methods that are solely relied upon for confirmation of PA. From a single patient outlined in that initial report, we are now finding that up to 50% or more of patients with presumed bilateral hyperplasia forms of PA, based on routine procedures, do not in fact have positive confirmatory tests based on other procedures (unpublished results).

For diagnostic stratification of ACS the situation is likely worse than for PA, although at least for the key test in this diagnosis, the DST, measurements of cortisol are invariably accurate by immunoassays. As indicated in the prospective study of Goh et al (10), while the DST is reasonably sensitive for indicating ACS, on follow-up of patients with positive initial tests there were subsequent normal test results in 44% of patients. This indicates a substantial proportion of false-positive results in patients with diagnosed ASC based on the DST alone. Since in the present study classification of ACS was based on a positive DST it is not surprising that diagnostic performance of the combination of steroid profiling and metanephrines for these patients was considerably inferior to that for patients with other functional tumors.

For the remaining patients with NFAI, the problems with a correct classification are potentially more difficult to navigate than for functional tumors, at least for the smaller lesions that remain unoperated. Larger tumors are usually resected, irrespective of evidence of functionality, and for many of these, such as myelolipomas, ganglioneuromas, hemangiomas, a diagnosis can be reached to ensure correct classification of NFAI. However, these represent a minority, and most lesions are small. In the present study the median average diameter was 1.8 cm. Thus, the vast majority of patients with a classification of NFAI remain unoperated.

Although for NFAI a high lipid content and other imaging characteristics can be of assistance in indicating an adenoma rather than a pheochromocytoma or ACC (11-13), this is not particularly useful for patients with ACS or PA and is mainly relevant for larger adrenal lesions. As shown by Akbulut et al (14), for smaller adrenal incidentaloma, CT imaging and washout characteristics can be similar in adenoma and pheochromocytoma. This is in further agreement with a case described in the present report of a small adrenal lesion that was initially classified as an adenoma based on initial imaging characteristics; it then took ten years for the tumor to be diagnosed as a pheochromocytoma after reaching a larger size when it produced clinical manifestations of catecholamine excess, a larger diagnostic signal and imaging characteristics that were no longer consistent with an adenoma.

Also as shown by Gagnon et al (15), small tumor size cannot be used to definitively exclude ACC. In that report the patient presented initially with a 2.9x1.9 cm incidentally discovered adrenal mass and no evidence of hormonal activity. The mass remained stable over yearly follow-up for another 4 years. However, 10 years later after presenting with renal colic the mass was found to have grown to 9x8.2 cm and the patient was diagnosed with ACC with hepatic metastases.

The above case along with the early case of pheochromocytoma in this report illustrates the importance of appropriate follow-up of patients with apparent NFAI. In the present series, most of the patients with NFAI were not followed up beyond a few years. This is in line with current guidelines that indicate if an adrenal mass shows no evidence of hormonal activity and remains stable in size and less than 4 cm after repeated imaging then there is no need for further follow-up imaging studies (4).

Among the 312 patients with a classification of NFAI there were significant numbers of patients with predictions that suggested other tumors (Table 2). Without appropriate follow-up it was not possible to assess whether any of these predictions were in fact correct; thus, diagnostic performance may have been underestimated. Indeed, among the NFAI there were reasonable numbers of patients with plasma concentrations of steroids elevated above age and/or sex specific upper cut-offs, including some with multiple elevations suggesting hormonal activity (Supplemental table 6).

Supplemental table 6. Plasma concentrations of selected steroids in patients with a classification of NFAI in whom at least three key steroids were elevated above upper cut-offs of reference intervals

