The Journal of Clinical Endocrinology & Metabolism, 2022, **XX**, 1–9
https://doi.org/10.1210/clinem/dgac271
Advance access publication 05 May 2022
Clinical Research Article

A Multiclassifier System to Identify and Subtype Congenital Adrenal Hyperplasia Based on Circulating Steroid Hormones

Lei Ye,[1,2,*] Zhiyun Zhao,[2,*] Huixia Ren,[3,*] Wencui Wang,[1,*] Wenzhong Zhou,[2,*] Sichang Zheng,[1] Rulai Han,[2] Jie Zhang,[1] Haorong Li,[1] Zhihan Wan,[1] Chao Tang,[3] Shouyue Sun,[1,] Weiqing Wang,[1,2,] and Guang Ning[1,2]

[1]Shanghai Clinical Center for Endocrine and Metabolic Diseases, Department of Endocrinology, Ruijin Hospital, Shanghai Jiaotong University, School of Medicine, Shanghai, China
[2]Shanghai Key Laboratory for Endocrine Tumors, Shanghai Institute of Endocrine and Metabolic Diseases, Shanghai, China
[3]Center for quantitative biology, Peking University, Beijing, China

*These authors contribute equally to this work.

Correspondence: Shouyue Sun, MD, Shanghai Clinical Center for Endocrine and Metabolic Diseases, Department of Endocrinology, Ruijin Hospital, Shanghai Jiaotong University, School of Medicine, 197 Ruijin 2nd Rd, Shanghai 200025, P.R. China. Email: ssy10926@rjh.com.cn; or Weiqing Wang, MD, Shanghai Clinical Center for Endocrine and Metabolic Diseases, Department of Endocrinology, Ruijin Hospital, Shanghai Jiaotong University, School of Medicine, 197 Ruijin 2nd Rd, Shanghai 200025, P.R. China. Email: wqingw61@163.com.

Abstract

Context: Measurement of plasma steroids is necessary for diagnosis of congenital adrenal hyperplasia (CAH). We sought to establish an efficient strategy for detection and subtyping of CAH with a machine-learning algorithm.

Methods: Clinical phenotype and genetic testing were used to provide CAH diagnosis and subtype. We profiled 13 major steroid hormones by liquid chromatography-tandem mass spectrometry. A multiclassifier system was established to distinguish 11β-hydroxylase deficiency (11βOHD), 17α-hydroxylase/17,20-lyase deficiency (17OHD), and 21α-hydroxylase deficiency (21OHD) in a discovery cohort (n = 226). It was then validated in an independent cohort (n = 111) and finally applied in a perspective cohort of 256 patients. The diagnostic performance on the basis of area under receiver operating characteristic curves (AUCs) was evaluated.

Results: A cascade logistic regression model, we named the "Steroidogenesis Score," was able to discriminate the 3 most common CAH subtypes: 11βOHD, 17OHD, and 21OHD. In the perspective application cohort, the steroidogenesis score had a high diagnostic accuracy for all 3 subtypes, 11βOHD (AUC, 0.994; 95% CI, 0.983-1.000), 17OHD (AUC, 0.993; 95% CI, 0.985-1.000), and 21OHD (AUC, 0.979; 95% CI, 0.964-0.994). For nonclassic 21OHD patients, the tool presented with significantly higher sensitivity compared with measurement of basal 17α-hydroxyprogesterone (17OHP) (0.973 vs 0.840, $P = 0.005$) and was not inferior to measurement of basal vs stimulated 17OHP (0.973 vs 0.947, $P = 0.681$).

Conclusions: The steroidogenesis score was biochemically interpretable and showed high accuracy in identifying CAH patients, especially for nonclassic 21OHD patients, thus offering a standardized approach to diagnose and subtype CAH.

Key Words: congenital adrenal hyperplasia, diagnosis, subtype, liquid chromatography-tandem mass spectrometry

Steroidogenesis is a both complex and subtle process that plays a key role in numerous cellular/physiological functions (1, 2). Adrenal steroids include glucocorticoids, mineralocorticoids, and androgens. A total of 14 enzymes or cofactors are required for adrenal steroidogenesis, producing 16 major steroid metabolites (3). Genetic deficiencies in steroidogenesis enzymes impair glucocorticoid biosynthesis, causing a group of diseases called congenital adrenal hyperplasia (CAH), which includes 9 subtypes (4). CAH is one of the most common autosomal recessive disorders. Classic CAH occurs in 1 in 10 000 to 1 in 20 000 live births while the estimated prevalence of nonclassic CAH ranges from 1 case per 200 persons to 1 case per 1000 persons (5-9).

The presentation of CAH is extremely variable, complicated, and sometimes occult, predominantly owing to the complexity of the steroidogenesis pathway and the numerous disorder subtypes (10, 11). Each steroidogenesis enzyme or cofactor catalyzes multiple reactions. Deficiencies in any step block downstream steroidogenesis, while increasing adrenocorticotropic hormone (ACTH) production, which stimulates the detour pathways. Patients with CAH, therefore, manifest with simultaneous multiple hormonal imbalances of glucocorticoids, mineralocorticoids, or sex steroids (4). Individual subtypes may show overlapping presentations. Moreover, as mutant enzyme activity can be either mildly affected or completely inactivated, a continuum of disease phenotypes have been reported, from potentially life-threatening to infertility, and in some cases, individuals are completely asymptomatic (12-14). For example, the clinical phenotype of female nonclassic patients is similar to polycystic ovarian syndrome,

Received: 29 November 2021. **Editorial Decision:** 26 April 2022. **Corrected and Typeset:** 27 May 2022

the most common reason for female infertility. Given the differential treatment approaches, accurate diagnosis is essential (15-18).

Current diagnosis of CAH often includes a tedious and uncertain workup. Take the most commonly occurring form of CAH, 21α-hydroxylase deficiency (21OHD) as example: although random measurement of 17α-hydroxyprogesterone (17OHP) may be informative, an ACTH stimulation test is usually required to confirm the diagnosis, especially for nonclassic 21OHD (6, 12). For other forms of CAH, such as 17α-hydroxylase/17,20-lyase deficiency (17OHD) and 11β-hydroxylase deficiency (11βOHD), diagnostic criteria are lacking, even when genetic data are available, and no guidelines or consensus statements have been published. All these challenges result in misdiagnosis or delayed diagnosis of CAH. The mortality associated with undiagnosed salt-wasting CAH is between 4% and 10% (19). The rates of spontaneous pregnancy and fertility are extremely low (0-10%) among women with salt-wasting CAH and moderately low (33-60%) in women with the simple-virilizing type (20, 21). Early diagnosis, therefore, improves not only survival but also enhances development of and potential for fertility.

Profiling of multiple steroids by liquid chromatography-tandem mass spectrometry (LC-MS/MS) in patients with steroid-hormone disorders has begun to reveal distinctive steroid patterns associated with these disorders (22-24). In 2017, 2 groups developed assays based on comparison to multiple steroid hormones that accurately detected heterozygote *Cyp21a2* mutation carriers (25, 26). More recently, urinary gas chromatography-mass spectrometry steroid metabotyping has been used to monitor treatment outcomes in children with CAH (27). Most recently, principal component analysis of 5 key 21OHD perturbed steroids was reported to improve the positive predictive value when screening 21OHD in neonates (28). However, for the most part data, interpretation remains largely based on individual steroid hormones or the individual ratio of the precursor and product of a specific enzyme (29). Interpretation of multiple steroid data based upon a data-learning process has been lacking.

