



HAL
open science

Combining metabolomics and machine learning models as a tool to distinguish non-classic 21-hydroxylase deficiency from polycystic ovary syndrome without adrenocorticotrophic hormone testing

Guillaume Bachelot, Anne Bachelot, Marion Bonnier, Joe-Elie Salem, Dominique Farabos, Severine Trabado, Charlotte Dupont, Peter Kamenicky, Muriel Houang, Jean Fiet, et al.

► To cite this version:

Guillaume Bachelot, Anne Bachelot, Marion Bonnier, Joe-Elie Salem, Dominique Farabos, et al.. Combining metabolomics and machine learning models as a tool to distinguish non-classic 21-hydroxylase deficiency from polycystic ovary syndrome without adrenocorticotrophic hormone testing. *Human Reproduction*, 2023, 38 (2), pp.266-276. 10.1093/humrep/deac254 . hal-03996582

HAL Id: hal-03996582

<https://hal.science/hal-03996582>

Submitted on 20 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Draft Manuscript For Review. Reviewers should submit their review at <http://mc.manuscriptcentral.com/humrep>

Combining metabolomics and machine learning models as a tool to distinguish non-classic 21OH deficiency from PCOS without ACTH testing

Journal:	<i>Human Reproduction</i>
Manuscript ID	HUMREP-22-0749.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Bachelot, Guillaume; Assistance Publique - Hôpitaux de Paris, Service de Biologie de La Reproduction-CECOS, Hôpital Tenon; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938; Assistance Publique - Hopitaux de Paris, Département de Métabolomique Clinique, Hôpital Saint Antoine</p> <p>Bachelot, Anne; Hopital Universitaire Pitie Salpetriere, Service d'Endocrinologie et Médecine de la reproduction</p> <p>BONNIER, Marion; Hopital Universitaire Pitie Salpetriere, Service d'Endocrinologie et Médecine de la reproduction</p> <p>Salem, Joe Elie; Assistance Publique - Hopitaux de Paris, CIC-1901 and Department of Pharmacology, Pitié-Salpêtrière Hospital</p> <p>Farabos, Dominique; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938</p> <p>Trabado, Severine; Assistance Publique - Hopitaux de Paris, Service de Génétique Moléculaire, Pharmacogénétique et Hormonologie, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, Le Kremlin-Bicêtre, France, INSERM UMR-U1185, Paris-Saclay University</p> <p>Dupont, Charlotte; Assistance Publique - Hopitaux de Paris, Hôpital Tenon, service de biologie de la reproduction CECOS; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938</p> <p>Kamenicky, Peter; Assistance Publique - Hopitaux de Paris, University Paris-Saclay, F-91405 Orsay, France, Department of Reproductive Endocrinology, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, F-94275, Le Kremlin-Bicêtre, France, INSERM UMR-S 1185, Paris-Saclay University</p> <p>Houang, Muriel; Assistance Publique - Hopitaux de Paris, Explorations Fonctionnelles Endocriniennes, Hôpital Armand Trousseau; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938</p>

	<p>Fiet, Jean; Assistance Publique - Hopitaux de Paris, Département de Métabolomique Clinique, Hôpital Saint Antoine; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938</p> <p>LE BOUC, Yves; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938</p> <p>Young, Jacques; Assistance Publique - Hopitaux de Paris, University Paris-Saclay, F-91405 Orsay, France, Department of Reproductive Endocrinology, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, F-94275, Le Kremlin-Bicêtre, France, INSERM UMR-S 1185, Paris-Saclay University</p> <p>LAMAZIERE, Antonin; Sorbonne University, Saint Antoine Research Center, INSERM UMR 938; Assistance Publique - Hopitaux de Paris, Département de Métabolomique Clinique, Hôpital Saint Antoine</p>
Keywords:	metabolomics, 21-Hydroxylase deficiency, late onset congenital adrenal hyperplasia, hirsutism, machine learning, algorithm, Polycystic ovary syndrome, hyperandrogenism
Subject Section:	Reproductive endocrinology

SCHOLARONE™
Manuscripts

1 **Combining metabolomics and machine learning models as a tool to**
2 **distinguish non-classic 21OH deficiency from PCOS without ACTH testing**

3 Running Title: Steroid profiling and machine learning to bypass ACTH test

4
5 Guillaume Bachelot^{1,2,3*}, Anne Bachelot⁴, Marion Bonnier⁴, Joe-Elie Salem⁵, Dominique
6 Farabos², Severine Trabado⁶, Charlotte Dupont^{1,2}, Peter Kamenicky⁷, Muriel Houang⁸, Jean
7 Fiet^{2,3}, Yves Le Bouc⁸, Jacques Young⁷, Antonin Lamazière^{2,3}

8
9 1- Service de Biologie de La Reproduction-CECOS, Hôpital Tenon, AP-HP, Sorbonne Université, 75020 Paris, France.

10 2- Sorbonne Université, Saint Antoine Research Center, INSERM UMR 938, 75012 Paris, France.

11 3- Département de Métabolomique Clinique, Hôpital Saint Antoine, AP-HP, Sorbonne Université, 27 Rue Chaligny, 75012 Paris, France.

12 4 - Service d'Endocrinologie et Médecine de La Reproduction, Centre de Référence Des Maladies Endocriniennes Rares de La Croissance et
13 du Développement, Centre Des Pathologies Gynécologiques Rares, Hôpital Pitié Salpêtrière (APHP), Sorbonne Université, 75013, Paris,
14 France.

15 5 - CIC-1901 and Department of Pharmacology, Pitié-Salpêtrière Hospital, AP-HP, Sorbonne Université, Inserm, 75013 Paris, France.

16 6 - Service de Génétique Moléculaire, Pharmacogénétique et Hormonologie, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, Le
17 Kremlin-Bicêtre, France, INSERM UMR-U1185, Paris-Saclay University, Le Kremlin Bicêtre, France

18 7 - University Paris-Saclay, F-91405 Orsay, France, Department of Reproductive Endocrinology, Assistance Publique-Hôpitaux de Paris,
19 Bicêtre Hospital, F-94275, Le Kremlin-Bicêtre, France, INSERM UMR-S 1185, Paris-Saclay University, Le Kremlin Bicêtre, F-94276,
20 France

21 8 - Explorations Fonctionnelles Endocriniennes, Hôpital Armand Trousseau, AP-HP, 26 Av Dr Netter, Paris 75012, France. Sorbonne
22 Université, Saint Antoine Research center, CRSA, INSERM, Sorbonne Université, Paris, France

23
24
25 * Correspondence: guillaume.bachelot@aphp.fr
26
27
28

29 **Abstract**

30 **Study question:** Can a combination of metabolomic signature and machine learning models
31 distinguish non-classic 21OH deficiency (NC21OHD) from PCOS without ACTH testing?

32
33 **Summary answer:** A single sampling methodology may be an alternative to the dynamic
34 ACTH test in order to exclude the diagnosis of NC21OHD in the presence of a clinical
35 hyperandrogenic presentation at any time of the menstrual cycle.

36 **What is known already:** The clinical presentation of patients with NC21OHD is similar with
37 that for other disorders of androgen excess. Currently, cosyntropin stimulation remains the
38 gold standard diagnosis of NC21OHD.

39 **Study design, size, duration:** The study was designed using a bicentric recruitment: an internal
40 training set included 19 women with NC21OHD and 19 controls used for developing the model;
41 a test set included 17 NC21OHD, 72 controls and 266 PCOS patients used to evaluate the
42 performance of the diagnostic strategy thanks to a machine learning approach.

43 **Participants/materials, setting, methods:** Fifteen steroid species were measured in serum by
44 liquid chromatography–mass spectrometry (LC–MS/MS). This set of 15 steroids (defined as
45 “steroidome”) used to map the steroid biosynthesis pathway was the input for our models.

46 **Main results and the role of chance:** From a single sample, modeling involving metabolic
47 pathway mapping by profiling 15 circulating steroids allowed to perfectly identify NC21OHD
48 from a confounding PCOS population. The constructed model using baseline LC-MS/MS-
49 acquired steroid fingerprinting successfully excluded all 17 NC21OHDs (sensitivity and
50 specificity of 100%) from 266 PCOS from an external testing cohort of originally 549 women,
51 without the use of ACTH testing. Blood sampling timing during the menstrual cycle phase did
52 not impact the efficiency of our model.

53 **Limitations, reasons for caution:** The main limitations were the use of a restricted and fully
54 prospective cohort as well as an analytical issue, as not all laboratories are equipped with mass
55 spectrometers able to routinely measure this panel of 15 steroids. Moreover, the robustness of
56 our model needs to be established with a larger prospective study for definitive validation in
57 clinical practice.

58 **Wider implications of the findings:** This tool makes it possible to propose a new semiology
59 for the management of hyperandrogenism. The model presents better diagnostic performances
60 compared to the current reference strategy. The management of patients may be facilitated by
61 limiting the use of ACTH tests. Finally, the modeling process allows a classification of steroid
62 contributions to rationalize the biomarker approach and highlight some underlying
63 pathophysiological mechanisms.

