Use of Steroid Profiling Combined With Machine Learning for Identification and Subtype Classification in Primary Aldosteronism

Graeme Eisenhofer, PhD; Claudio Durán, MS; Carlo Vittorio Cannistraci, PhD; Mirko Peitzsch, PhD; Tracy Ann Williams, PhD; Anna Riester, MD; Jacopo Burrello, MD; Fabrizio Buffolo, MD; Aleksander Prejbisz, MD; Felix Beuschlein, MD; Andrzej Januszewicz, MD; Paolo Mulatero, MD; Jacques W. M. Lenders, MD; Martin Reincke, MD

Abstract

IMPORTANCE Most patients with primary aldosteronism, a major cause of secondary hypertension, are not identified or appropriately treated because of difficulties in diagnosis and subtype classification. Applications of artificial intelligence combined with mass spectrometry-based steroid profiling could address this problem.

OBJECTIVE To assess whether plasma steroid profiling combined with machine learning might facilitate diagnosis and treatment stratification of primary aldosteronism, particularly for patients with unilateral adenomas due to KCNJ5 sequence variants.

DESIGN, SETTING, AND PARTICIPANTS This diagnostic study was conducted at multiple tertiary care referral centers. Steroid profiles were measured from June 2013 to March 2017 in 462 patients tested for primary aldosteronism and 201 patients with hypertension. Data analyses were performed from September 2018 to August 2019.

MAIN OUTCOMES AND MEASURES The aldosterone to renin ratio and saline infusion tests were used to diagnose primary aldosteronism. Subtyping was done by adrenal venous sampling and follow-up of patients who underwent adrenalectomy. Statistical tests and machine-learning algorithms were applied to plasma steroid profiles. Areas under receiver operating characteristic curves, sensitivity, specificity, and other diagnostic performance measures were calculated.

RESULTS Primary aldosteronism was confirmed in 273 patients (165 men [60%]; mean [SD] age, 51 [10] years), including 134 with bilateral disease and 139 with unilateral adenomas (58 with and 81 without somatic KCNJ5 sequence variants). Plasma steroid profiles varied according to disease subtype and were particularly distinctive in patients with adenomas due to KCNJ5 sequence variants, who showed better rates of biochemical cure after adrenalectomy than other patients. Among patients tested for primary aldosteronism, a selection of 8 steroids in combination with the aldosterone to renin ratio showed improved effectiveness for diagnosis over either strategy alone. In contrast, the steroid profile alone showed superior performance over the aldosterone to renin ratio for identifying unilateral disease, particularly adenomas due to KCNJ5 sequence variants. Among 632 patients included in the analysis, machine learning–designed combinatorial marker profiles of 7 steroids alone both predicted primary aldosteronism in 1 step and subtyped patients with unilateral adenomas due to KCNJ5 sequence variants at diagnostic sensitivities of 69% (95% CI, 68%-71%) and 85% (95% CI, 81%-88%), respectively, and at specificities of 94% (95% CI, 93%-94%) and 97% (95% CI, 97%-98%), respectively. The validation series yielded comparable diagnostic performance.

CONCLUSIONS AND RELEVANCE Machine learning–designed combinatorial plasma steroid profiles may facilitate both screening for primary aldosteronism and identification of patients with unilateral adenomas due to KCNJ5 sequence variants.
Abstract (continued)
adenomas due to KCNJ5 sequence variants, who are most likely to show benefit from surgical intervention.


Introduction
Applications of artificial intelligence, including machine learning, are gaining increasing recognition for informing medical decision-making.1-4 Machine learning may be particularly useful in heterogeneous disorders where there is a need for stratification to guide therapy.5-8 One such disorder is primary aldosteronism (PA), a common cause of secondary hypertension with 2 main subtypes for which treatment stratification is crucial but difficult.9,10 With a prevalence of 5% to 7% among unselected patients with hypertension and up to 20% among patients with severe hypertension, PA affects large numbers of patients and is associated with considerable morbidity exceeding that of patients with primary hypertension (PHT) and similar elevations of blood pressure.11,12

The aforementioned considerations highlight the importance of effective methods for diagnosis and treatment of PA, which must allow for stratification according to unilateral vs bilateral hypersecretion of aldosterone.9,10 Cure of the former can be achieved by adrenalectomy, whereas mineralocorticoid receptor antagonists are indicated for the bilateral subtype. Attaining this stratification is not simple and usually requires adrenal venous sampling (AVS), a technically demanding, expensive, time-consuming, and not infallible procedure.9,13-16 In 2 independent studies,13,16 discordant lateralization results were observed in 24% to 28% of patients who underwent AVS with vs without adrenocorticotropin. In another study,14 clinical outcomes did not differ according to determination of unilateral disease by AVS vs radiological imaging. In a fourth study,15 there were no significant differences in rates of biochemical cure (76% vs 69%) in patients younger than 65 years who underwent adrenalectomy according to AVS lateralization ratios larger vs smaller than 4.