The data-learning process has been successfully applied in disease classification and the management of commonly occurring diseases (30-32). In the present study, we sought to establish a multiclassifier system to diagnose CAH, a hereditary disease, through measurement of a panel of 13 steroid hormone levels in a single baseline blood draw.

Materials and Methods

Study Design and Participants

The study was approved by Ruijin Hospital Ethics Committee. All enrolled participants or their legal guardians provided informed consent and were patients seen at Ruijin Hospital. To establish a classification system and validate its diagnostic performance in CAH, we conducted a 3-phase, single-center study according to Standards for Reporting of Diagnostic Accuracy guidelines (33). In the first phase, 140 patients (110 CAH cases and 30 case-like patients) were enrolled between July 4, 2007 and March 30, 2016 to serve as a discovery cohort (Fig. 1A). An independent cohort of 69 patients (54 CAH cases and 15 case-like patients) were enrolled between June 22, 2007 and March 23, 2016 and served as the validation cohort (Fig. 1B). The clinical diagnosis of 21OHD was based on the 2018 Endocrine Society Clinical Practice Guideline (6). The clinical diagnoses of 11βOHD, 17OHD, and others were based on the definitions provided by the *Williams Textbook of Endocrinology, 14th Edition* (3). The recommended hormone measurements and genetic analysis are described in the next section. All CAH patients were confirmed to carry biallelic pathogenic mutations of *Cyp21a2*, *Cyp17a1*, *Cyp11b1*, or other CAH-associated genes. A diagnosis of CAH was ultimately excluded in case-like patients based on complete clinical and molecular data. A total of 146 sex- and age-matched (±5 years) healthy controls were included (Table 1). The prospective cohort included 271 subjects enrolled from May 4, 2016 to December 25, 2020 (Table 2), with at least 1 of the following features: (1) atypical genitalia and abnormal growth; (2) hyperandrogenism as irregular menses, hirsutism, or acne; (3) hypertension and/or hypokalemia; (4) adrenal insufficiency; (5) newly identified adrenal masses or hyperplasia during the evaluation of nonadrenal disease; (6) infertility; or (7) other suspicious findings including increased 17OHP, pedigree analysis, etc.

Sample Collection and Genetic Testing

Blood samples were obtained at 800 am (female patients during the menstruation cycle days 3-5). Basal levels of plasma-free testosterone, 17OHP, androstenedione, aldosterone, and plasma renin activity were measured by radioimmunoassay (Beckman Coulter Corp). Plasma and urinary-free cortisol levels were assessed using Access Immunoassay Systems (Beckman Coulter Inc., Fullerton, CA, USA); ACTH levels, by ELSA-ACTH immunoradiometric assay (Cisbio Bioassays, France); and other sex hormones including dehydroepiandrosterone sulfate, sex hormone-binding globulin, luteinizing hormone, follicle-stimulating hormone, estradiol, progesterone and testosterone, by chemiluminescence assay (Abbott Laboratories, Abbott Park, IL, USA). Stimulated 17OHP were measured before and 60 min after the intramuscular administration of 0.250 mg of synthetic 1-24 ACTH (Synacthen; Ciba SA, Gron, France). Genetic testing of all patients was conducted with a targeted gene panel for CAH, which included 9 genes (*Cyp21a2*, *Cyp17a1*, *Cyp11b1*, *Hsd3b2*, *Cyp11a1*, *Por*, *Star*, *Cyp11b2*, and *H6pd*) (34). Variant nucleotides were interpreted according to American College of Medical Genetics criteria, and all variants considered pathogenic or likely pathogenic were validated by Sanger sequencing (35). Multiplex ligation-dependent probe amplification was used to detect *Cyp21a2* large structure variations (P050-C1 CAH assay kit MRC Holland, The Netherlands).

Plasma Steroid Measurement by LC-MS/MS

Plasma concentration of 13 circulating steroid hormones were measured using LC-MS/MS including pregnenolone, 17-hydropregnenolone, 17-hydroprogesterone, 11-deoxycortisol, cortisol, cortisone, progesterone, 11-deoxycorticosterone, corticosterone, aldosterone, androstenedione, testosterone, and estradiol. Chromatography was performed using a Phenomenex Kinetex C18 2.1*100 mm, 2.6 μm, 100 Å column. All reference materials and internal standards were obtained from Dr. Ehrenstorfer GmbH. Double charcoal-stripped human serum (Beckman Coulter) and Liquicheck Immunoassay Plus Control (BIO-RAD) were used as quality controls. LC-MS/MS was performed using a