ID	Age yr	Sex (M/F)	BMI kg/m ²	TD cm	11-DOC ng/mL	DOC ng/mL	CORT ng/mL	17-OHP ng/mL	P ng/mL	21-DOC ng/mL	AED ng/mL	DHEA ng/mL
1	52	F	28.1	2.8	1.7	0.2	5.7	35.0	3.4	0.6	2.6	5.6
				<i>Upper cut-offs</i>	0.6	0.1	6.3	0.8	1.3	0.1	1.6	6.6
2	63	M	25.3	1.7	1.4	0.1	8.9	0.9	0.2	0.2	1.1	4.1
				<i>Upper cut-offs</i>	0.9	0.1	6.8	1.4	0.2	0.2	1.8	5.6
3	56	M	26.3	0.8	1.4	0.2	10.7	1.1	0.1	0.0	1.2	3.9
				<i>Upper cut-offs</i>	0.9	0.1	7.5	1.4	0.2	0.2	1.9	6.7
4	50	M	24.2	5.0	1.2	0.2	4.7	1.5	0.1	0.1	1.1	4.8
				<i>Upper cut-offs</i>	0.9	0.1	8.3	1.4	0.2	0.2	1.9	7.8
5	77	M	29.3	1.1	1.1	0.1	7.4	0.8	0.1	0.1	1.1	2.5
				<i>Upper cut-offs</i>	0.9	0.1	5.8	1.4	0.2	0.2	1.7	3.9
6	51	M		3.0	1.0	0.2	21.4	2.1	0.4	0.5	1.9	16.2
				<i>Upper cut-offs</i>	0.9	0.1	8.1	1.4	0.2	0.2	1.9	7.5
7	56	F	32.8	5.1	0.7	0.1	1.9	1.4	0.3	0.7	0.1	0.3
				<i>Upper cut-offs</i>	0.6	0.1	5.5	0.8	1.3	0.1	1.5	6.0
8	66	M	27.8	2.2	0.8	0.1	7.0	3.0	0.2	1.2	1.2	6.8
				<i>Upper cut-offs</i>	0.9	0.1	6.6	1.4	0.2	0.2	1.8	5.2
9	52	M	24.0	1.1	0.6	0.1	7.1	2.7	0.5	0.8	1.1	2.2
				<i>Upper cut-offs</i>	0.9	0.1	8.0	1.4	0.2	0.2	1.9	7.4

Abbreviations: BMI, body mass index; TD, tumor diameter; 11-DOC, 11-deoxycortisol; DOC, 11-deoxycorticosterone; CORT, corticosterone; 17-OHP, 17-hydroxyprogesterone; P, progesterone; 21-DOC, 21-deoxycortisol; AED, androstenedione; DHEA, dehydroepiandrosterone. Plasma concentrations of the six steroids are shown according to age- and/or sex-specific upper cut-offs of reference intervals and where increased above those cut-offs are shown in yellow highlight. Tumor diameters are shown as mean diameters.

Future Perspectives and avoidance of Study Limitations

With consideration of the study limitations outlined above, it is of no surprise that the combinations of steroid profiling, plasma free metanephrines and tumor size were less than optimal for correct classification of some adrenal tumors, particularly NFAI, ACS and to a lesser extent PA. Diagnostic classification of these adrenal lesions largely depends on routine endocrine testing, which is currently far from accurate. To overcome this limitation, there is need for a fully prospective study design in which alternatives to routine diagnostic tests must be employed to improve upon classifications. Such alternatives may include, as outlined here, multidimensional approaches that include imaging characteristics, steroid profiles and metanephrines aided by artificial intelligence to provide probabilities of disease. However, even then there must be alternative and independent confirmatory tests and/or “gold standards” for final classification, the latter likely to be only achieved by long-term follow-up of all patients for whom the diagnosis is not fully accurate and 100% clear.

In contrast to NFAI, ACS and PA, diagnostic classifications achieved using the combination of plasma steroids, metanephrines and tumor size were more accurate for pheochromocytoma and ACC; this reflects in part accurate final confirmation of disease by histopathology; for pheochromocytoma, there is additional diagnostic accuracy provided by use of plasma free metanephrines for identifying and excluding disease (1). For identification of ACC, diagnostic performance was significantly improved by inclusion of tumor size into models. However, as outlined in the manuscript it would be more ideal not to rely on this parameter so that these malignant tumors can be identified at an earlier stage. For that purpose, imaging characteristics would provide a more appropriate parameter than lesion size. For that again, a carefully designed fully prospective study is necessary, which for pragmatic purposes should take into account the relevant imaging characteristics for both MRI and CT. Without a fully prospective study that provides for long-term follow-up to confirm and exclude all types of adrenal tumors it is not possible to appropriately establish the utility of steroid profiling or any other method or combination of methods for detection of ACC or other tumors; this is particularly important for detection at an earlier stage when there may be improved prognosis for patients achieved by timely intervention.