64 **Study funding/competing interest(s):** This study was supported by “Agence Française de
65 Lutte contre le dopage” and DIM Région Ile de France; This study was supported by the French
66 institutional PHRC 2010-AOR10032 funding source and APHP). All authors declare no
67 competing financial interests.

68 **Key words:**

69 21-Hydroxylase deficiency, late onset congenital adrenal hyperplasia, hirsutism, machine
70 learning, algorithm, Polycystic ovary syndrome, hyperandrogenism.

71

72 **Introduction**

73 21-hydroxylase deficiency (21OHD) is a rare genetic disorder with autosomal recessive
74 inheritance caused by deleterious mutations in *CYP21A2* (1). 21OHD is the most common
75 cause (95%) of congenital adrenal hyperplasia (CAH) (Speiser *et al.*, 1992). The different
76 clinical forms of 21OHD depend on the severity of the mutations. Besides the classical form
77 (C21OHD), which is due to complete or severe 21OH deficiency revealed at birth, there exists
78 the nonclassical form (NC21OHD), which is due to partial enzymatic defect, with late-onset
79 symptoms; diagnosis is usually made in late childhood or after puberty (3). The partial
80 preservation of cortisol and aldosterone secretion in patients with NC21OHD (Kamenický *et*
81 *al.*, 2020) is explained by the existence of at least one of the alleles of a *CYP21A2* mutation
82 encoding a 21-hydroxylase that retains a minimum residual enzyme activity of 50% (New *et*
83 *al.*, 2013). NC21OHD is most often revealed in adolescent girls and young women by a clinical
84 pattern of hyperandrogenism accompanied by oligomenorrhea associated with increased
85 circulating testosterone (Admoni *et al.*, 2006). Infertility and episodes of miscarriage may also
86 occur in some women (New, 2006). The clinical presentation of patients with NC21OHD is
87 similar to that seen in a common female condition, polycystic ovarian syndrome (PCOS),
88 wherein signs of hyperandrogenism associated with oligomenorrhea and infertility are also very
89 prevalent (Matheson and Bain, 2019). As a result, many national and international consensus
90 recommend that the diagnosis of NC21OHD be systematically ruled out before the diagnosis
91 of PCOS is definitively confirmed (Teede *et al.*, 2018). To exclude NC21OHD, many
92 consensus recommend that an assay of 17-hydroxyprogesterone (17OHP) be performed in
93 the morning during the follicular phase and that the diagnosis of NC21OHD be considered if
94 the basal serum concentration of 17OHP is increased (Azziz *et al.*, 1999; Speiser *et al.*, 2000).
95 The thresholds proposed in the literature are variable and depend on the 17OHP assay method
96 used—that is, either immunoassay or mass spectrometry (Makela and Ellis, 1988). Thus,

97 depending on the centers and methods, the diagnosis of NC21OHD will be considered if
98 baseline 17OHP levels are ranging between 2 and 4 ng/mL (Azziz *et al.*, 1999; Young *et al.*,
99 2010). In such cases, it is recommended that ACTH stimulation be performed and the diagnosis
100 established if the 17OHP peak value is greater than 10 ng/mL (Dewailly *et al.*, 1986; Bello *et*
101 *al.*, 2017). However, a significant percentage of women with NC21OHD may have circulating
102 levels of 17OHP close to or even below 2 ng/mL (Bidet *et al.*, 2009, 2010, Escobar-Morreale
103 *et al.*, 2008). In addition, patients with increased serum 17OHP levels, but without NC21OHD,
104 have also been reported (Rosenfield and Ehrmann, 2016).

105 This raises the question of systematizing the ACTH test in all women with a clinical
106 presentation compatible with PCOS syndrome, so as to exclude NC21OHD. However, medico-
107 economic and practical obstacles limit the generalization of ACTH stimulation test. First, PCOS
108 is a very prevalent disease whose frequency has been estimated at between 5 and 10% of
109 women, whereas the prevalence of NC21OHD is much lower, estimated at nearly 1/1000 in
110 most countries (Kuttann *et al.*, 1985). In addition, many non-specialized centers treating women
111 suspected of having PCOS do not have easy access to this dynamic test. On the other hand,
112 making the diagnosis of NC21OHD is a primordial issue, as it has been demonstrated that more
113 than 50% of patients affected by this condition carry severe loss-of-function CYP21A2
114 mutations in one allele (Bidet *et al.*, 2009). These patients are at high risk of giving birth to
115 children affected by the classical form of CAH if their partner also carries a severe heterozygous
116 mutation (probability evaluated at 1/60 people in the general population) (Bidet *et al.*, 2009).

117 The concept of profiling or multiplexing the circulating steroidome is an emerging
118 paradigm shift, as this method appears to be more accurate than single-steroid immunoassay
119 quantitation to stratify or classify patients from different conditions. Specifically, in a recent
120 preliminary work, investigators have shown that, compared to an isolated measure of serum
121 17OHP, simultaneous measurements of a panel of several steroids in the basal state more

122 accurately distinguish subjects with NC21OHD from a control population (Turcu *et al.*, 2020a).
123 In parallel to these advances in hormonal evaluation, machine learning (ML) approaches are
124 emerging as interesting tools in diagnostic procedures due to their classification strength
125 compared to classical univariate analysis traditionally used in clinical practice. As NC21OHD
126 is an adrenal disorder and PCOS has an ovarian origin, we hypothesized that an evaluation of
127 a steroid panel including both predominantly adrenal steroids and others with a dual adrenal
128 and ovarian origin would help to better distinguish between these two conditions.

129 In this work, we combined hormonal data obtained by simultaneous measurements of a
130 set of circulating steroids measured by liquid chromatography-mass spectrometry (LC–
131 MS/MS) in the basal state (metabolomic approach) together with a machine learning approach
132 in order to build a model that could reliably distinguish women with NC21OHD from both
133 normal women and females affected by PCOS without the use of an ACTH stimulation test.

134

135 **Subjects, materials and methods**

136 **Study overview, population and data source**

137 *Study overview*

138 As shown in Figure 1, the study was designed using a bicentric recruitment: an internal training
139 set included 38 female subjects (19 with NC21OHD and 19 healthy controls) from the
140 Department of Reproductive Endocrinology, Bicêtre Hôpital, Le Kremlin-Bicêtre, France
141 (KB); an external test set included 355 women (17 NC21OHD, 72 controls and 266 PCOS)
142 from the Department of Reproductive Endocrinology, Pitié-Salpêtrière Hospital, Paris, France
143 (PSL), evaluated from 2019 to 2020.

144 The training set was used for developing the machine learning model. The test set was
145 used to evaluate the performance of the model. PCOS patients (with a clinical presentation
146 similar to those with NC21OHD) from PSL hospital were included in the test set to confirm

147 that the model succeeded in classifying PCOS patients out of the NC21OHD group. The
148 structure of this test cohort was designed to be in accordance with the prevalence of those
149 conditions in a context of hyperandrogenism.

150 The ethics committee (“Comité de Protection des Personnes”) approved the study. All the
151 participants signed a written informed consent.

152

153 Subjects - population

154 Training set (KB): Nineteen women with NC21OHD with biallelic *CYP21A2* mutations were
155 included as the training set in this study. These patients were previously evaluated in a clinical
156 study (Kamenický *et al.*, 2020) prospective case–control open-label blinded-endpoint non-
157 inferiority (ClinicalTrials.gov Id: NCT01862380) at a tertiary referral medical center and
158 clinical investigation center (2013 – 2016). This study was approved by the local ethics
159 committee (Comité de Protection des Personnes Ile de France VII, approval no. 11–034) and
160 the competent authority (ANSM). All subjects gave written informed consent to participate to
161 the study. In brief, these eligible NC21OHD patients were premenopausal women aged 18–45
162 years with biallelic *CYP21A2* mutations (homozygotes or compound heterozygotes) and serum
163 17OHP concentration >10 ng/mL after stimulation with ACTH (250µg of Synacthen®)
164 performed at inclusion visit. Main clinical and genetics characteristics of these patients have
165 been described elsewhere (Kamenický *et al.*, 2020). In brief, seven of these patients carried the
166 mild p.V281L *CYP21A2* mutation in the homozygous state and 12 were compound
167 heterozygotes with at least one minor mutation (for details see Supplementary Table SI).
168 Twelve patients received, at some point in their lives, hydrocortisone (10–20 mg per day) and/or
169 dexamethasone (0.25–0.5 mg per day), prescribed transiently for important hyperandrogenism
170 or impaired fertility, but all had received these treatments more than 1 year before inclusion.
171 Nineteen age- and BMI-matched healthy women, whose characteristics have been detailed
172 elsewhere (Kamenický *et al.*, 2020), were included as controls. These eligible controls were

173 premenopausal healthy women with serum 17OHP <2.0 ng/mL after stimulation with 250 µg
174 of ACTH.