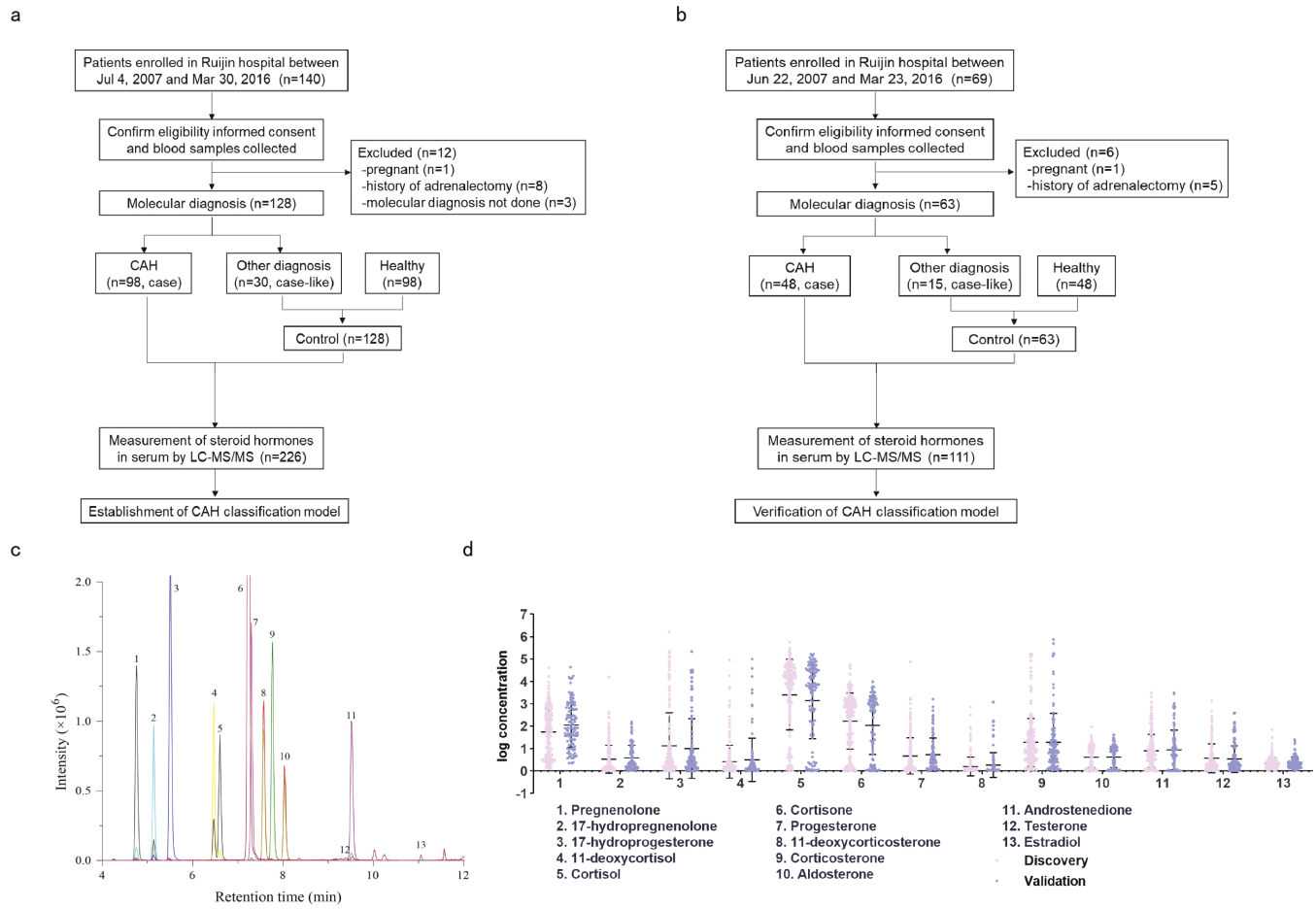

Figure 1. Discovery and validation cohorts. (A and B) Flowchart of participants enrolled in discovery and validation cohorts. (C) The retention time of 13 steroid hormones. (D) The concentration of 13 steroid hormones in discovery and validation cohorts. Data are presented as mean ± SD. Abbreviations: CAH, congenital adrenal hyperplasia; LC-MS/MS, liquid chromatography-tandem mass spectrometry.

500-µL mixture of calibration standards, patient plasma and quality control with methanol, internal quality control and water. The mixture was transferred to Cleanert solid-liquid extraction 96 well plate (500µL/3 mL HC5003Q-9DW) and finally washed with ethyl acetate:n-hexane (1:1, v/v). The extracts were collected, evaporated with nitrogen, reconstituted in methanol, and loaded onto the automatic sampler of the ultra-performance liquid chromatographic unit at a flow rate of 0.5 µL/minute. The mobile phases were (1) dH_2O and (2) methanol, each containing 2 mM ammonium formate. The injection model included 20 µL with full-loop injection. MS/MS was performed using Analyst (V1.6.1, ABSciex) in the positive mode with electrospray ionization. Data were collected using multiple reaction monitoring. To assess the laboratory-developed LC-MS/MS method, the following parameters were validated: the linearity, recovery, and limit of quantity (Land precision). For all steroid metabolites, the absolute extraction recovery rate was 40% to 60%. The relative extraction recovery rates, evaluated after internal standard correction, were between 87.2% and 106.5%. For peak values, the relative extraction recovery rates were 95.8% to 112%. We established an upper limit of quantitation for each steroid to allow for measurements at high pathological concentrations, while the lower limit of quantitation was sufficiently low to allow for the quantification of most steroids in healthy individuals. The intra-assay precision rate was 3% to 8%, and interassay precision rate was <10% for all steroids.

Model Establishment

For the purpose of CAH diagnosis, a CAH classification model was developed. We tested our data set with 8 different multiclassification models, including linear logistic regression, linear support vector classifier with L1/L2 penalty, support vector classifier with radial basis function kernel, decision-making tree, naive Bayes, random forest, and gradient-boosting classifier (36). We finally applied linear logistic regression model because it showed the highest prediction accuracy. We used the natural logarithm of the 13 metabolites concentration instead of using the original concentration. This turned out to be a critical step for the classification performance as it enhanced the magnitude of differences. Under log transformation, the subtraction between coefficients was equivalent to the ratio between the original concentrations. All the samples were first labeled as 11βOHD and non-11βOHD samples. The 11βOHD score was acquired by training a linear logistic regression model. Similarly, the remaining non-11βOHD samples were further labeled as 17OHD and non-17OHD patients. We obtained the 17OHD score by training a second linear logistic regression model. Finally, the non-11βOHD and non-17OHD samples were labeled as 21OHD and controls. The 21OHD score was defined on the third logistic regression model. Altogether, we constructed a 3D diagnosis space based on the cascade logistic regression model. The healthy control located around the original point, with the 11βOHD patients, 17OHD

patients, and 21OHD patients distributed along x, y, and z axis.

Statistical Analysis

Statistical analyses were performed using SPSS (version 22.0), PASS (version 11.0), GraphPad Prism (version 8.0), and Python (version 3.9.5). Continuous variables were presented as median [interquartile range (IQR)], and categorial variables are presented as frequency (%). Performance metrics (ie, sensitivity, specificity, positive predictive value, and negative predictive value) were calculated according to the efficient-score method described by Newcombe (37). Chi-square tests and Fisher's exact probability were used to calculate the *P*-value in comparison of the diagnosis performance between distinct assays. Mann-Whitney U test was used for continuous variable comparisons between 2 groups. The receiver operating characteristic curve was used to predict the cutoffs. *P*-value < 0.05 was considered statistically significant. We hypothesized that the steroidogenesis score can effectively distinguish CAH patients from others, and the area under the receiver operating characteristic curve (AUC) of the index was >0.8. Based on the preliminary results that AUC in discovery and validation cohort was 0.9 at $\alpha = 0.05$ (1-sided), $\beta = 0.1$, and the ratio between groups was 1:1, the sample size of the prospective cohort was estimated. At least 83 CAH patients and 83 patients with another diagnosis needed to be enrolled. Since there is no follow-up in this study, the loss to follow-up rate is not considered.

Results

Establishment of Steroidogenesis Score to Diagnose and Subtype CAH

We profiled 13 major circulating steroid hormones by LC-MS/MS [Supplementary Table 1 (38)]. The retention time was linear over certain orders of magnitude for all steroids (Fig. 1C). We then measured the 13 steroid hormones in discovery cohort. As in Figure 1A, 128 patients clinically and genetically assessed for CAH were included in the data analysis. The cohort was determined to include 66 cases with 21OHD, 22 with 17OHD, 10 with 11βOHD, and 30 case-like patients. The median age of CAH cases was 22 (IQR 18-29) years. Common manifestations included atypical genitalia and abnormal growth (55.1%), followed by irregular menses, hirsutism, and acne (25.5%); infertility (8.2%); and adrenal hyperplasia or incidentaloma (4.1%) during physical workup (Table 1). There was no significant difference in the 12 steroid hormone levels between discovery and validation cohort except pregnenolone ($P = 0.018$) (Fig. 1D).