Apart from the difficulties and limited effectiveness of AVS for subtype classification, there are also problems with earlier steps in the diagnosis of PA. Although the aldosterone to renin ratio (ARR) offers a time-honored method for screening, there is considerable overlap of ratios among patients with and without PA,17,18 thus, at ARR cutoffs selected to optimize diagnostic sensitivity, there are many false-positives, leading to the need for confirmatory studies.19,20 Such multiple steps, poor standardization, requirements to consider antihypertensive medications, and difficulties with AVS all represent barriers to diagnostic stratification; consequently, most patients remain undiagnosed and are not appropriately treated.9,21 Improved approaches for diagnostic stratification are therefore needed.

With the aforementioned considerations in mind, we examined the use of mass spectrometry-based steroid profiling combined with machine learning for diagnostic stratification, with the hypothesis that this approach at screening might facilitate case detection and also allow for subtype classification. This hypothesis was based on findings that distinct steroid profiles in adrenal venous plasma of patients with bilateral and unilateral PA translated to similarly distinct profiles in peripheral plasma.22 Patients with unilateral aldosterone-producing adenomas (APAs) due to somatic sequence variants of KCNJ5 have particularly distinct steroid profiles.23 These patients also have larger and more clearly visualized APAs and show the most favorable outcomes after adrenalectomy.24-27 The use of steroid profiles to identify these patients may, therefore, be especially useful. Thus, the primary objective of this study was to establish whether steroid profiling could facilitate both identification and subtype classification of patients with PA, particularly those with unilateral APAs due to KCNJ5 sequence variants.
Methods

Patients
This diagnostic study was approved by the Klinikum der Ludwig-Maximilians-Universität München, University of Turin, Technische Universität Dresden, and Institute of Cardiology (Warsaw). The study follows the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guideline. All patients provided written informed consent under protocols approved by ethics committees at the 4 tertiary clinical care centers where patients were referred for testing.

The study involved mass spectrometry–based steroid profiling of plasma specimens from 462 patients tested for primary aldosteronism between June 13, 2013, and March 8, 2017. Follow-up of patients was completed by July 31, 2018. Patient data and specimens were derived from studies and registries with a focus on hypertension and PA, including the Conn registry in Germany, the European Network for Studies of Adrenal Tumors registry, and the Prospective Monoamine-Producing Tumor study.

Patients were tested for PA according to 1 or more of several criteria: office blood pressure greater than 150/100 mm Hg, therapy-resistant hypertension, or hypertension associated with hypokalemia or hemorrhagic stroke, an adrenal incidentaloma, or obstructive sleep apnea. Other forms of secondary hypertension were excluded when relevant. Testing for PA followed standard practice guidelines, including use of the ARR, confirmatory testing, and AVS to distinguish unilateral from bilateral PA (eAppendix 1 in the Supplement).

PA was confirmed in 304 patients included into the study according to selective sampling of both adrenal veins or in whom adrenalectomies were performed without AVS because of young age and imaging evidence of a single adrenal adenoma (eFigure 1 in the Supplement). Among these patients, 116 were defined by AVS to have bilateral disease. After exclusions among the others with unilateral disease, there remained 157 patients for whom Sanger sequencing for somatic sequence variants of KCNJ5 was performed in resected tumor specimens (eAppendix 1 in the Supplement). Two and 16 of the respective 60 and 97 patients with and without KCNJ5 sequence variants did not experience complete biochemical cure and were reassigned as having bilateral PA according to the primary aldosteronism surgical outcome (PASO) classification system. PA was excluded in 158 patients, who were subsequently defined to have PHT. A further 201 patients with hypertension from a reference population were included to enhance patient numbers for generation of machine-learning algorithms (eTable 1 in the Supplement).

Steroid Profiling
Steroid profiling was performed using liquid chromatography with tandem mass spectrometry, with details outlined in eAppendix 1 and eTable 2 in the Supplement. Measurements included 15 adrenal steroids: aldosterone, 18-oxocortisol, 18-hydroxycortisol, cortisol, cortisone, 11-deoxycortisol, 21-deoxycortisol, corticosterone, 11-deoxycorticosterone, progesterone, 17-hydroxyprogesterone, pregnenolone, androstenedione, dehydroepiandrosterone, and dehydroepiandrosterone-sulfate. Reference intervals were established as discussed elsewhere (eTable 3 in the Supplement).