175 Exclusion criteria for all subjects were pregnancy, breastfeeding, acute infectious
176 disease (in particular, diarrhoea) within the last 7 days before inclusion, coronary artery disease,
177 abnormal electrocardiogram, epilepsy, drug allergy, blood hemoglobin <10 g/dL, and ongoing
178 treatment with diuretics or drugs interfering with renin angiotensin aldosterone system (i.e.,
179 angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, and
180 mineralocorticoid receptor inhibitors).

181

182 Test set (PSL): Two groups of patients with hyperandrogenism and/or hyperandrogenemia were
183 recruited from this institution between 2019 and 2020 (Flow Chart). We included 266 PCOS
184 patients and 17 women with NC21OHD patients that were selected from routine clinical
185 practice.

186 Thirteen of these 17 N21OHD patients carried the mild p.V281L CYP21A2 mutation
187 in the homozygous state, and 4 were compound heterozygotes (V281L/P30L, IVS2-13C-
188 1>G/P30L, IVS2-13C-1>G/V281L, 126C>T, 113G>A, 110T>C/126C>T, 113G>A) with at
189 least one minor mutation (for details see Supplementary Table SI). The diagnosis of PCOS was
190 based on 2018 International Guidelines which resume the 2003 Rotterdam criteria (Teede *et*
191 *al.*, 2018). PCOS diagnosis was based on the presence of 2 out of the following 3 criteria:
192 irregular menstrual cycles, clinical hyperandrogenism and/or hyperandrogenemia and the
193 polycystic ovary on ultrasound after excluding hyperprolactinemia and Cushing syndrome
194 (Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004). CAH
195 diagnosis was excluded according to the guidelines, i.e. a 17OHP value below 10 ng/mL after
196 ACTH testing.

197

198 Data from seventy-two healthy female controls were also included from a multicenter
199 prospective observational case-control study (2011–2014) exploring early cardiovascular
200 damage in adult men and women with CAH due to late onset 21 α -hydroxylase deficiency
201 (CARDIOHCS [NCT01807364]) (Rosenbaum *et al.*, 2019).

202 Exclusion criteria were: known history of cardiovascular disease and pregnancy.
203 Women using combined contraceptive were excluded.

204 All patients gave written informed consent to participate. This study was performed in
205 accordance with the Declaration of Helsinki. Healthy volunteers of this study were recruited by
206 the Centre d'Investigation Clinique Paris Est (CIC-1421, Pitié-Salpêtrière Hospital, Paris).

207

208 LC-MS/MS steroid profiling - data source and acquisition - variables

209 Fifteen steroid species (21-deoxycortisol (21-DF), 11 β -hydroxyandrostenedione (11 β -OHA),
210 androstenedione (D4A), testosterone (T), 17-OH-progesterone (17OHP), 17-OH-pregnenolone
211 (17-OHPreg), dehydroepiandrosterone (DHEA), aldosterone (ALDO), cortisone (E), cortisol
212 (F), corticosterone (B), 11-deoxycorticosterone (DOC), 21-deoxycorticosterone (21-DB),
213 progesterone (P) and pregnenolone (Preg)) were measured in serum by LC–MS/MS as
214 described elsewhere (Fiet *et al.*, 2017). This set of 15 steroids (defined as the “steroidome”)
215 used to map the steroid biosynthesis pathway were the input for our models. All the steroid
216 analyses were performed in the same clinical laboratory (Saint-Antoine Hospital, Paris –
217 France). There were no missing values.

218

219 **Data processing, machine learning algorithm implementation, performance measures** 220 **and statistical analysis**

221 After performing regular statistical analysis (mean comparison, Mann–Whitney–Wilcoxon
222 using Python 3.7.6 environment and the SciPy Package version 1.4.1), we trained a machine

223 learning model using the steroidome (set of 15 steroids) as input to detect (or predict)
224 NC21OHD from both normal and PCOS women without the use of a ACTH stimulation test.
225 We used orthogonal partial least squares (OPLS) as the standard metabolomic machine learning
226 algorithm. As previously shown in Bachelot et al. and Agnani et al. (Bachelot *et al.*, 2021;
227 Agnani *et al.*, 2022), we applied the OPLS algorithm for its performance on metabolomic data
228 and the possibility to obtain easily interpretable score plot. The OPLS model was performed
229 using SIMCA 16.0.1 software (Umetrics, Sweden).

230 Statistical modeling was performed using a multicentric design (two cohorts of patients
231 from two independent clinical institutions) as described in Figure 1: the OPLS model was
232 obtained by training the algorithm on the training set, and the performance of this model was
233 assessed on the test set.

234 As a classification task, the model performances were evaluated using confusion matrix
235 and accuracy score (the fraction of correct predictions).

236 Eventually, an additional evaluation was performed. We compared the performance of
237 the model versus the standard cut-off based on the basal 17OHP concentration of 2 ng/mL and
238 10 ng/mL.

239 **Results**

240 *Study cohorts – baseline characteristics*

241 In the training set, as previously described in Kamenicky et al., controls were matched to
242 NC21OHD for age and BMI (Kamenický *et al.*, 2020).

243 In the test set, (i.e., PSL) PCOS and NC21OHD patients exhibited similar baseline
244 characteristics (Table I), except for oligomenorrhea, follicle number (right and left) and serum
245 LH. As expected, these parameters were increased in PCOS compared with NC21OHD.

246 As a bicentric study, basal steroidomic profiles were compared in the different control
247 (KB vs. PSL) and NC21OHD (KB vs. PSL) groups. As shown in Supplementary Table SII,

248 levels of 17OHP, progesterone, DOC, 21-DF, corticosterone, aldosterone, and cortisone were
249 homogeneous in patients and controls from both institutions. In the control groups, some steroid
250 species—such as testosterone, 21 deoxycortisol, and 17OH-pregnenolone—exhibited
251 differences between the two institutions, thereby reflecting natural inter-individual variabilities
252 while remaining within the same orders of magnitude and within the normal value ranges.

253 Figure 2A and figure 2B showed respectively the circulating basal concentrations of
254 17OHP and circulating concentrations of 17OHP after ACTH stimulation in the 2 groups of
255 NC21OHD patients (n = 36) and PCOS patients (n = 266), which were significantly different
256 ($p < 0.0001$). Some overlapping among NCAH and PCOS patients was observed for basal
257 17OHP concentrations. However, all PCOS (n = 266) exhibited a concentration well below the
258 10 ng/mL cutoff values of 17OHP after injection. Likewise, none of the NC21OHD patients (n
259 = 36) exhibited post-stimulation 17OHP values below 10 ng/mL. Thus, there was strictly no
260 overlap between NC21OHDs (with a minimum stimulated 17OHP concentration=16.40ng/mL)
261 and PCOSs (with a maximum elevated 17OHP concentration=4.55ng/mL). These data are in
262 accordance with the literature on 21-hydroxylase deficiency, which proposed that when a
263 patient was genotyped with a biallelic mutation, the peak after ACTH was consistently above
264 10 ng/mL (Bidet *et al.*, 2009, 2010). These results excluded the possibility that any of the PCOS
265 patients in the test set would be affected by 21-hydroxylase deficiency.

266

267 *From model to diagnosis*

268 The main goal of our study was to avoid the ACTH stimulation test through the use of
269 mathematical algorithms and metabolomics. For that purpose, we built the model described in
270 Figure 1, which was implemented with the circulating basal steroids signature (15 steroids) and
271 trained to efficiently differentiate two groups of individuals: female controls and NC21OHD
272 patients.

273

274 Machine learning modeling (Figure 1)

275 Step 1: Circulating basal steroids signatures were acquired from LC–MS/MS and preprocessing
276 (normalization).

277 Step 2: OPLS-DA was performed on the training set (KB). The corresponding observation score
278 plots (Figure 3A) confirmed the separation between NC21OHD (red dots) and control (green
279 dots) individuals. Indeed, the horizontal direction of the score plot will capture variation
280 between the two groups (NC21OHD and controls). The Y axis represents the decision
281 boundary, and allowed us to define two areas: in the left quadrants, NC21OHD patients with
282 specific steroids signature; in the right quadrants, controls with their associated steroids pattern.

283

284 Step 3: Once the OPLS-DA model had been built and trained on training set (KB), the model
285 was then tested on an independent external test set from PSL; the corresponding individuals
286 were plotted (black stars) regardless of their label (Figure 3B). Second, NC21OHD (red stars)
287 and control (green stars) were colored according to their diagnostic status (Figure 3C).
288 Accuracy score on independent test set was 1.00. All 17 NC21OHD and 72 controls from PSL
289 were therefore correctly classified by the OPLS-DA model.

290

291 Extension of the algorithm to solve an additional practical clinical issue: to distinguish women
292 with NC21OHD from PCOS women.