A CAH classification algorithm was developed based on a cascade logistic regression model, which we termed the "steroidogenesis score" to obtain a multiclassifier system for the discrimination of 11βOHD, 17OHD, 21OHD, and controls (Fig. 2A and 2B). The 4 groups were classified by the combinatory cutoff of 3 scores: 11βOHD score, 17OHD score, and 21OHD score. Using the steroidogenesis score classifier, 8 out of 10 11βOHD patients were correctly classified with the sensitivity as 0.800 and specificity as 1.000 (11βOHD score higher than the cutoff); 22 out of 22 patients with 17OHD were correctly clustered by 11βOHD and 17OHD score with the sensitivity of 1.000 and the specificity of 1.000 (lower 11βOHD score and higher 17OHD score than the cutoffs); and 62 out of 66 patients with 21OHD were distinguished from the controls with the sensitivity of 0.939 and specificity of 1.000 (lower 11βOHD, 17OHD score and higher 21OHD score than the cutoffs). The 30 case-like patients and 98 healthy subjects were all correctly clustered as controls (lower 11βOHD, 17OHD, and 21OHD scores than the cutoffs). The diagnostic accuracy based on the steroidogenesis score was 0.991, 1.000, and 0.982 for 11βOHD, 17OHD, and 21OHD, respectively.

Specific steroids contributed greatly to how the steroidogenesis score was able to differentiate between CAH subtypes. As shown in Figure 2C, 11-deoxycortisol and corticosterone contributed mostly to the steroidogenesis score, indicative of 11βOHD, while corticosterone, cortisol, and androstenedione had the most impact on the score for 17OHD. Lastly, 17OHP and corticosterone were the biggest drivers of the score for 21OHD. The previous analysis found that the derived steroidogenesis score for each patient sample depended on the levels of multiple hormones but that they were not independent of each other. For example, corticosterone contributed significantly to the scores of all 3 subtypes of CAH patients, thus illustrating the necessity of multiple hormone measurement to allow classifier-based diagnosis.

Verification of Steroidogenesis Score

To verify the effectiveness of the steroidogenesis score, we performed steroid hormone LC-MS/MS and classifier analysis in an independent validation cohort. A total of 48 CAH cases, 15 case-like patients, and 48 healthy subjects were included (Fig. 1B, Table 1). The steroidogenesis score showed consistent diagnostic performance (Table 3): for 11βOHD and 17OHD, the accuracy was 1.000 and 0.991, respectively, and for 21OHD, the sensitivity was 0.909 (95% CI 0.745-0.976) and specificity was 0.949 (95% CI 0.867-0.983).

Application of Steroidogenesis Score in Prospective Cohort

The prospective cohort were recruited between May 4, 2016 and December 25, 2020 (Fig. 3A). A total of 256 patients underwent classification using the steroidogenesis score. The median age was 24 (IQR 18-31) years, and 69.9% were females. Patients aged 15 to 30 years accounted for 54.7% (140/256). The most common referral reasons were atypical genitalia and abnormal growth (42.6%), followed by hyperandrogenism classified by irregular menses, hirsutism, and acne (31.3%) and hypertension and/or hypokalemia (14.1%) (Table 2).

The steroidogenesis score classified 162 patients as CAH (11βOHD = 11, 17OHD = 19, and 21OHD = 132) and 94 patients excluded as CAH (Fig. 3A). Molecular diagnosis found biallelic pathogenic mutations of *Cyp17a1* in 23 patients, *Cyp11b1* in 10 patients, and *Cyp21a2* in 126 patients. Therefore, within the CAH-positive group, using genetic analysis as a diagnostic standard, 95.0% patients were correctly classified by the steroidogenesis score. Overall, as in Figure 3B and Supplementary Table 3 (38), the steroidogenesis score presented with sensitivity of 0.900 (95% CI 0.541-0.995) for 11βOHD, 0.826 (95% CI 0.605-0.943) for 17OHD, and 0.976 (95% CI 0.927-0.994) for 21OHD (38). Meanwhile, specificity showed good performance in 11βOHD 0.992 (95% CI 0.968-0.999), 17OHD 1.000 (95% CI 0.980-1.000), and 21OHD 0.931 (95% CI 0.869-0.966). Overall diagnostic

Table 1. Baseline characteristics of discovery and validation cohorts

Characteristics	Discovery cohort			Validation cohort		
	Case	Case-like	Healthy	Case	Case-like	Healthy
n	98	30	98	48	15	48
Sex						
Female	78 (79.6)	23 (76.7)	78 (79.6)	37 (77.1)	11 (73.3)	38 (79.2)
Male	20 (20.4)	7 (23.3)	20 (20.4)	11 (22.9)	4 (26.7)	10 (20.8)
Age, years	22 (18-29)	21 (13-32)	25 (18-30)	19 (16-25)	23 (17-27)	23 (16-24)
≤15	13 (13.3)	10 (33.3)	13 (13.3)	8 (16.7)	3 (20.0)	12 (25.0)
>15 to ≤30	62 (63.3)	12 (40.0)	63 (64.3)	21 (43.8)	12 (80.0)	34 (70.8)
>30 to ≤45	19 (19.4)	7 (23.3)	18 (18.4)	10 (20.8)	0 (0.0)	2 (4.2)
>45 years	4 (4.1)	1 (3.3)	4 (4.1)	1 (2.1)	0 (0.0)	0 (0.0)
Manifestations						
Atypical genitalia and abnormal growth	54 (55.1)	8 (26.7)	—	23 (47.9)	5 (33.3)	—
Irregular menses, hirsutism, acne	25 (25.5)	12 (40.0)	—	8 (16.7)	9 (60.0)	—
Hypertension and/or hypokalemia	15 (15.3)	6 (20.0)	—	12 (25.0)	0 (0.0)	—
Adrenal insufficiency	4 (4.1)	2 (6.7)	—	5 (10.4)	2 (13.3)	—
Adrenal hyperplasia or mass	4 (4.1)	2 (6.7)	—	2 (4.2)	0 (0.0)	—
Infertility	8 (8.2)	1 (3.3)	—	4 (8.3)	1 (6.7)	—
Diagnosis						
21OHD, n	66	—	—	33	—	—
Classic (SW and SV)	34 (51.5)	—	—	23 (69.7)	—	—
Nonclassic	31 (47.0)	—	—	9 (27.3)	—	—
Undetermined	1 (1.5)	—	—	1 (3.0)	—	—
17OHD	22	-	—	11	—	—
11βOHD	10	—	—	4	—	—
Other diagnosis[a]	0	30	—	0	15	—

Data are given as n, n (%), or median (interquartile range).
Abbreviations: 11βOHD, 11β-hydroxylase deficiency; 17OHD, 17α-hydroxylase/17,20-lyase deficiency; 21OHD, 21α-hydroxylase deficiency; SV, simple virilizing; SW, salt-wasting.
[a]Other diagnosis included polycystic ovary syndrome, precocious puberty, hirsutism, unknown hypertension and hypokalemia, adrenal hyperplasia, adrenal insufficiency, etc.