Statistical Analysis
Statistical analyses used JMP Pro statistical software version 14 (SAS Institute). Unless otherwise specified, significance was defined as \( P < .05 \). Statistical tests were 2-tailed and included Fisher exact test and the Mann-Whitney U test. Nominal logistic modeling was used to assess for associations of the presence versus absence of a KCNJ5 sequence variant with PASO criteria based-outcomes according to sex and age as additional covariates. Associations are shown according to whole model and likelihood ratio tests. Data for steroids were normalized by logarithmic transformation before analyses, including for generation of geometric means and 95% CIs. Least-squares multivariable models were used to assess differences in plasma steroids according to patient group, age, sex, and
assay batch. Differences among patient groups were assessed using the Tukey honest significance test. Logistic regression was used to generate receiver operating characteristic curves, with selections of steroids in profiles based on both stepwise regression and likelihood ratios for each steroid. Differences between areas under receiver operating characteristic curves (AUROCs) and data from confusion matrices were used to assess performance of logistic regression models. Data were normalized according to upper cutoffs of reference intervals, which for most of the plasma steroids were specific for either or both age and sex (eTable 3 in the Supplement). Data analyses were performed from September 2018 to August 2019.

**Machine Learning**

In brief, the machine-learning workflow involved 3 phases (eFigure 2 in the Supplement): data preparation, model learning, and external validation. Data preparation included several procedures for normalization, batch correction, and, in some models, adjustments for age and sex (see eTable 3, eTable 4, and eTable 5 in the Supplement). At this stage, each of the 13 different data sets was subdivided into 2 different proportions for learning and external validation data sets, as outlined in eAppendix 1 in the Supplement. After data preparation, machine-learning tasks for feature selection, model training, and sample classification in the second model learning phase were performed according to different algorithms, with their application in this phase restricted to learning data sets. Feature selection involved the use of 4 different algorithms to identify specific steroid combinations that provided either optimal segregation of patients with and without PA or identification of those with unilateral disease due to KCNJ5 sequence variants among all patients.

Several combinations of the aforementioned procedures were investigated for optimized data analysis and assessed according to 9 machine-learning algorithms corresponding to variations of 4 commonly used models in medicine: random forest (RF), support vector machine (SVM), linear discriminant analysis, and logistic regression. A total of 585 models arising from 13 data sets and 9 machine-learning algorithms were tested, each involving a 10 times, 5-fold cross-validation step (eFigure 3 in the Supplement). Optimal classification, determined as part of the final validation phase according to either AUROC or F scores, was determined according to external validations achieved by application of algorithms for each of the 585 models to external validation data sets.

**Results**

**Final Study Population**

PA was confirmed in 273 patients (165 men [60%]; mean [SD] age, 51 [10] years). In addition to the 201 patients of the reference hypertension population, after screening and subtype classification, there were 158 patients classified with PHT (134 with bilateral PA and 139 with unilateral PA) (eTable 1 in the Supplement). Among those with unilateral PA, 58 had APAs with KCNJ5 sequence variants and 81 did not and were designated as having wild-type KCNJ5.

**Genotype-Related Therapeutic Outcomes and Patient Group Reclassification**

Among patients who underwent adrenalectomy because of presumed unilateral PA, those with APAs with KCNJ5 sequence variants were, on average, 5.8 years younger (mean [SD] age, 47.4 [10.8] years vs 51.3 [10.3] years) and were 2.7-fold more likely to be female (47 women [78.3%] vs 28 women [28.9%]) compared with those with wild-type KCNJ5 APAs (Table 1). According to the PASO classification, the presence of KCNJ5 sequence variants conferred significantly better clinical and biochemical outcomes after adrenalectomy compared with the absence of KCNJ5 sequence variants. However, logistic modeling indicated that improved blood pressure control in patients with variant vs wild-type KCNJ5 APAs was accounted for by the younger age and female predominance of patients with KCNJ5 sequence variants. In contrast, the presence of a KCNJ5 sequence variant remained independently associated with biochemical cure. The overall postadrenalectomy biochemical cure
rate in this study was 88.5%; the cure rates were 96.6% for patients with KCNJ5 sequence variants and 83.6% for patients without KCNJ5 sequence variants.