293 As mentioned above, one important clinical issue is the differentiation between NC21OHD and
294 polycystic ovary syndrome (PCOS) in female patients with hyperandrogenism (Table I).
295 Therefore, PCOS patients from PSL were tested in the model. The objective was to ensure that
296 the algorithm was able to accurately identify and segregate the NC21OHD patients, not only
297 from controls but also from PCOS patients. After being tested on the NC21OHD/controls

298 model, PCOS patients were plotted on OPLS-DA (Figure 3D); all were located in the controls
299 area (266 out 266).

300 The Variable Importance for the Projection (VIP) plot summarized the importance
301 (contribution) of the steroids in the model (Figure 4). The VIP plot displays the VIP values as
302 a column plot with confidence intervals (95%) derived from jack-knifing. The most segregative
303 basal steroid species were those with VIP values larger than 1, which included 21DF, 21DB,
304 11BOHA, 17OHP, A, B, and T. Interestingly, except for compound B, all these steroids
305 precursors are upstream of the enzymatic blockade observed in subjects affected by NC21OHD.

306 This strategy was able to set up an algorithm that is also capable of efficiently stratifying
307 all the patients of the test set without going through an ACTH stimulation test.

308

309 Comparison of score performance of the model versus 17OHP thresholds proposed in the
310 literature to detect NC21OHD.

311 To demonstrate the contribution of the model compared to the classical routine procedure,
312 based on basal thresholds concentrations of 17OHP, we chose to compare the confusion matrix
313 of the model against the one defined with thresholds of basal 2 ng/mL and 10 ng/mL.

314 As shown in Table II, despite the use of LC–MS/MS (more sensitive and more specific
315 than immunoassay), using a 17OHP cut-off value of 2 ng/mL (Table II), the test has a sensitivity
316 of 94.1% and a specificity of 92.6% for the identification of N21HOD cases. It corresponded
317 to one false negative and 25 false positives on the test set (PSL). Using a 17OHP cut-off value
318 of non-stimulated 10 ng/mL (Table II), the test showed a sensitivity of 29.4% and a specificity
319 of 100%.

320 Based on our data, the ROC analysis (Youden's index) suggested a basal 17OHP
321 measured by LC-MS/MS cutoff level of 1.76 ng/mL as the optimal cutoff. This cut-off provided

322 94.1% sensitivity, but yielded 36 false positives results, leading to a more widespread use of
323 genetic testing.

324 The use of the model (Table II), exhibiting 100% sensitivity and specificity of the
325 classical threshold, was more efficient than the use of the above-indicated thresholds.

326

327 *Blood sampling timing during the menstrual cycle phase does not impact the efficiency of our*
328 *model*

329 During the luteal phase of the menstrual cycle, there is a documented increase in serum
330 progesterone and 17OHP levels of ovarian origin (Young *et al.*, 2010). These hormonal changes
331 interfere with the diagnostic thresholds of 17OHP used. For this reason, it is recommended
332 (Young *et al.*, 2010) to collect samples for the determination of these steroids during the
333 follicular phase. However, identifying the phases of the cycle is not always easy in women with
334 oligomenorrhea. We therefore evaluate the impact of the luteal phase of the cycle on the
335 performance of our model.

336 In our set test (PSL), among 338 non-NC21OHD women, including controls and PCOS,
337 progesterone serum levels were increased: in 31, progesterone levels were above 10 ng/mL and
338 in 32, levels were between 5 ng/mL and 10 ng/mL, suggesting a blood sampling during the
339 luteal phase. Twenty-five out of 31 women with increased progesterone levels also had elevated
340 17OHP levels (above the 2 ng/mL threshold) resulting in misclassification (false positives)
341 using the above-indicated classic thresholds. Despite this confounding parameter, the model
342 succeeded in correctly classifying all these women as non-affected by NC21OHD (Table II).

343

344 **Discussion**

345 To our knowledge, this proof-of-concept study is the first attempt to use a relevant steroid
346 profile combined with a machine learning approach in order to build a model that reliably

347 distinguishes women with NC21OHD from controls as well as PCOS women without the use
348 of an ACTH stimulation test and regardless of the phase of the menstrual cycle.

349 We were confronted by a dual challenge: from a clinical point of view, NC21OHD and
350 PCOS patients presented a similar symptomatology (Papadakis *et al.*, 2019) and, from a
351 biological/biochemical point of view, it was necessary to determine and quantify relevant
352 biomarkers to allow a reliable identification of these patients.

353 A basal 17OHP level cut-off value of 2 ng/mL is recommended for screening (Carmina
354 *et al.*, 2017), although, ideally, the reference value should be established at each laboratory.
355 However, the election of a cut-off basal 17OHP value for the screening of 21-hydroxylase
356 deficient in NC21OHD remains a matter of debate. Indeed, the diagnosis of NC21OHD would
357 not have been made for 2–8% of the compound heterozygous patients carrying mild and severe
358 CYP21A2 mutations without a stimulation test by ACTH, despite a basal 17OHP value < 2
359 ng/mL in large cohort (Livadas *et al.*, 2015). Finally, sampling timing requirements (i.e., early
360 follicular phase of the menstrual cycle), particularly in PCOS/NC21OHD patients with irregular
361 cycles, make the procedure even more difficult in clinical practice. In this work, the patient
362 inclusion was designed to be as close as possible to the epidemiology of clinical
363 hyperandrogenism. The replication cohort (i.e., the cohort used to evaluate the performance of
364 the diagnostic strategy) was composed of 62% PCOS patients, 6% NC21OHD and 32%
365 controls.

366 We have shown here that isolated basal levels of serum 17OHP cut-offs, even quantified
367 using LC–MS/MS, do not allow sufficient sensibility and specificity to effectively distinguish
368 NC21OHD from PCOS. Indeed, despite using the LC–MS/MS method, which is more sensitive
369 and specific than the conventional immunological assays (Wong *et al.*, 1992; Minutti *et al.*,
370 2004; Fiet *et al.*, 2017), the mere application of the classic thresholds of 2 or 10 ng/mL for basal

371 17OHP on the patients of the replication cohort yields many misclassifications (1 false negative
372 and 25 false positives for 2ng/mL basal 17OHP cut-off).

373 Indeed, in the literature, studies have focused on establishing new cutoffs : among these,
374 Escobar *et al.*, using a RIA assay and on a study including 6 NC21OHD patients proposed a
375 threshold of 1.7 ng/mL. More recently, Grethe A. Ueland *et al.*, based on LC-MS/MS assay
376 which included 8 genetically confirmed NC21OHD, suggested 1.5 ng/mL (Escobar-Morreale
377 *et al.*, 2008; Ueland *et al.*, 2022). By applying these cut-offs to our patients below 2 ng/mL,
378 false positives increased to 45 and 38 for the 1.5 and 1.7 cut-offs respectively. In accordance
379 with these two studies, the cut-off of 1.76 based on our cohort of 355 individuals, confirmed
380 this high rate of false positives (n=36), rendering circulating basal 17OHP as a single biomarker
381 unsuitable for diagnosis. Furthermore, the false-negative patient remained so, independently of
382 the chosen cut-off point. These findings strongly support the multivariate, scoring approach,
383 which would change the semiology by using a more complete panel of the steroid biosynthetic
384 pathway rather than a single LC-MS/MS 17OHP concentration.

385 The combination of a 15-steroid profile with adapted data analysis maximizes the
386 classification performance to exclude the diagnosis of NC21OHD in the context of
387 hyperandrogenism.

388 Turcu *et al.* recently proposed a logistic regression (LR) restricted to a small set of 3
389 circulating steroids (17OHP, 21DF, corticosterone) measured by LCMS to discriminate
390 between patients with NC21OHD and unaffected individuals (Turcu *et al.*, 2020a). This
391 preliminary monocentric study exhibited an interesting performance to distinguish healthy
392 women from NC21OHD. However, it did not evaluate the ability of this approach to
393 differentiate NC21OHD from a larger representative cohort of women with PCOS. In addition,
394 the absence of an external test set did not fully allow an unbiased evaluation of the performance
395 of the logistic regression. Considering the vast heterogeneity of PCOS clinical expression and

396 its epidemiology, the inclusion by Turcu et al. of only 27 patients with PCOS in the group of
397 “unaffected individuals” may have been biased and not representative of the well-recognized
398 wide hormonal spectrum observed in women with PCOS. However, this strategy (multivariate
399 vs 17OHP only) is mathematically relevant because it reduces the risk of spurious results (type
400 I error/false positive and type II error/false negative) (Wold *et al.*, 2001; Eriksson *et al.*, 2013);
401 however, one may argue that in Turcu et al., the authors focused only on few pre-selected
402 steroids. In our model, we used a less restricted strategy including 15 steroids to more
403 comprehensively map the steroidal pathway, incorporating steroids that are predominantly—if
404 not exclusively—of adrenal origin (21DB, 21DF, and 11 β OH androstenedione). In our
405 experience, these steroids are particularly adapted to compare the level hormone synthesis by
406 each gland, distinguishing adrenal from ovarian disease (Fiet *et al.*, 2017; Pretorius *et al.*, 2017;
407 Turcu *et al.*, 2020b; Agnani *et al.*, 2022).