Table 2. Clinical characteristics of prospective cohort

Characteristics	Prospective cohort
n	256
Sex	
Female	179 (69.9)
Male	77 (30.1)
Age, years	24 (18-31)
≤15	45 (17.6)
>15 to ≤ 30	140 (54.7)
>30 to ≤45	54 (21.1)
>45	17 (6.6)
Manifestations	
Atypical genitalia and abnormal growth	109 (42.6)
Irregular menses, hirsutism, acne	80 (31.3)
Hypertension and/or hypokalemia	36 (14.1)
Adrenal insufficiency	30 (11.7)
Adrenal hyperplasia or mass	25 (9.8)
Infertility	16 (6.3)

Data are given as n, n (%), or median (interquartile range).

accuracy was 0.988 in 11βOHD, 0.984 in 17OHD, and 0.953 in 21OHD.

Steroidogenesis Score Compared With 17OHP Assay Alone in Diagnosis of 21OHD

Diagnosis of 21OHD currently requires either basal 17OHP (17OHP 0') or ACTH-stimulated 17OHP (17OHP 60') greater than 10 ng/mL. We compared the diagnostic performance of the steroidogenesis score to the 17OHP assay in the polled cohort. A total of 336 patients with 17OHP assay results and 175 patients genotyped as 21OHD were included in the comparison. There was no significant difference between the steroidogenesis score, basal 17OHP, and combined 17OHP (17OHP 0'&60') either in sensitivity or specificity [Supplementary Table 4 (38)]. However, for nonclassic patients, who accounted for 38.2% of all 21OHD patients, the steroidogenesis score classified 97.3% patients as 21OHD, with significantly higher sensitivity compared to 17OHP 0' (0.973 vs 0.840, $P = 0.005$) and was not inferior to 17OHP 0'&60' (0.973 vs 0.947, $P = 0.681$) [Supplementary Table 5 (38)]. Twenty-one patients were referred because of incidental findings of bilateral adrenal masses. Four patients showed elevated 17OHP 0'&60'>10 ng/mL, which can be diagnosed as

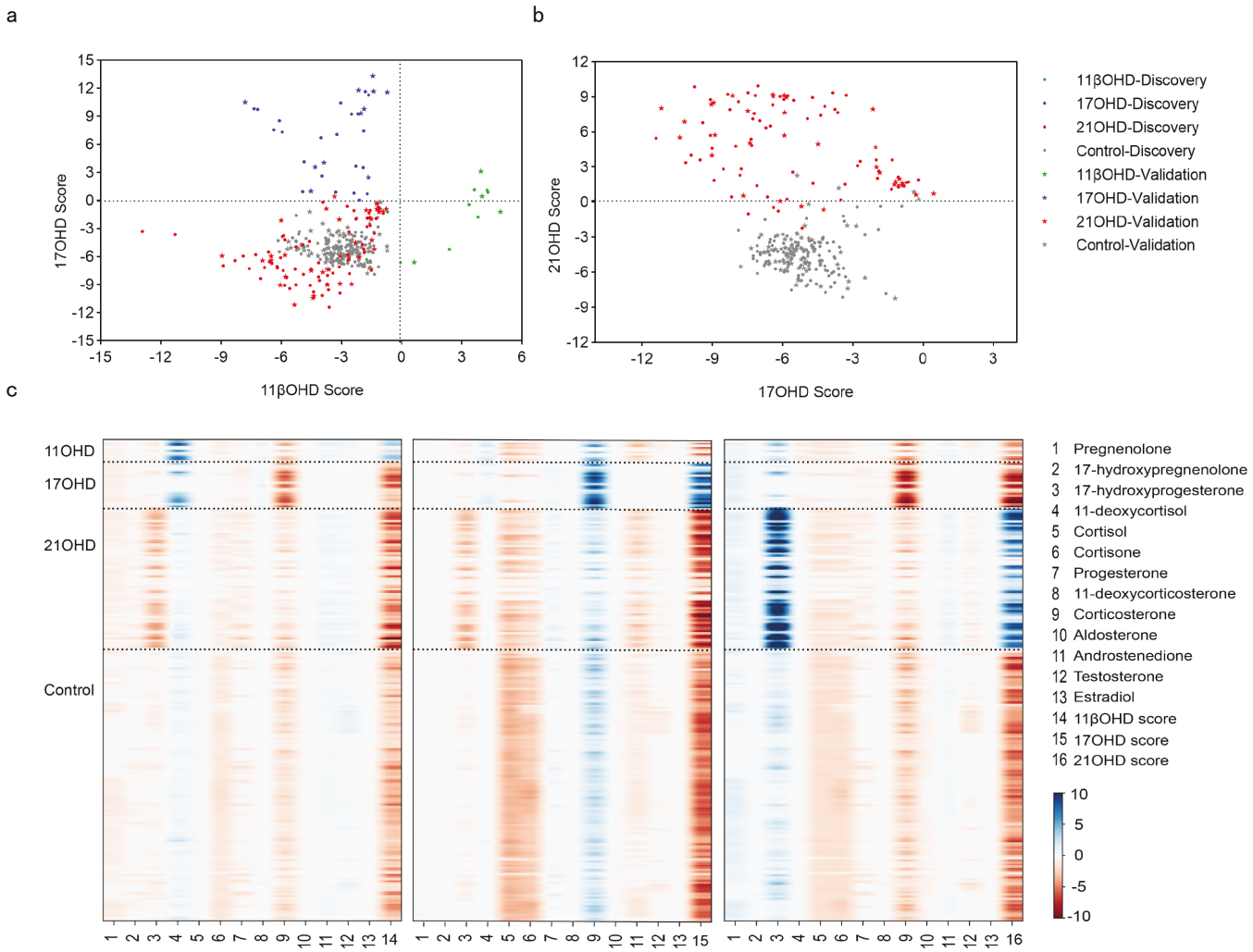

Figure 2. Spatial distribution of cases with 11βOHD, 17OHD, 21OHD, and control subjects in discovery and validation cohorts. (A) Cases with 11βOHD (green) and 17OHD (blue) were clustered. (B) Cases with 21OHD (red) and control (gray) subjects were clustered. The dots indicate participants in discovery cohort and stars represent in validation cohort. (C) Thirteen steroid hormones contributed differently to each score. The dashed line denoted the cutoff of disease classification.

Table 3. Performance metrics of steroidogenesis score in validation cohort (n = 111)

Subtype	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy
11βOHD	1.000 (0.396-1.000)	1.000 (0.957-1.000)	1.000 (0.396-1.000)	1.000 (0.957-1.000)	1.000
17OHD	1.000 (0.679-1.000)	0.990 (0.938-0.999)	0.917 (0.598-0.996)	1.000 (0.953-1.000)	0.991
21OHD	0.909 (0.745-0.976)	0.949(0.867-0.983)	0.882 (0.716-0.962)	0.961 (0.883-0.990)	0.937

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

21OHD according to the published guidelines. None carried biallelic *Cyp21a2* mutations. Consistent with molecular findings, the steroidogenesis score classified none of the 4 patients as 21OHD [Supplementary Table 6 (38)].