Steroid Profiles

With least squares adjustments of sex, age, and assay batch, all plasma steroids showed some differences among the 5 patient groups (eTable 4 in the Supplement). Plasma 18-oxocortisol showed differences among all groups but especially the group with unilateral APAs due to KCNJ5 sequence variants, in whom plasma concentrations were 6.2- to 10.3-fold higher than all other groups (Table 2). Plasma 18-hydroxycortisol in the variant KCNJ5 group was also 3.3- to 4.0-fold higher than in other groups. Plasma aldosterone in the 2 unilateral disease groups, which did not differ, were higher than in the other 3 groups. Other steroids were either similarly increased in patients with PA

---

**Table 1. Comparisons of Age, Sex, and Primary Aldosteronism Surgical Outcome Clinical and Biochemical Outcomes in Patients With Adrenal Venous Sampling–Lateralized Evidence of Unilateral Adrenal Aldosterone Secretion According to the Presence or Absence of KCNJ5 Sequence Variants in Resected Adenomas**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients, No. (%)</th>
<th>Wild-type KCNJ5 (n = 97)</th>
<th>Variant KCNJ5 (n = 60)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
<td>53.1 (10.3)</td>
<td>47.4 (10.8)</td>
<td>.002</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>28 (28.9)</td>
<td>28 (28.9)</td>
<td>47 (78.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Male</td>
<td>69 (71.1)</td>
<td>69 (21.1)</td>
<td>13 (27.7)</td>
<td></td>
</tr>
<tr>
<td>Clinical outcomes of primary aldosteronism surgery†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete cure</td>
<td>20 (20.6)</td>
<td>20 (20.6)</td>
<td>24 (40.0)</td>
<td></td>
</tr>
<tr>
<td>Partial cure</td>
<td>52 (53.6)</td>
<td>52 (53.6)</td>
<td>30 (50.0)</td>
<td>.008</td>
</tr>
<tr>
<td>Failure</td>
<td>25 (25.8)</td>
<td>25 (25.8)</td>
<td>6 (10.0)</td>
<td></td>
</tr>
<tr>
<td>Biochemical outcomes of primary aldosteronism surgery†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete cure</td>
<td>81 (83.6)</td>
<td>81 (83.6)</td>
<td>58 (96.6)</td>
<td></td>
</tr>
<tr>
<td>Partial cure</td>
<td>7 (7.2)</td>
<td>7 (7.2)</td>
<td>1 (1.7)</td>
<td>.04</td>
</tr>
<tr>
<td>Failure</td>
<td>9 (9.2)</td>
<td>9 (9.2)</td>
<td>1 (1.7)</td>
<td></td>
</tr>
</tbody>
</table>

† In the multivariate analyses for clinical outcomes, likelihood ratios were 9.34 for age impact (P = .009), 6.01 for sex impact (P = .05), and 1.42 for KCNJ5 impact (P = .49), with P < .001 for the whole model.

Table 2. Plasma Concentrations of Steroids in Reference Patients With Hypertension, Patients With Primary Hypertension, and Patients With Bilateral Primary Aldosteronism or Unilateral Primary Aldosteronism Without and With KCNJ5 Sequence Variants

<table>
<thead>
<tr>
<th>Steroid</th>
<th>Plasma concentration, least square geometric mean (95% CI), nmol/L*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hypertension</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
</tr>
<tr>
<td>Aldosterone</td>
<td>0.091 (0.077-0.106)</td>
</tr>
<tr>
<td>18-Oxocortisol</td>
<td>0.026 (0.022-0.031)</td>
</tr>
<tr>
<td>18-Hydroxycortisol</td>
<td>1.62 (1.41-1.84)</td>
</tr>
<tr>
<td>Corticosterone</td>
<td>4.28 (3.60-5.01)</td>
</tr>
<tr>
<td>11-Deoxycorticosterone</td>
<td>0.063 (0.052-0.075)</td>
</tr>
<tr>
<td>11-Deoxycortisol</td>
<td>1.332 (1.074-1.618)</td>
</tr>
<tr>
<td>21-Deoxycortisol</td>
<td>0.039 (0.030-0.050)</td>
</tr>
<tr>
<td>Cortisol</td>
<td>237 (213-262)</td>
</tr>
<tr>
<td>Cortisone</td>
<td>47.2 (43.2-51.4)</td>
</tr>
<tr>
<td>Androstenedione</td>
<td>2.47 (2.25-2.70)</td>
</tr>
<tr>
<td>Dehydroepiandrosterone</td>
<td>8.70 (7.66-9.81)</td>
</tr>
<tr>
<td>Dehydroepiandrosterone-sulfate</td>
<td>3401 (3061-3758)</td>
</tr>
<tr>
<td>17-Hydroxyprogesterone</td>
<td>1.10 (0.96-1.24)</td>
</tr>
<tr>
<td>Progesterone</td>
<td>0.336 (0.269-0.411)</td>
</tr>
<tr>
<td>Pregnenolone</td>
<td>1.95 (1.61-2.32)</td>
</tr>
</tbody>
</table>

* Geometric means and 95% CIs were derived from the exponents of logarithmically transformed data. For whole model differences, see eTable 4 in the Supplement.
or showed differing patterns or decreases or increases compared with patients with hypertension according to the particular subtype of PA.