408 Mathematically, a set of variables is considered to be collinear if there are one or more
409 relations or associations between those variables. These 15 steroids are interrelated in their
410 biosynthetic pathway, resulting in a high degree of collinearity. Consequently, and based on
411 our previous work (Bachelot *et al.*, 2021; Agnani *et al.*, 2022), we used the orthogonal partial
412 least squares approach (OPLS): the latter is less affected by multicollinearity than logistic
413 regression because its components are uncorrelated (i.e., orthogonal) (Wold *et al.*, 2001; Pérez-
414 Enciso and Tenenhaus, 2003; Kjeldahl and Bro, 2010). On a larger set of patients (n = 366),
415 our model allowed us to successfully exclude all (100%) patients with NC21OHD from the
416 independent test cohort (including PCOS, controls and NC21OHD). Practically, we could
417 reduce the dimension of our system by considering fewer variables (e.g., the four most
418 discriminating steroids: 21DF, 21DB, 11BOHA, 17OHP) for an equally efficient
419 NC21OHD/PCOS differential diagnosis. From a practical point of view, it must be stressed that

420 a quantitative profiling by LC–MS/MS of the 15 steroids in a 10-minute chromatographic run
421 is not a limiting factor and indeed facilitates its real-life clinical application (Fiet *et al.*, 2017).

422 Beyond the diagnostic tool, our model including 15 circulating steroids provides new
423 insights into biochemical alterations in the context of hyperandrogenism and its debated
424 potential ovarian/adrenal origins. The selection and ranking of variables by the algorithm are
425 understandable given the previous results obtained on each of these molecular species, which
426 are often considered individually. In our model, six steroids were found to be particularly
427 significant (with VIP values >1); three metabolites are ranked with higher discriminative power
428 than 17OHP (21DF, 21DB, and 11BOHA) and three other steroids are ranked lower
429 (androstenedione, corticosterone, and testosterone).

430 For years, 17OHP has been used—and is still used by many practitioners—to diagnose
431 NC21OHD (Carmina *et al.*, 2017). However, as stated before we have shown here the interest
432 of 21-deoxysteroid metabolites in the diagnosis of NC21OHD. Indeed, 21DF and 21DB—being
433 strictly adrenal metabolites—may represent more specific biomarkers than 17OHP, which has
434 a double ovarian and adrenal origin (Turcu *et al.*, 2014). The exclusively adrenal origin of these
435 steroids is likely an explanation for the significant improvement of the diagnostic performance
436 of our model compared to the isolated use of a simple 17OHP. In women with PCOS, the main
437 source of androstenedione is in the ovary, and this androgen is not submitted to the direct action
438 of adrenal 11-hydroxylase; consequently, the 11 β hydroxyandrostenedione serum
439 concentration found was similar to that in normal subjects, and the ratio of androstenedione to
440 11-OHA was higher than in normal women (Polson *et al.*, 1988).

441 Because of the ovarian and adrenal origin, basal 17OHP should be measured during the
442 follicular phase to detect NC21OHD. Indeed, an ovarian surge of 17OHP occurs until 3 ng/mL
443 during the luteal phase of the menstrual cycle. Moreover, slightly elevated 17OHP has been
444 reported in women with PCOS (Pall *et al.*, 2010). The inclusion of several patients (n = 31,

445 9.2%) in our series who were in the luteal phase (i.e., with serum progesterone level above 10
446 ng/mL) suggests that our model is accurate independently of the timing of blood sampling.

447 Finally, in our study, using all 15 steroids, the model selected and ranked 11 beta-
448 hydroxyandrostenedione as one of the most discriminating variables for isolating the
449 NC21OHD population from the rest of the PCOS and controls without a priori knowledge.
450 Concentrations of 11 beta-hydroxyandrostenedione were found to be higher in NC21OHD than
451 in PCOS and controls. The choice to measure 11OH-androstenedione to distinguish 21-
452 hydroxylase deficiency from PCOS patients appears reasonable and is indicated to add
453 diagnostic value when corroborated with the 21 DF, 21 DB and the rest of the panel.

454 This study has some potential limitations to consider, including the use of a restricted
455 and fully prospective cohort. The second limitation is an analytical issue because not all
456 laboratories, especially in France, are equipped with mass spectrometers able to routinely
457 measure this panel of 15 steroids. Finally, the robustness of our model needs to be established
458 with a larger prospective study for definitive validation in clinical practice. The definitive
459 confirmation of our results will also allow us to reserve expensive and complex genetic analyses
460 of CYP21A2 for patients with a high probability of NC21OHD. An open question that will also
461 require further studies is the ability of our approach to distinguish NC21OHD from other causes
462 of chronic anovulation, such as hyperprolactinemia and Cushing's syndrome, which are also
463 differential diagnoses of PCOS (Pugeat *et al.*, 2010).

464 The main perspective of this work would be to overcome the requirement to interrupt
465 hormonal contraceptive treatment. In fact, the diagnosis should usually be made at least 3
466 months after stopping all hormonal treatment (e.g., corticosteroids, hormonal contraception);
467 this can also be inconvenient for the patient.

468

469 **Conclusion**

470 The aim of this work was to develop a unique sampling methodology to serve as an alternative
471 to the dynamic ACTH test to exclude the diagnosis of NC21OHD in the presence of a clinical
472 hyperandrogenic presentation at any time of the menstrual cycle, facilitating easier management
473 and greater patient comfort. This methodology was based on a combination of machine learning
474 with metabolic profiling of 15 molecular species (steroids) in LC–MS/MS at the base level.
475 This first proof of concept is promising; however, it will require further evaluation in real
476 clinical situations in the future.

477 **Author's roles:** G.B. performed lipidomics analysis, statistical modeling analysis, drafted the
478 manuscript, participated in the study conception and design, interpretation of data. M.B, J.E.S,
479 D.F, S.T, C.D., P.K, M.H, J.F, Y.LB participated in acquisition of data, patients care, and in
480 critical revision of the manuscript for intellectual content. J.Y, A.B supervised the study,
481 participated in study conception and design, drafting the manuscript and in critical revision of
482 the manuscript for intellectual content. A.L. supervised the study, lipidomics analysis, statistical
483 modeling analysis, participated in study conception and design, interpretation of data and
484 drafting the manuscript.

485 **Acknowledgements:** We thank Nicolas Helin and Bettina Ribault for their technical support.

486 **Funding:** This study was supported by Agence Française de Lutte contre le dopage and DIM
487 Région Ile de France; This study was supported by the French institutional PHRC 2010-
488 AOR10032 funding source and APHP).

489 **Conflict of interest:** All authors declare no competing financial interests.

490 **Data availability statement :** The data underlying this article will be shared on reasonable
491 request to the corresponding author.

492

493 The ethics committee (“Comité de Protection des Personnes”) approved the study.

494

495

496

497 **Reference list**

- 498 Admoni O, Israel S, Lavi I, Gur M, Tenenbaum-Rakover Y. Hyperandrogenism in carriers of
499 CYP21 mutations: the role of genotype. *Clin Endocrinol (Oxf)* 2006;**64**:645–651.
- 500 Agnani H, Bachelot G, Eguether T, Ribault B, Fiet J, Le Bouc Y, Netchine I, Houang M,
501 Lamazière A. A proof of concept of a machine learning algorithm to predict late-onset
502 21-hydroxylase deficiency in children with premature pubic hair. *J Steroid Biochem*
503 *Mol Biol* 2022;**220**:106085.
- 504 Azziz R, Hincapie LA, Knochenhauer ES, Dewailly D, Fox L, Boots LR. Screening for 21-
505 hydroxylase-deficient nonclassic adrenal hyperplasia among hyperandrogenic women:
506 a prospective study. *Fertil Steril* 1999;**72**:915–925.
- 507 Bachelot G, Lévy R, Bachelot A, Faure C, Czernichow S, Dupont C, Lamazière A, Alifert
508 Collaborative Group. Proof of concept and development of a couple-based machine
509 learning model to stratify infertile patients with idiopathic infertility. *Sci Rep*
510 2021;**11**:24003.
- 511 Bello R, Lebenthal Y, Lazar L, Shalitin S, Tenenbaum A, Phillip M, Vries L de. Basal 17-
512 hydroxyprogesterone cannot accurately predict nonclassical congenital adrenal
513 hyperplasia in children and adolescents. *Acta Paediatr* 2017;**106**:155–160.
- 514 Bidet M, Bellanné-Chantelot C, Galand-Portier M-B, Golmard J-L, Tardy V, Morel Y, Clauin
515 S, Coussieu C, Boudou P, Mowzowicz I, *et al.* Fertility in women with nonclassical
516 congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *J Clin Endocrinol*
517 *Metab* 2010;**95**:1182–1190.
- 518 Bidet M, Bellanné-Chantelot C, Galand-Portier M-B, Tardy V, Billaud L, Laborde K,
519 Coussieu C, Morel Y, Vaury C, Golmard J-L, *et al.* Clinical and molecular
520 characterization of a cohort of 161 unrelated women with nonclassical congenital