Discussion

In this study, we developed a highly interpretable steroid hormone-driven algorithm, namely the steroidogenesis score, to identify and subtype CAH, a group of autosomal recessive hereditary diseases with genetic defects that cause imbalances in multiple steroid hormone pathways. The steroidogenesis score showed high accuracy in diagnosing 3 common CAH subtypes. Most important, for nonclassic 21OHD, the most common and occult CAH subtype, the generation of the steroidogenesis score by a single baseline plasma measurement showed significantly higher diagnostic sensitivity comparing to serum baseline 17OHP, the currently used approach in clinic.

Since each steroidogenesis enzyme catalyzes multiple reactions, impaired enzyme activity causes accumulation of multiple substrates and reduces production of multiple downstream steroids. Genetic defect of individual enzymes therefore is predicted to create a specific pattern of multiple steroid

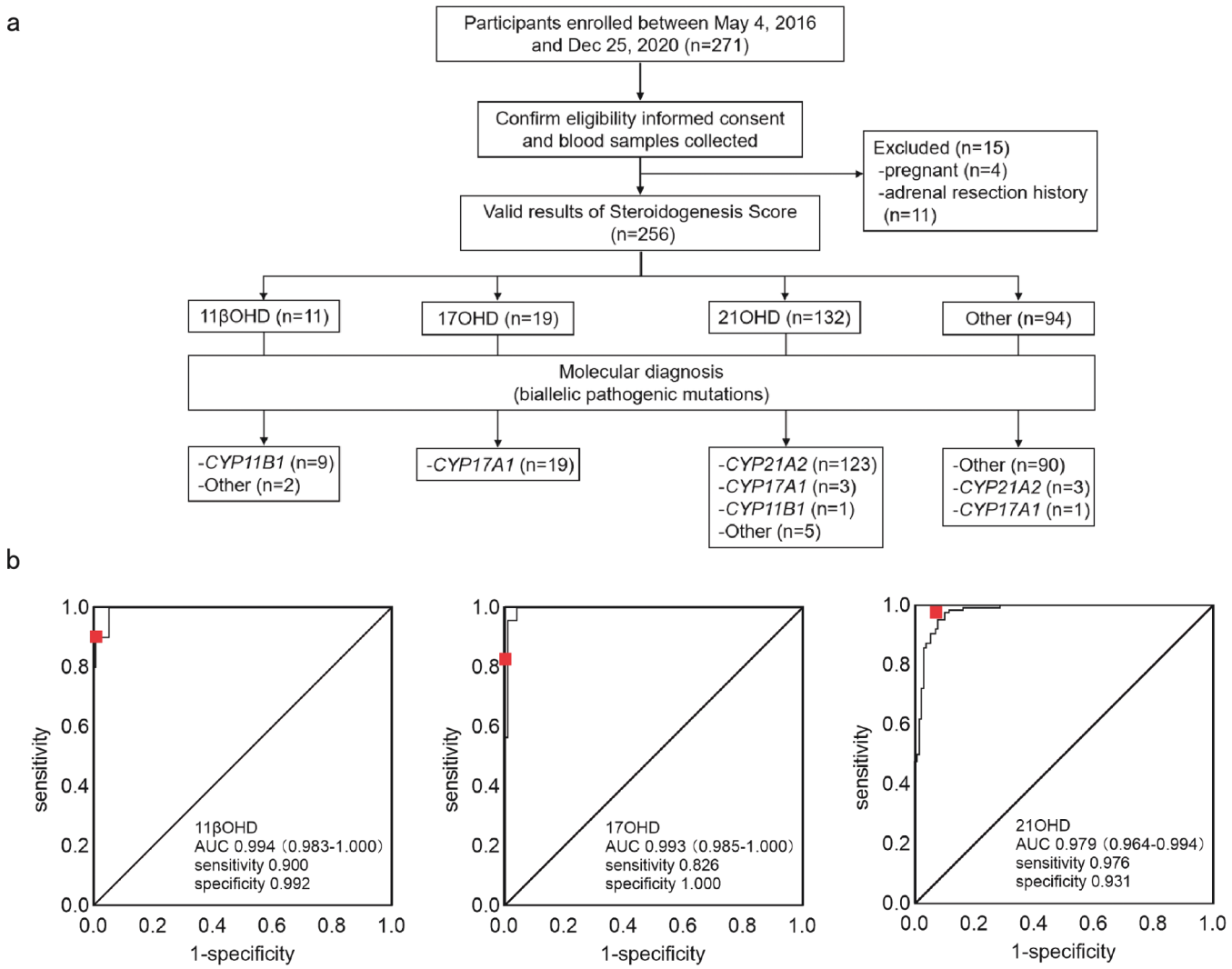

Figure 3. Steroidogenesis score for diagnosis of 21OHD, 17OHD, and 11βOHD in prospective cohort. (A) Flowchart of participants enrolled in prospective cohort. (B) Receiver operating characteristic (ROC) curves of steroidogenesis score. The red dots on each ROC curve indicated the performance of the steroidogenesis score with the cutoff.

hormone imbalances. To detect these patterns and assign them to CAH disease subtypes, we interrogated the levels of 13 circulating steroid hormones observed using a single LC-MS/MS run. By inputting individual steroid levels using a cascade logistic regression model, we established the steroidogenesis score. The model turned out to be highly interpretable. First, the main contributing hormones were all directly up- or downstream of the defected enzymes [Supplementary Figure 1 (38)]: 11-deoxycortisol is the substrate of 11β-hydroxylase in the glucocorticoid pathway; corticosterone is the product of 11β-hydroxylase in the mineralocorticoid pathway; androstenedione is the product of 17α-hydroxylase/17,20-lyase in the sex steroids pathway; and 17OHP is the substrate of 21-hydroxylase in the glucocorticoid pathway (38). In addition, the steroids contributing to the derivation of individual scores was consistent with the known biochemical process (4). In the model, corticosterone remarkably contributed to the 3 scores, and the signs of the coefficients in the 17OHD score were opposite to the 11βOHD and 21OHD scores. Consistent with the biochemical reaction, corticosterone accumulated in patients with 17OHD, while severely it was deficient in patients with 11βOHD and 21OHD. Finally, the model identified the most sensitive multihormonal scores to improve diagnostic accuracy. In 11βOHD patients, 11-deoxycortisol was accumulated, and corticosterone was deficient. Under log transformation, the opposite signs between 11-deoxycortisol and corticosterone coefficients was equivalent to the ratio between the original concentrations. Thus, the *Cyp11b1* mutation patients had significantly lower ratio of corticosterone/11-deoxycortisol compared to the control group. Similarly, the *Cyp21a2* mutation group had a lower ratio of corticosterone/17OHP, and the *Cyp17a1* mutation group had a ratio lower ratio of androstenedione/corticosterone.

We were able to create a score that showed a high degree of specificity using the discovery cohort and that the specificity effectively translated into the validation and prospective cohorts. It was a matter of importance to distinguish 21OHD from 11βOHD, which both present with hyperandrogenism and only 11βOHD with elevated mineralocorticoids. However, hypertension in children is usually occult and difficult to identify, sometimes leading to misdiagnosis (39-42). In our study, 3 patients aged 1, 2, and 4 years at first visit were clinically diagnosed as 21OHD because of atypical genitalia

and elevated 17OHP levels but classified as 11βOHD by the steroidogenesis score and confirmed by molecular diagnosis. Primary hypertension is an extremely common disease in adult patients (43). A 35-year-old female patient presented with hypertension (150/90 mmHg), hyperandrogenism, and infertility and was clinically diagnosed as 11βOHD. However, the steroidogenesis score identified this patient as 21OHD and subsequent genotyping identified biallelic *Cyp21a2* heterozygous pathogenic mutations. Hypertension in this patient was later proven to be primary in nature, thus providing an example of superior performance for the steroidogenesis score.