**Diagnostic Test Performance of the ARR and Plasma Steroids**

From differences in AUROCs, a selected panel of 8 steroids was less effective than the ARR (difference in AUROC, 0.053; 95% CI, 0.006 to 0.099; \( P = .03 \)) for differentiating patients with PA from those with hypertension, but more effective (difference in AUROC, 0.107; 95% CI, 0.037 to 0.176; \( P = .003 \)) for distinguishing the unilateral variant \( \text{KCNJ5} \) APA group from others (Figure 1). Combination of the steroid profile with the ARR was nevertheless more effective for discriminating...
and RF models. For the SVM model, cortisone, 11-deoxycortisol, and androstenedione replaced 3 places, with 11-deoxycorticosterone following in fourth and fifth places, respectively, for the SVM (Figure 2). For both models, aldosterone, 18-oxocortisol, and 18-hydroxycortisol occupied the top 3 steroids within machine-learning approaches, combinatorial markers composed of up to 7 steroids were identified that offered best performance for discriminating patient groups (eFigure 5, eFigure 6, eFigure 7, eFigure 8, and eFigure 10 in the Supplement). Among those steroids, aldosterone, 18-oxocortisol, and 18-hydroxycortisol commonly occupied the top 3 places for discriminatory power. The next steroid with useful discriminatory power was 11-deoxycorticosterone, followed by several others depending on the model.

From confusion matrices, steroid profiles correctly identified nearly 3 times more patients with variant KCNJ5 unilateral disease than the ARR (eTable 6 in the Supplement). With the addition of the steroid profile to the ARR, the diagnostic yield of patients correctly identified with PA increased from 69.6% (95% CI, 64.2%-75.0%) to 81.3% (95% CI, 76.7%-85.9%) at respective diagnostic specificities of 89.2% (95% CI, 84.3%-94.1%) and 89.9% (95% CI, 85.2%-94.6%).

**Steroid Profiling With Machine Learning**

After batch corrections (eFigure 4 and eFigure 5 in the Supplement) and using feature selection within machine-learning approaches, combinatorial markers composed of up to 7 steroids were identified that offered best performance for discriminating patient groups (eFigure 5, eFigure 6, eFigure 7, eFigure 8, and eFigure 9, and eFigure 10 in the Supplement). Among those steroids, aldosterone, 18-oxocortisol, and 18-hydroxycortisol commonly occupied the top 3 places for discriminatory power. The next steroid with useful discriminatory power was 11-deoxycorticosterone, followed by several others depending on the model.

The final selection of models for optimal classification was reduced to 21 best models according to either AUROCs or F scores (eTable 7 in the Supplement). Among these, an RF model provided optimal performance for the classification of patients with and without PA, whereas a nonlinear (radial basis function kernel) SVM model was optimal for patients with unilateral variant KCNJ5 APAs (Figure 2). For both models, aldosterone, 18-oxocortisol, and 18-hydroxycortisol occupied the top 3 places, with 11-deoxycorticosterone following in fourth and fifth places, respectively, for the SVM and RF models. For the SVM model, cortisol, 11-deoxycortisol, and androstenedione replaced

---

**Figure 2. Results for the 2 Best Machine-Learning Models**

Panel A shows the subtree from 1 decision tree of 500 primary aldosteronism vs other groups. Panels B and D show a support vector machine (SVM) with a nonlinear kernel model for the differentiation of unilateral KCNJ5 APAs from primary aldosteronism vs other groups. Panel A shows the subtree from 1 decision tree of 500 new samples. Panel B outlines the mathematical formula used in SVM to predict new sample scores, where $k$ is the number of binary SVM models created for the 1 vs 1 approach for multiclass SVM training, $x$ is the new sample to be predicted, $n$ is the number of support vectors for the $k$th binary SVM, $a$ and $b$ are the parameters learned from the training step of the $k$th binary SVM, $y$ is the class of the respective $k$th support vector (1 or −1), $G(x, X)$ is the dot product between the $j$th support vector hyperplane measures in the binary $k$th SVM $y$ with the (new) sample measurements $x$. Correlation networks from the respective selected features for each model are shown in panels C and D, with nodes in brown showing common features. DHEA indicates dehydroepiandrosterone.
corticosterone, 17-hydroxyprogesterone, and dehydroepiandrosterone as selected features of the RF model.