- 521 adrenal hyperplasia due to 21-hydroxylase deficiency and 330 family members. *J Clin*
522 *Endocrinol Metab* 2009;**94**:1570–1578.
- 523 Carmina E, Dewailly D, Escobar-Morreale HF, Kelestimur F, Moran C, Oberfield S, Witchel
524 SF, Azziz R. Non-classic congenital adrenal hyperplasia due to 21-hydroxylase
525 deficiency revisited: an update with a special focus on adolescent and adult women.
526 *Hum Reprod Update* 2017;**23**:580–599.
- 527 Dewailly D, Vantyghem-Haudiquet MC, Sainsard C, Buvat J, Cappoen JP, Ardaens K,
528 Racadot A, Lefebvre J, Fossati P. Clinical and biological phenotypes in late-onset 21-
529 hydroxylase deficiency. *J Clin Endocrinol Metab* 1986;**63**:418–423.
- 530 El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. *Lancet* 2017;**390**:2194–
531 2210.
- 532 Eriksson L, Byrne T, Johansson E, Trygg J, Vikström C. *Multi- and Megavariate Data*
533 *Analysis Basic Principles and Applications* [Internet]. 2013; Umetrics
534 Academy Available from: <https://books.google.fr/books?id=58qLBQAAQBAJ>.
- 535 Escobar-Morreale HF, Sanchón R, San Millán JL. A Prospective Study of the Prevalence of
536 Nonclassical Congenital Adrenal Hyperplasia among Women Presenting with
537 Hyperandrogenic Symptoms and Signs. *The Journal of Clinical Endocrinology &*
538 *Metabolism* 2008;**93**:527–533.
- 539 Fiet J, Le Bouc Y, Guéchet J, Hélin N, Maubert MA, Farabos D, Lamazière A. A Liquid
540 Chromatography/Tandem Mass Spectrometry Profile of 16 Serum Steroids, Including
541 21-Deoxycortisol and 21-Deoxycorticosterone, for Management of Congenital
542 Adrenal Hyperplasia. *J Endocr Soc* 2017;**1**:186–201.
- 543 Kamenický P, Blanchard A, Lamaziere A, Piedvache C, Donadille B, Duranteau L, Bry H,
544 Gautier J-F, Salenave S, Raffin-Sanson M-L, *et al*. Cortisol and Aldosterone

- 545 Responses to Hypoglycemia and Na Depletion in Women With Non-Classic 21-
546 Hydroxylase Deficiency. *J Clin Endocrinol Metab* 2020;**105**:dgz005.
- 547 Kjeldahl K, Bro R. Some common misunderstandings in chemometrics. *J Chemometrics*
548 2010;**24**:558–564.
- 549 Kuttann F, Couillin P, Girard F, Billaud L, Vincens M, Boucekkine C, Thalabard JC,
550 Maudelonde T, Spritzer P, Mowszowicz I. Late-onset adrenal hyperplasia in
551 hirsutism. *N Engl J Med* 1985;**313**:224–231.
- 552 Livadas S, Dracopoulou M, Dastamani A, Sertedaki A, Maniati-Christidi M, Magiakou A-M,
553 Kanaka-Gantenbein C, Chrousos GP, Dacou-Voutetakis C. The spectrum of clinical,
554 hormonal and molecular findings in 280 individuals with nonclassical congenital
555 adrenal hyperplasia caused by mutations of the CYP21A2 gene. *Clin Endocrinol (Oxf)*
556 2015;**82**:543–549.
- 557 Makela SK, Ellis G. Nonspecificity of a direct 17 alpha-hydroxyprogesterone
558 radioimmunoassay kit when used with samples from neonates. *Clin Chem*
559 1988;**34**:2070–2075.
- 560 Matheson E, Bain J. Hirsutism in Women. *Am Fam Physician* 2019;**100**:168–175.
- 561 Minutti CZ, Lacey JM, Magera MJ, Hahn SH, McCann M, Schulze A, Cheillan D, Dorche C,
562 Chace DH, Lymp JF, *et al*. Steroid profiling by tandem mass spectrometry improves
563 the positive predictive value of newborn screening for congenital adrenal hyperplasia.
564 *J Clin Endocrinol Metab* 2004;**89**:3687–3693.
- 565 New MI. Extensive clinical experience: nonclassical 21-hydroxylase deficiency. *J Clin*
566 *Endocrinol Metab* 2006;**91**:4205–4214.
- 567 New MI, Abraham M, Gonzalez B, Dumic M, Razzaghy-Azar M, Chitayat D, Sun L, Zaidi
568 M, Wilson RC, Yuen T. Genotype-phenotype correlation in 1,507 families with

- 569 congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. *Proc Natl Acad*
570 *Sci U S A* 2013;**110**:2611–2616.
- 571 Pall M, Azziz R, Beires J, Pignatelli D. The phenotype of hirsute women: a comparison of
572 polycystic ovary syndrome and 21-hydroxylase-deficient nonclassic adrenal
573 hyperplasia. *Fertility and Sterility* 2010;**94**:684–689.
- 574 Papadakis G, Kandaraki EA, Tseniklidi E, Papalou O, Diamanti-Kandarakis E. Polycystic
575 Ovary Syndrome and NC-CAH: Distinct Characteristics and Common Findings. A
576 Systematic Review. *Front Endocrinol* 2019;**10**:388.
- 577 Pérez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial
578 least squares discriminant analysis (PLS-DA) approach. *Hum Genet* 2003;**112**:581–
579 592.
- 580 Polson DW, Reed MJ, Franks S, Scanlon MJ, James VHT. Serum 11 β -
581 Hydroxyandrostenedione as an Indicator of the Source of Excess Androgen
582 Production in Women with Polycystic Ovaries*. *The Journal of Clinical*
583 *Endocrinology & Metabolism* 1988;**66**:946–950.
- 584 Pretorius E, Arlt W, Storbeck K-H. A new dawn for androgens: Novel lessons from 11-
585 oxygenated C19 steroids. *Mol Cell Endocrinol* 2017;**441**:76–85.
- 586 Pugeat M, Déchaud H, Raverot V, Denuzière A, Cohen R, Boudou P. Recommandations pour
587 l'exploration des hyperandrogénies. 2010;30.
- 588 Rosenbaum D, Gallo A, Lethielleux G, Bruckert E, Levy BI, Tanguy ML, Dulon J,
589 Dahmoune N, Salem JE, Bittar R, *et al.* Early central blood pressure elevation in adult
590 patients with 21-hydroxylase deficiency. *Journal of Hypertension* 2019;**37**:175–181.
- 591 Rosenfield RL, Ehrmann DA. The Pathogenesis of Polycystic Ovary Syndrome (PCOS): The
592 Hypothesis of PCOS as Functional Ovarian Hyperandrogenism Revisited. *Endocr Rev*
593 2016;**37**:467–520.

- 594 Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003
595 consensus on diagnostic criteria and long-term health risks related to polycystic ovary
596 syndrome (PCOS). *Hum Reprod* 2004;**19**:41–47.
- 597 Speiser PW, Dupont J, Zhu D, Serrat J, Buegeleisen M, Tusie-Luna MT, Lesser M, New MI,
598 White PC. Disease expression and molecular genotype in congenital adrenal
599 hyperplasia due to 21-hydroxylase deficiency. *J Clin Invest* 1992;**90**:584–595.
- 600 Speiser PW, Knochelhauer ES, Dewailly D, Fruzzetti F, Marcondes JA, Azziz R. A
601 multicenter study of women with nonclassical congenital adrenal hyperplasia:
602 relationship between genotype and phenotype. *Mol Genet Metab* 2000;**71**:527–534.
- 603 Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ,
604 International PCOS Network. Recommendations from the international evidence-
605 based guideline for the assessment and management of polycystic ovary syndrome.
606 *Hum Reprod* 2018;**33**:1602–1618.
- 607 Turcu A, Smith JM, Auchus R, Rainey WE. Adrenal androgens and androgen precursors-
608 definition, synthesis, regulation and physiologic actions. *Compr Physiol* 2014;**4**:1369–
609 1381.
- 610 Turcu AF, El-Maouche D, Zhao L, Nanba AT, Gaynor A, Veeraraghavan P, Auchus RJ,
611 Merke DP. Androgen excess and diagnostic steroid biomarkers for nonclassic 21-
612 hydroxylase deficiency without cosyntropin stimulation. *Eur J Endocrinol*
613 2020a;**183**:63–71.
- 614 Turcu AF, Wannachalee T, Tsodikov A, Nanba AT, Ren J, Shields JJ, O’Day PJ, Giacherio
615 D, Rainey WE, Auchus RJ. Comprehensive Analysis of Steroid Biomarkers for
616 Guiding Primary Aldosteronism Subtyping. *Hypertension* 2020b;**75**:183–192.
- 617 Ueland GÅ, Dahl SR, Methlie P, Hessen S, Husebye ES, Thorsby PM. Adrenal steroid
618 profiling as a diagnostic tool to differentiate polycystic ovary syndrome from