The steroidogenesis score also performed well in the diagnosis of nonclassic 21OHD. Nonclassic 21OHD is an important cause of androgen excess, and patients usually only present with hirsutism, acne, irregular menses, infertility, or miscarriages (44, 45). The phenotype in female patients greatly overlaps with polycystic ovarian syndrome, the most common infertility disorder in women of reproductive age, affecting 10% to 15% of women worldwide (46). Diagnosis of nonclassic 21OHD typically requires a cosyntropin stimulation test. One recent study used logistic regression modeling of combined basal sera 17OHP4, 21dF, and corticosterone to discriminate patients with and without nonclassic 21OHD (n = 32 vs n = 54) (47). Consistent with their findings, the scores defining 21OHD in our study were predominantly based upon the level of 17OHP and corticosterone. We enrolled a total of 86 patients with nonclassic 21OHD, all with genotyping-confirmed biallelic mutations of *Cyp21a2* gene. We found the steroidogenesis score superior to serum basal 17OHP measurement and not inferior to the cosyntropin stimulation test when diagnosing nonclassic 21OHD. It is noteworthy that patients with incidentally found bilateral adrenal masses may show false-positive findings in cosyntropin stimulation test (48-50), which we found to be excluded as CAH by the steroidogenesis score.

A limitation of the current study is the relatively small sample size. When establishing a data-learning approach, large data sets are usually required owing to the heterogeneous risk factors and confounders (51, 52). However, CAH is caused by a monogene deficiency, and all reported genes are involved in the steroidogenesis pathway. Data sets pertaining to CAH patients are more homogeneous in nature, and therefore, we felt a relatively small sample size merited testing. Further studies with larger sample sizes, especially for 17OHD and other rare CAH subtypes, should be performed in future. Of 9 patients misclassified as 21OHD by the steroidogenesis score, 3 cases carried biallelic pathogenic mutations in *Cyp17a1*, 2 cases with *Hsd3b2* mutations, and 2 cases with *Star* mutations. Additional improvements might be realized from enrolling more complex disease (eg, functional adrenal mass, adrenal carcinoma) to expand the training data.

In summary, by using a data-learning approach, we have established a steroidogenesis score based upon the basal level of 13 circulating steroid hormones, which was able to diagnose and subtype autosomal recessive hereditary CAH. This steroidogenesis score is biochemically interpretable and showed high diagnostic accuracy, especially for nonclassic 21OHD patients. The test is easily conducted with a single blood draw. We believe that the use of data-learning approaches such as the steroidogenesis score may greatly increase the diagnostic accuracy and efficiency for this rare disease and possibly others.

Acknowledgments

We are grateful to the participants for their consent to participate in this study. We would like to thank all the staff at the Department of Endocrine and Metabolic Diseases who contributed to this work.

Financial Support

This study was supported by the National Natural Science Foundation of China (grant 81570702 to L.Y., grant 81621061 to G.N.), the National Key Research and Development Program of China (grant 2016YFC0901503 to L.Y.).

Author Contributions

W-Q.W., S-Y.S., and G.N. conceived and designed the project. W-Q.W., L.Y., Z-Y.Z., W-Z.Z., and W-C.W. managed the study. S-Y.S., L.Y. W-Q.W and S-C.Z. made clinical diagnosis, recruited subjects, and collected samples and clinical phenotypes. R-L.H., J.Z., H-R.L., and Z-H.W. performed steroid profiling. Z-Y.Z, H-X.R., and C.T. performed steroidogenesis score data analyses. L.Y., Z-Y.Z., W-Z.Z., and W-C.W. wrote the manuscript. W-Q.W., S-Y.S., and G.N. contributed to text revision and discussion.

Disclosure Summary

The authors have nothing to disclose and declare no potential conflicts of interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*. 2011;32(1):81-151.
2. Miller WL. Steroidogenesis: unanswered questions. *Trends Endocrinol Metab*. 2017;28(11):771-793.
3. Stewart PM, Krone NP. CHAPTER 15 - the adrenal cortex. In: Melmed S, Polonsky KS, Larsen PR, Kronenberg HM, eds. *Williams Textbook of Endocrinology (Twelfth Edition)*. Philadelphia: W.B. Saunders; 2011:479-544.
4. El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. *Lancet*. 2017;390(10108):2194-2210.
5. Merke DP, Auchus RJ. Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *N Engl J Med*. 2020;383(13):1248-1261.
6. Speiser PW, Arlt W, Auchus RJ, *et al*. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2018;103(11):4043-4088.
7. Li Z, Huang L, Du C, *et al*. Analysis of the screening results for congenital adrenal hyperplasia involving 7.85 million newborns in China: a systematic review and meta-analysis. *Front Endocrinol*. 2021;12:624507.
8. Speiser PW, Dupont B, Rubinstein P, Piazza A, Kastelan A, New MI. High frequency of nonclassical steroid 21-hydroxylase deficiency. *Am J Hum Genet*. 1985;37(4):650-667.
9. Hannah-Shmouni F, Morissette R, Sinaii N, *et al*. Revisiting the prevalence of nonclassic congenital adrenal hyperplasia in US Ashkenazi Jews and Caucasians. *Genet Med*. 2017;19(11):1276-1279.