Performance of RF and SVM models upon external validation was similar or even appeared to exceed that of the learning series (Table 3), according to 10 cross-validations in 5-folds (eAppendix 2 and eFigure 2 in the Supplement). Comparisons of AUROCs and F scores indicated that the SVM model performed nearly as well as the RF model for identifying patients with PA, but was consistently better for identifying those with variant KCNJ5 APAs. Using the RF model to identify patients with PA, among all 632 patients, diagnostic sensitivity in the learning series was 69% (95% CI, 68%-71%) and specificity was 94% (95% CI, 93%-94%). The external validation series yielded sensitivity of 85% and specificity of 100%. To identify patients with variant KCNJ5 APAs, the SVM model yielded a diagnostic sensitivity of 85% (95% CI, 81%-88%) at a specificity of 97% (95% CI, 97%-98%). The external validation series yielded respective values of sensitivity of 100% and specificity of 98%.

As outlined in eAppendix 2 in the Supplement, the aforementioned measures of diagnostic performance were derived using learning ratios of 90% optimal for training and testing (thus, 10% Table 3. Confusion Matrices and Diagnostic Performance for the 2 Machine-Learning Models (RF-Gini and SVMnl-RFE) for the Learning (Training and Testing) and External Validation Series of Patients With PHT, B-PA, and Unilateral Primary Aldosteronism With and Without KCNJ5 Sequence Variants

<table>
<thead>
<tr>
<th>Actual groups</th>
<th>Predicted groups</th>
<th>Diagnostic performance^c</th>
<th>External validation^b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Learning^a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHT</td>
<td>B-PA</td>
<td>Wild-type KCNJ5</td>
<td>Variant KCNJ5</td>
</tr>
<tr>
<td>RF-Gini</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion matrices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHT</td>
<td>60.4</td>
<td>2.4</td>
<td>1.2</td>
</tr>
<tr>
<td>B-PA</td>
<td>9</td>
<td>11</td>
<td>2.9</td>
</tr>
<tr>
<td>Wild-type KCNJ5</td>
<td>4.1</td>
<td>5.4</td>
<td>3.4</td>
</tr>
<tr>
<td>Variant KCNJ5</td>
<td>2</td>
<td>1.7</td>
<td>2</td>
</tr>
<tr>
<td>Diagnostic performance^c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>94 (93-94)</td>
<td>42 (40-45)</td>
<td>26 (24-29)</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>69 (68-71)</td>
<td>89 (89-90)</td>
<td>94 (94-95)</td>
</tr>
<tr>
<td>AUROC</td>
<td>0.815 (0.807-0.825)</td>
<td>0.657 (0.645-0.670)</td>
<td>0.599 (0.585-0.613)</td>
</tr>
<tr>
<td>PPV, %</td>
<td>80 (79-81)</td>
<td>52 (50-55)</td>
<td>38 (32-40)</td>
</tr>
<tr>
<td>NPV, %</td>
<td>89 (88-90)</td>
<td>86 (85-87)</td>
<td>90 (89-91)</td>
</tr>
<tr>
<td>F score</td>
<td>0.863 (0.858-0.870)</td>
<td>0.464 (0.440-0.480)</td>
<td>0.309 (0.279-0.336)</td>
</tr>
<tr>
<td>SVMnl-RFE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion matrices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHT</td>
<td>63.2</td>
<td>1.2</td>
<td>0.2</td>
</tr>
<tr>
<td>B-PA</td>
<td>10.6</td>
<td>10.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Wild-type KCNJ5</td>
<td>2.8</td>
<td>3.6</td>
<td>6.6</td>
</tr>
<tr>
<td>Variant KCNJ5</td>
<td>1.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Diagnostic performance^c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>98 (97-98)</td>
<td>45 (43-47)</td>
<td>45 (42-49)</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>70 (69-73)</td>
<td>94 (94-95)</td>
<td>98 (88-98)</td>
</tr>
<tr>
<td>AUROC</td>
<td>0.841 (0.831-0.850)</td>
<td>0.695 (0.684-0.706)</td>
<td>0.716 (0.700-0.736)</td>
</tr>
<tr>
<td>PPV, %</td>
<td>81 (80-83)</td>
<td>69 (67-71)</td>
<td>78 (74-81)</td>
</tr>
<tr>
<td>NPV, %</td>
<td>95 (94-95)</td>
<td>86 (85-87)</td>
<td>92 (91-93)</td>
</tr>
<tr>
<td>F score</td>
<td>0.888 (0.883-0.895)</td>
<td>0.537 (0.516-0.556)</td>
<td>0.562 (0.528-0.593)</td>
</tr>
</tbody>
</table>

Abbreviations: AUROC, area under the receiver operating characteristic curve; B-PA, bilateral primary aldosteronism; NPV, negative predictive value; PHT, primary hypertension; PPV, positive predictive value.