- 619 nonclassic congenital adrenal hyperplasia: pinpointing easy screening possibilities and
620 normal cutoff levels using liquid chromatography tandem mass spectrometry. *Fertility
621 and Sterility* 2022;**118**:384–391.
- 622 Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics.
623 *Chemometrics and Intelligent Laboratory Systems* 2001;**58**:109–130.
- 624 Wong T, Shackleton CH, Covey TR, Ellis G. Identification of the Steroids in Neonatal
625 Plasma that Interfere with 17 Alpha-Hydroxyprogesterone Radioimmunoassays.
626 *Clinical Chemistry* 1992;**38**:1830–1837.
- 627 Young J, Tardy V, Brac de la Perrière A, Bachelot A, Morel Y. Dépistage et prise en charge
628 du déficit en 21-hydroxylase à révélation tardive chez les femmes avec
629 hyperandrogénie. *Annales d'Endocrinologie* 2010;**71**:e15–e19.
- 630

631 **Figure and Table legends**

632

633 **Figure 1 : Study overview, step-by-step from steroidome quantitation to explainable**
634 **clinical decision**

635 Step 1 : 15 steroid species were measured in serum by liquid chromatography–mass
636 spectrometry (LC-MS/MS)

637 Step 2 : orthogonal partial least squares (OPLS-DA) was performed on the training set (KB).
638 The Y axis represents the decision boundary and allowed to define 2 areas : on the left
639 quadrants, non-classic 21OH deficiency (NC21OHD) patients with specific steroids signature,
640 and on the right quadrants controls with their associated steroids pattern.

641 Step 3 : Once the OPLS-DA model built, the model was then tested on an independent external
642 test set from PSL. After measuring their steroidome (similarly to step 1), the individuals of the
643 test cohort were put into the model. The goal was to distinguishes women with NC21OHD from
644 both normal and polycystic ovary syndrome (PCOS) women without the use of an ACTH
645 stimulation test.

646

647 **Figure 2 : Serum 17OHP levels before (A) after (B) ACTH stimulation in the two patient**
648 **groups, NC21OHD (n=36) and PCOS (n=266). p values were calculated using a Mann**
649 **Whitney test.**

650

651 **Figure 3 : Segregation of patients according to their basal steroid profiles.**

652 A. Plan construction and patient projection of the training set : 19 controls (green dots) and 19
653 non-classic 21OH deficiency (NC21OHD) (red dots). This plane represents the output of the
654 model, and each point a patient defined by its steroidome. The vertical axis separates the plane

655 into 2 areas: left area corresponding to steroidomes consistent with a diagnosis of NC21OHD,
656 and right area with steroidomes of healthy women.

657 B. Projection of the 355 women of the test set (black stars) on the plane created by the model,
658 according to their steroidome. At this stage, no diagnosis is considered. However, their location
659 on the plane (right or left side of the vertical axis) allows to see if their steroidome is compatible
660 or not with the diagnosis of 21OHD.

661 C. Projection of 17 NC21OHD patients (red stars) and 72 controls (green stars) from the test
662 set on the plan, following annotation in accordance with their actual diagnosis.

663 D. Projected confounding population of 266 polycystic ovary syndrome (PCOS) patients (blue
664 stars) after annotation based on their actual diagnosis.

665

666 **Figure 4 : Major contributing steroids to distinguish non-classic 21OH deficiency**
667 **(NC21OHD) from controls and polycystic ovary syndrome (PCOS) within the model**

668 The VIP (Variable Importance for the Projection) plot displays the VIP values as a column plot
669 with confidence intervals (95%) derived from jack-knifing (jack-knife uncertainty bars). VIP
670 plot summarized the importance or contribution of the steroids in the model. The most
671 segregative steroid species were those with VIP values larger than 1 including 21-deoxycortisol
672 (21-DF), 21-deoxycorticosterone (21-DB), 11 β -hydroxyandrostenedione (11 β -OHA), 17-OH-
673 progesterone (17OHP), androstenedione (A), corticosterone (B), and testosterone (T).

674

675 **Table I : Clinical and biological characteristics of polycystic ovary syndrome (PCOS) and**
676 **non-classic 21OH deficiency (NC21OHD) patients from test set**

677 § : 1 patient with unknown BMI, 4 patients with unknown FSH, 3 patients with unknown LH,
678 2 patients with unknown Estradiol, 1 patient with unknown 17 OHP one hour after ACTH

679 stimulation, 58 patients with unknown Follicular number right, 60 patients with unknown
680 Follicular number left
681 §§ : 2 patients with unknown FSH, LH and Estradiol, 1 patient with unknown 17OHP one hour
682 after ACTH stimulation, 6 patients with unknown Follicular number left and right
683 Data are expressed as median (interquartile range). Unpaired two-tailed Mann-Whitney U test.
684 *p < 0.05; **p < 0.001; n.s. not significant

685

686 **Table II : Comparison of the discriminating power of the basal [17OHP] threshold used**
687 **for screening (2ng/mL); the threshold used for diagnosis (10ng/mL) as well as the model**
688 **in identifying non-classic 21OH deficiency (N21HOD) cases.**

689 Contingency table using basal [17OHP] cut-off value of 2 ng/mL has a sensitivity of 94.1% and
690 a specificity of 92.6% ; Contingency table using basal [17OHP] cut-off value of 10ng/mL
691 showed a sensitivity of 29.4% and a specificity of 100%. Contingency table using the model
692 developed from the 15 steroids in prime rate. The model exhibiting sensitivity and specificity
693 of 100%.

694

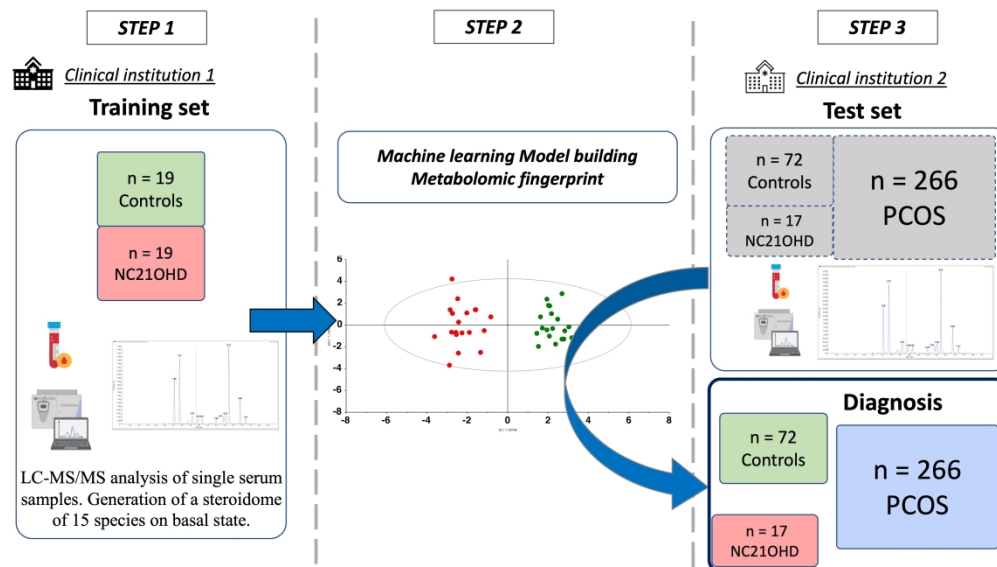


Figure 1 : Study overview, step-by-step from steroidome quantitation to explainable clinical decision
Step 1 : 15 steroid species were measured in serum by liquid chromatography–mass spectrometry (LC-MS/MS)

Step 2 : orthogonal partial least squares (OPLS-DA) was performed on the training set (KB). The Y axis represents the decision boundary and allowed to define 2 areas : on the left quadrants, non-classic 21OH deficiency (NC21OHD) patients with specific steroids signature, and on the right quadrants controls with their associated steroids pattern.

Step 3 : Once the OPLS-DA model built, the model was then tested on an independent external test set from PSL. After measuring their steroidome (similarly to step 1), the individuals of the test cohort were put into the model. The goal was to distinguishes women with NC21OHD from both normal and polycystic ovary syndrome (PCOS) women without the use of a ACTH stimulation test.

335x190mm (330 x 330 DPI)

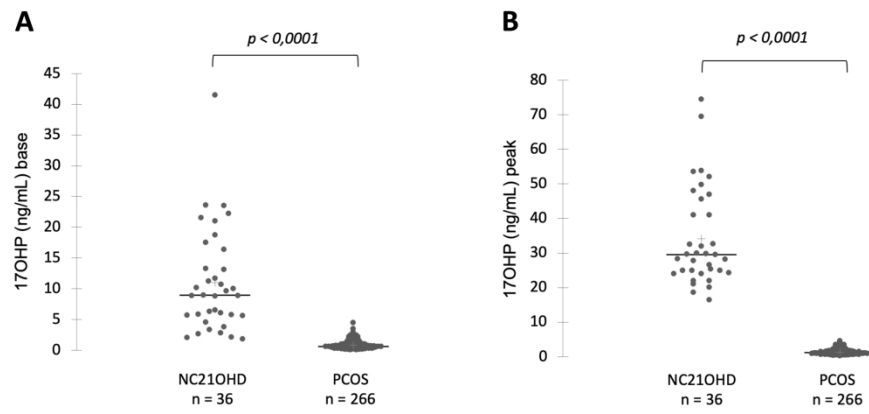


Figure 2 : Serum 17OHP levels before (A) after (B) ACTH stimulation in the two patient groups, NC21OHD (n=36) and PCOS (n=266). p values were calculated using a Mann Whitney test.