10. New MI. Extensive clinical experience: nonclassical 21-hydroxylase deficiency. *J Clin Endocrinol Metab*. 2006;91(11):4205-4214.
11. Riedl S, Röhl FW, Bonfig W, *et al*. Genotype/phenotype correlations in 538 congenital adrenal hyperplasia patients from Germany and Austria: discordances in milder genotypes and in screened versus prescreening patients. *Endocr Connect*. 2019;8(2):86-94.
12. Carmina E, Dewailly D, Escobar-Morreale HF, *et al*. Non-classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency revisited: an update with a special focus on adolescent and adult women. *Hum Reprod Update*. 2017;23(5):580-599.
13. Gidlöf S, Falhammar H, Thilén A, *et al*. One hundred years of congenital adrenal hyperplasia in Sweden: a retrospective, population-based cohort study. *Lancet Diabetes Endocrinol*. 2013;1(1):35-42.
14. Rosenfield RL. Normal and premature adrenarche. *Endocr Rev*. 2021;**42**(6):783-814.
15. Nordenström A, Falhammar H. Management of endocrine disease: diagnosis and management of the patient with non-classic CAH due to 21-hydroxylase deficiency. *Eur J Endocrinol*. 2019;180(3):R127-R145.
16. Chrousos GP, Loriaux DL, Mann DL, Cutler GB Jr. Late-onset 21-hydroxylase deficiency mimicking idiopathic hirsutism or polycystic ovarian disease. *Ann Intern Med*. 1982;96(2):143-148.
17. Di Fede G, Mansueto P, Pepe I, Rini GB, Carmina E. High prevalence of polycystic ovary syndrome in women with mild hirsutism and no other significant clinical symptoms. *Fertil Steril*. 2010;94(1):194-197.
18. Escobar-Morreale HF, Carmina E, Dewailly D, *et al*. Epidemiology, diagnosis and management of hirsutism: a consensus statement by the androgen excess and polycystic ovary syndrome society. *Hum Reprod Update*. 2012;18(2):146-170.
19. Watson MS, Mann MY, Lloyd-Puryear MA, Rinaldo P, Howell RR, eds. Newborn screening: toward a uniform screening panel and system. *Genet Med*. 2006;8(suppl 1):1s-252s.
20. Mulaikal RM, Migeon CJ, Rock JA. Fertility rates in female patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *N Engl J Med*. 1987;316(4):178-182.
21. Lo JC, Grumbach MM. Pregnancy outcomes in women with congenital virilizing adrenal hyperplasia. *Endocrinol Metab Clin North Am*. 2001;30(1):207-229.
22. Rauh M. Steroid measurement with LC-MS/MS in pediatric endocrinology. *Mol Cell Endocrinol*. 2009;301(1-2):272-281.
23. 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(9):773-781.
24. Chortis V, Bancos I, Nijman T, *et al*. Urine steroid metabolomics as a novel tool for detection of recurrent adrenocortical carcinoma. *J Clin Endocrinol Metab*. 2020;105(3):e307-e318.
25. Hines JM, Bancos I, Bancos C, *et al*. High-resolution, accurate-mass (HRAM) mass spectrometry urine steroid profiling in the diagnosis of adrenal disorders. *Clin Chem*. 2017;63(12):1824-1835.
26. Kulle AE, Riepe FG, Hedderich J, *et al*. LC-MS/MS based determination of basal- and ACTH-stimulated plasma concentrations of 11 steroid hormones: implications for detecting heterozygote CYP21A2 mutation carriers. *Eur J Endocrinol*. 2015;173(4):517-524.
27. Kamrath C, Hartmann MF, Pons-Kühnemann J, Wudy SA. Urinary GC-MS steroid metabotyping in treated children with congenital adrenal hyperplasia. *Metab Clin Exp*. 2020;112(9):154354.
28. Lasarev MR, Bialk ER, Allen DB, Held PK. Application of principal component analysis to newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2020;105(8):e2930-e2940.
29. Hicks RA, Yee JK, Mao CS, *et al*. Precursor-to-product ratios reflect biochemical phenotype in congenital adrenal hyperplasia. *Metabolomics*. 2014;10(1):123-131.
30. Taylor DR, Ghataore L, Couchman L, *et al*. A 13-steroid serum panel based on LC-MS/MS: use in detection of adrenocortical carcinoma. *Clin Chem*. 2017;63(12):1836-1846.
31. Beaulieu-Jones B, Finlayson SG, Chivers C, *et al*. Trends and focus of machine learning applications for health research. *JAMA Netw Open*. 2019;2(10):e1914051.
32. Antoniou T, Mamdani M. Evaluation of machine learning solutions in medicine. *CMAJ*. 2021;193(36):e1425-e1429.
33. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
34. Wang W, Han R, Yang Z, *et al*. Targeted gene panel sequencing for molecular diagnosis of congenital adrenal hyperplasia. *J Steroid Biochem Mol Biol*. 2021;211:105899.
35. Richards S, Aziz N, Bale S, *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
36. Pedregosa F, Varoquaux G, Gramfort A, *et al*. Scikit-learn: Machine learning in python. *J Mach Learn Res*. 2011;12(85):2825-2830.
37. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998;17(8):857-872.
38. Ye L, Zhao Z, Ren H, *et al*. Data from a multi-classifier system to identify and subtype congenital adrenal hyperplasia based on circulating steroid hormones. Uploaded March 31, 2022. *Figshare*. 2022. https://doi.org/10.6084/m9.figshare.19487660.v3
39. Zhang M, Liu Y, Sun S, *et al*. A prevalent and three novel mutations in CYP11B1 gene identified in Chinese patients with 11-beta hydroxylase deficiency. *J Steroid Biochem Mol Biol*. 2013;133:25-29.
40. Hinz L, Pacaud D, Kline G. Congenital adrenal hyperplasia causing hypertension: an illustrative review. *J Hum Hypertens*. 2018;32(2):150-157.
41. Bulsari K, Falhammar H. Clinical perspectives in congenital adrenal hyperplasia due to 11β-hydroxylase deficiency. *Endocrine*. 2017;55(1):19-36.
42. Ardhanari S, Kannuswamy R, Chaudhary K, Lockette W, Whaley-Connell A. Mineralocorticoid and apparent mineralocorticoid syndromes of secondary hypertension. *Adv Chronic Kidney Dis*. 2015;22(3):185-195.
43. Lu J, Lu Y, Wang X, *et al*. Prevalence, awareness, treatment, and control of hypertension in China: data from 1·7 million adults in a population-based screening study (China PEACE million persons project). *Lancet*. 2017;390(10112):2549-2558.
44. Reichman DE, White PC, New MI, Rosenwaks Z. Fertility in patients with congenital adrenal hyperplasia. *Fertil Steril*. 2014;101(2):301-309.
45. New MI, Ghizzoni L, Meyer-Bahlburg H, Khattab A, Reichman D, Rosenwaks Z. Fertility in patients with nonclassical congenital adrenal hyperplasia. *Fertil Steril*. 2019;111(1):13-20.
46. Yildiz BO, Bozdag G, Yapici Z, Esinler I, Yarali H. Prevalence, phenotype and cardiometabolic risk of polycystic ovary syndrome under different diagnostic criteria. *Hum Reprod*. 2012;27(10):3067-3073.
47. Turcu AF, El-Maouche D, Zhao L, *et al*. Androgen excess and diagnostic steroid biomarkers for nonclassic 21-hydroxylase deficiency without cosyntropin stimulation. *Eur J Endocrinol*. 2020;183(1):63-71.
48. Calissendorff J, Juhlin CC, Sundin A, Bancos I, Falhammar H. Adrenal myelolipomas. *Lancet Diabetes Endocrinol*. 2021;**9**(11):767-775.
49. Bourdeau I, El Ghorayeb N, Gagnon N, Lacroix A. Management of endocrine disease: differential diagnosis, investigation and therapy of bilateral adrenal incidentalomas. *Eur J Endocrinol*. 2018;179(2):R57-R67.
50. Libé R, Coste J, Guignat L, *et al*. Aberrant cortisol regulations in bilateral macronodular adrenal hyperplasia: a frequent finding in a prospective study of 32 patients with overt or subclinical Cushing's syndrome. *Eur J Endocrinol*. 2010;163(1):129-138.
51. Capper D, Jones DTW, Sill M, *et al*. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018;555(7697):469-474.
52. Zeevi D, Korem T, Zmora N, *et al*. Personalized nutrition by prediction of glycemic responses. *Cell*. 2015;163(5):1079-1094.