^a For learning series, numbers in confusion matrices reflect 5-folds of patients (ie, 569/5 = 114 patient for each fold) with evaluations of each fold performed 10 times within each learning series (thus, numbers represent the mean of 50 confusion matrices).

^b For the external validation series numbers reflect the learning proportions (90:10) and 10% (63) of the total number of patients (632) in the analysis.

^c Values for diagnostic performance in learning series are shown with 95% CI, whereas those for validation series are not.

for external validation), which was particularly important for the limited population of 58 patients with variant KCNJ5 APAs. Thus, measures of diagnostic performance for the SVM model of both learning and external validation series, but particularly the latter, showed improvement as the learning ratio increased from 50% to 90% (eFigure 11 in the Supplement). For the RF model, for which population sizes of PA and hypertensive groups were both relatively large, measures of diagnostic performance showed little difference between learning and validation series until the learning ratio reached 90% (eFigure 12 in the Supplement).

Discussion

To our knowledge, this study is the first to demonstrate the application of multidimensional pattern recognition and machine learning for analysis of steroidomic data in the diagnosis of PA. This approach offers the potential for more efficient and effective diagnostic stratification than the traditional series of multiple studies involving single end-point measures in relation to given cutoff values. Stratification was achieved by distinct steroid profiles among subgroups of patients with PA. From those profiles, panels of steroids were identified that can facilitate diagnosis of PA and, in the same screening step, identify patients with variant KCNJ5 APAs for triaging as candidates likely to show beneficial therapeutic outcomes from further interventions.

Among the steroids in the panel with distinctive profiles, 18-oxocortisol and 18-hydroxycortisol stood out from the others for identifying patients with unilateral APAs due to somatic KCNJ5 sequence variants and were consistently among the top 4 steroids selected by machine-learning algorithms. Previous studies have identified those hybrid steroids to be produced in excess in some patients with PA, but only recently has it been clarified that elevations of these steroids are linked to APAs with sequence variants of KCNJ5. The 2 hybrid steroids appear to be formed by actions of aldosterone synthase on 11-deoxycortisol, which is normally produced in the zona fasciculata and converted there to cortisol by 11β-hydroxylase. Production of the hybrid steroids by variant KCNJ5 APAs is explained by their zona fasciculata phenotype and their expression of both CYP11B1 and CYP11B2.

The benefits of AVS over imaging to establish unilateral PA are well established. Nevertheless, the originally suggested high diagnostic accuracy of AVS for indicating unilateral disease has not been supported by some subsequent studies involving postadrenalectomy follow-up. In the present study, the 88.5% postadrenalectomy biochemical cure rate lies between those found previously and is similar to that found in a single prospective study. The failure of adrenalectomy to cure PA may reflect asymmetric bilateral disease in some patients. Aldosterone-producing cell clusters have been identified in the zona glomerulosa of aging adrenal glands and in the adrenal glands of patients with PA due to bilateral adrenal aldosterone hypersecretion; in both cases, cells of those clusters are characterized by high rates of sequence variants of CACNAID, but not KCNJ5. This raises the possibility that KCNJ5 sequence variants might be characteristic of unilateral adenomas. Nevertheless, 2 of our patients with variant KCNJ5 adenomas did not experience complete biochemical cure after adrenalectomy, suggesting that KCNJ5 sequence variants are not strictly associated with unilateral disease. Nevertheless, failure to reach cure in the KCNJ5 sequence variant group was rare, confirming findings that patients with variant KCNJ5 adenomas show more clinical benefit after adrenalectomy than others. As we further establish here, the benefit in terms of biochemical cure is independent of age and sex, further highlighting the importance of triaging patients with variant KCNJ5 APAs for further interventions.

There have been other studies that combined steroid profiling with machine learning, but to our knowledge, this is the first to apply a combinatorial marker design strategy to PA. The potential benefits for diagnostic stratification of PA are multiple. First, during screening it may be possible to more effectively distinguish patients with PA from those with other causes of hypertension. Second, by identifying within the same screening step patients with unilateral APAs due to KCNJ5 sequence variants, it should be possible to immediately triage those patients for AVS; alternatively, with clear
imaging evidence of a unilateral adenoma, it may be possible to directly proceed to an adrenalectomy without AVS. These considerations underscore the potential advantages of moving away from traditional unidimensional approaches (e.g., ARR) for diagnostic stratification to multidimensional approaches that take advantage of today’s computational power for applications of artificial intelligence.