Although the PMT study had a prospective design and did involve follow-up, this was primarily focused at excluding or confirming pheochromocytoma rather than other tumors. Similarly, the PROSALDO trial also features follow-up, but again this follow-up is directed at further confirming or excluding PA rather than other tumors. The present study took advantage of the data and banked samples from the PMT study, the PROSALDO trial and the ENSAT-registry and biobanking protocol; thus, the present study was not specifically designed for the follow-up of all adrenal incidentaloma and can only be classified as retrospective in nature. For a fully prospective clinical trial that covers the diagnostic utility of any novel method or combination of methods to simultaneously and efficiently identify all types of adrenal incidentaloma, the study design must take into account and address the limitations outlined here. Furthermore, to move beyond whether any new method provides improved diagnosis to whether there are also improved outcomes for patients, then the design should include some kind of randomization of new and routine procedures with inclusion of appropriate outcome measures to assess differences in patient outcomes.

References

1. Eisenhofer G, Prejbisz A, Peitzsch M, et al. Biochemical diagnosis of chromaffin cell tumors in patients at high and low risk of disease: Plasma versus urinary free or deconjugated o-methylated catecholamine metabolites. *Clin Chem* 2018; 64:1646-1656.
2. Eisenhofer G, Peitzsch M, Kaden D, et al. Reference intervals for plasma concentrations of adrenal steroids measured by LC-MS/MS: Impact of gender, age, oral contraceptives, body mass index and blood pressure status. *Clin Chim Acta* 2017; 470:115-24.
3. Eisenhofer G, Peitzsch M, Kaden D, et al. Reference intervals for LC-MS/MS measurements of plasma free, urinary free and urinary acid-hydrolyzed deconjugated normetanephrine, metanephrine and methoxytyramine. *Clin Chim Acta* 2019; 490:46-54.
4. Fassnacht M, Arlt W, Bancos I, et al. Management of adrenal incidentalomas: European society of endocrinology clinical practice guideline in collaboration with the european network for the study of adrenal tumors. *Eur J Endocrinol* 2016; 175:G1-G34.

5. Bancos I, Taylor AE, Chortis V, et al. Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the eurine-act study: A prospective test validation study. *Lancet Diabetes Endocrinol* 2020; 8:773-781.
6. Mantero F, Terzolo M, Arnaldi G, et al. A survey on adrenal incidentaloma in Italy. Study Group on Adrenal Tumors of the Italian Society of Endocrinology. *J Clin Endocrinol Metab.* 2000; 85:637-44.
7. Ahn SH, Kim JH, Baek SH et al. Characteristics of Adrenal Incidentalomas in a Large, Prospective Computed Tomography-Based Multicenter Study: The COAR Study in Korea *Yonsei Med J.* 2018; 59:501-510.
8. 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.
9. 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.
10. Goh Z, Phillips I, Hunt PJ, Soule S, Cawood TJ. Three-year follow up of adrenal incidentalomas in a New Zealand centre. *Intern Med J* 2020; 50:350-356
11. Marty M, Gaye D, Perez P, et al. Diagnostic accuracy of computed tomography to identify adenomas among adrenal incidentalomas in an endocrinological population. *Eur J Endocrinol* 2018; 178:439-446.
12. Petersenn S, Richter PA, Broemel T, R et al. Computed tomography criteria for discrimination of adrenal adenomas and adrenocortical carcinomas: analysis of the German ACC registry. *Eur J Endocrinol.* 2015;172:415-22.
13. Canu L, Van Hemert JAW, Kerstens MN, et al. Ct characteristics of pheochromocytoma: Relevance for the evaluation of adrenal incidentaloma. *J Clin Endocrinol Metab* 2019; 104:312-318.
14. Akbulut S, Erten O, Kahramangil B, et al. A Critical analysis of computed tomography washout in lipid-poor adrenal incidentalomas. *Ann Surg Oncol.* 2021; 28:2756-2762.
15. Gagnon N, Boily P, Alguire C, et al. Small adrenal incidentaloma becoming an aggressive adrenocortical carcinoma in a patient carrying a germline APC variant. *Endocrine.* 2020; 68:203-209.