**Limitations**

As detailed in eAppendix 3 in the *Supplement*, the present analysis has limitations that are typical of retrospective diagnostic studies, including dependence on traditional methods to establish patient classifications. It thus cannot be guaranteed that PA was excluded in all patients designated as having primary hypertension or that some cases of bilateral PA were incorrectly classified. Although patient follow-up ensured that final cases of unilateral PA were correctly classified, reliance on Sanger sequencing for identifying *KCNJ5* sequence variants is not 100% sensitive; it is, thus, possible that some wild-type *KCNJ5* cases may have been incorrectly classified. Measurements of the ARR at sampling time points different from those used for steroid profiling, batch effects, and inconsistencies in supine and seated blood sampling represent other limitations. Seated sampling, which increases plasma renin and aldosterone, likely accounts for the higher levels of aldosterone and 18-oxocortisol in patients with PHT who were screened for PA compared with those of the reference hypertensive population, for whom sampling was performed with patients in the supine position. As outlined in the eAppendix 3 in the *Supplement*, because sampling for steroid profiles among patients with PA was mainly performed with patients in the supine position, this may have adversely impacted the performance of steroid profiles for distinguishing patients with PA from those with PHT.

**Conclusions**

These findings suggest that plasma steroid profiles obtained during initial screening for PA can improve case detection beyond that possible using the ARR alone. Moreover, the use of distinctive profiles to identify patients with unilateral variant *KCNJ5* APAs further illustrates the potential of steroid profiling for disease stratification at a single screening step. Along with advances in functional imaging and other measurements, such as the angiotensin peptidome, steroid profiling combined with machine learning may facilitate more rapid identification of patients with PA for appropriate therapeutic interventions. As detailed in eAppendix 3 in the *Supplement*, such strategies are now being tested in further patient populations, and with those developments it may become possible to screen more than the small proportion of patients with PA who are currently tested and treated according to disease subtype.
REFERENCES
Steroid Profiling and Machine Learning for Diagnosing Primary Aldosteronism


SUPPLEMENT.
eAppendix 1. Supplemental Methods
eFigure 1. SPISCA Study Patient Flow Diagram
eTable 1. Demographic Data for the Five Final Groups of the Study Population
eTable 2. Interassay Coefficients of Variation (CV) for Three Different Quality Control Materials Containing the Steroids at Differing Plasma Concentrations (ng/mL)
eTable 3. Upper Cutoffs of Reference Intervals for Plasma Steroids
eFigure 2. Workflow for the Analysis of the Plasma Steroidomic Data by Machine Learning (ML) According to Three Steps
eTable 4. Multivariate Analysis Using a Model That Takes Into Account Sex, Age and Assay Batch to Establish Differences Among the Five Patient Groups: 1. Reference Hypertensives; 2. Primary Hypertensives; 3. Bilateral PA; 4. Unilateral PA Without KCNJ5 Mutations; and 5 Unilateral PA With KCNJ5 Mutations
eTable 5. Normalizations and Methods for the S4 Selected Top Machine Learning Models
eAppendix 2. Supplemental Results
eTable 6. Confusion Matrices and Diagnostic Performance From Logistic Regression Analyses of Aldosterone-Renin Ratios (ARR), Steroid Profiles and Combined ARR and Steroid Profiles for Patients With Primary Hypertension (PHT) or Primary Aldosteronism (PA) According to Bilateral Disease or Unilateral KCNJ5 Wildtype (KCNJ5$^{WT}$) or Mutation-Positive (KCNJ5$^{MUT}$) Disease
eFigure 3. AUC for the Assessment of Batch Influence and Correction for Each Steroid
eFigure 4. Outputs Forest Plots by the Tool exploBATCH Quantifying the Batch Effects
eFigure 5. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series)for the First Set of Nine ML Algorithms According to Criteria 1 and Evaluations by AUC
eFigure 6. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series) for the Second Set of Nine ML Algorithms According to Criteria 2 and Evaluations by AUC
eFigure 7. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series) for the Third Set of Nine ML Algorithms According to Criteria 3 and Evaluations by AUC
eFigure 8. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series) for the Fourth Set of Nine ML Algorithms According to Criteria 3 and Evaluations by F-Score
neFigure 9. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series) for the Fifth Set of Nine ML Algorithms According to Criteria 2 and Evaluations by F-Score
eFigure 10. Listings of Models With Selected Features and Bar Graphs of Diagnostic Performance (External Validation Series) for the Sixth Set of Nine ML Algorithms According to Criteria 2 and Evaluations by F-Score
eTable 7. Final 21 Unique Top Performing ML Models
eFigure 11. Measures of Diagnostic Performance for Identification of Patients With KCNJ5 Mutation+ve APAs Using the SVMnl-RFE Model
eAppendix 3. Supplemental Discussion
eFigure 12. Measures of Diagnostic Performance for Identification of Patients With Primary Hypertension Using the RF-Gini Model
eReferences