316x135mm (300 x 300 DPI)

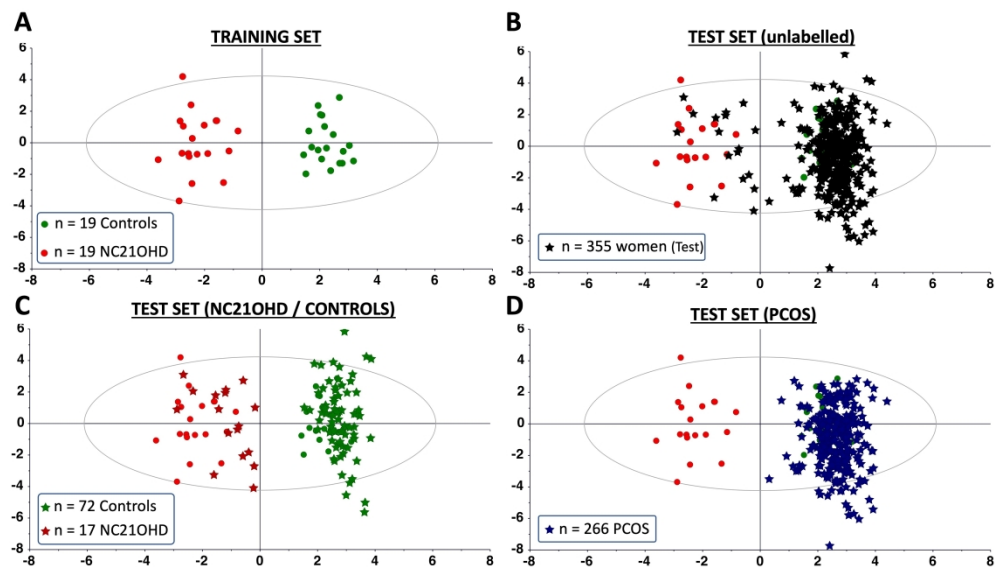


Figure 3 : Segregation of patients according to their basal steroid profiles.

- A. Plan construction and patient projection of the training set : 19 controls (green dots) and 19 non-classic 21OH deficiency (NC21OHD) (red dots). This plane represents the output of the model, and each point a patient defined by its steroidome. The vertical axis separates the plane into 2 areas: left area corresponding to steroidomes consistent with a diagnosis of NC21OHD, and right area with steroidomes of healthy women.
- B. Projection of the 355 women of the test set (black stars) on the plane created by the model, according to their steroidome. At this stage, no diagnosis is considered. However, their location on the plane (right or left side of the vertical axis) allows to see if their steroidome is compatible or not with the diagnosis of NC21OHD.
- C. Projection of 17 NC21OHD patients (red stars) and 72 controls (green stars) from the test set on the plan, following annotation in accordance with their actual diagnosis.
- D. Projected confounding population of 266 polycystic ovary syndrome (PCOS) patients (blue stars) after annotation based on their actual diagnosis.

335x194mm (330 x 330 DPI)

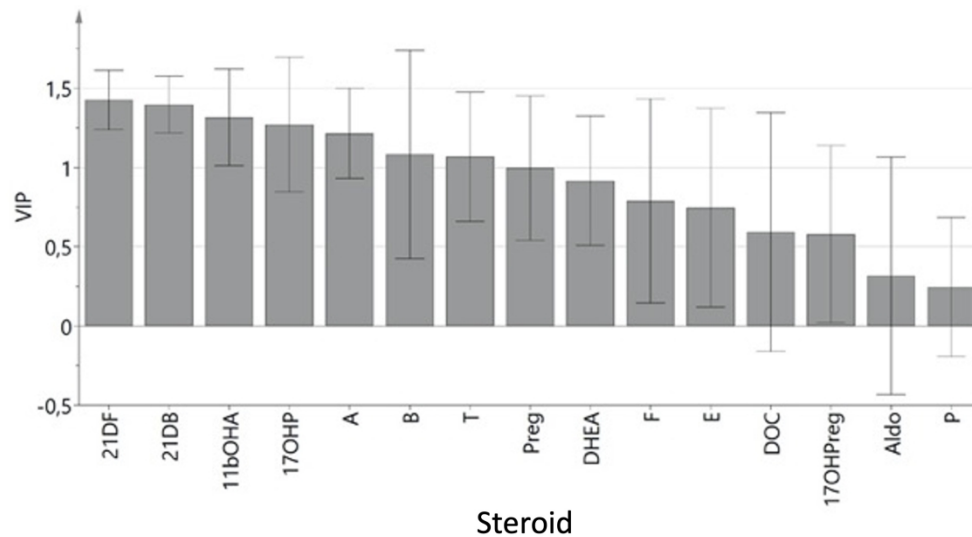


Figure 4 : Major contributing steroids to distinguish non-classic 21OH deficiency (NC21OHD) from controls and polycystic ovary syndrome (PCOS) within the model

The VIP (Variable Importance for the Projection) plot displays the VIP values as a column plot with confidence intervals (95%) derived from jack-knifing (jack-knife uncertainty bars). VIP plot summarized the importance or contribution of the steroids in the model. The most segregative steroid species were those with VIP values larger than 1 including 21DF, 21DB, 11BOHA, 17OHP, A, B, and T.

(21-deoxycortisol (21-DF), 11 β -hydroxyandrostenedione (11 β -OHA), androstenedione (D4A), testosterone (T), 17-OH-progesterone (17OHP), 17-OH-pregnenolone (17-OHPreg), dehydroepiandrosterone (DHEA), aldosterone (ALDO), cortisone (E), cortisol (F), corticosterone (B), 11-deoxycorticosterone (DOC), 21-deoxycorticosterone (21-DB), progesterone (P) and pregnenolone (Preg))

211x120mm (330 x 330 DPI)

	PCOS (n=266)§	NC21OHD (n=17)§§	P-value
<u>Baseline characteristics</u>			
Age (years)	27.8 (23.9-31.2)	30.8 (27.8-36.9)	*
BMI (kg/m ²)	25.6 (21.3-32.4)	26 (22.4-33.7)	n.s
Oligomenorrhea %	87.55	23.53	**
Hirsutism %	71.43	64.71	n.s
Acne %	45.49	29.41	n.s
Infertility %	15.04	17.65	n.s
Weight gain %	28.95	11.76	n.s
Baldness %	13.16	0	n.s
Follicle number right (n)	24 (19-32)	15 (9-22)	*
Follicle number left (n)	20 (16-30)	12 (6-13)	*
FSH (UI/L)	5.5 (4.2-6.4)	5.80 (4.0-7.2)	n.s
LH (UI/L)	11.2 (7.1-16.2)	6.7 (4.4-9.4)	*
Estradiol (pg/mL)	60.0 (45.0-100.3)	66.0 (49.0-102.3)	n.s
AMH (ng/mL)	7.7 (5.2 – 11.8)	5.3 (3.0 – 8.1)	**
<u>LC-MS/MS steroids</u>			
Basal 17-OH-progesterone (ng/mL)	0.60 (0.40-0.94)	6.07 (3.81-9.99)	**
17-OH-progesterone (ng/mL) after stimulation with ACTH	1.2 (0.9-1.6)	28.9 (25.3-45.4)	**
Basal 21-deoxycorticosterone (ng/mL)	10.00 (10.00-10.00)	15.64 (10.00-39.45)	**
Basal 21-deoxycortisol (ng/mL)	0.01 (0.01-0.01)	0.87 (0.37-2.11)	**
Basal 11 β -hydroxyandrostenedione (ng/mL)	1.12 (0.74-1.61)	2.64 (1.82-4.89)	**
Basal androstenedione (ng/mL)	1.93 (1.39-2.49)	2.53 (2.04-3.50)	**
Basal testosterone (ng/mL)	0.56 (0.41-0.74)	0.56 (0.42-0.77)	n.s
Basal progesterone (ng/mL)	0.12(0.08-0.25)	0.52(0.39-3.88)	**

Table I : Clinical and biological characteristics of polycystic ovary syndrome (PCOS) and non-classic 21OH deficiency (NC21OHD) patients from test set

§ : 1 patient with unknown BMI, 4 patients with unknown FSH, 3 patients with unknown LH, 2 patients with unknown Estradiol, 1 patient with unknown 17 OHP one hour after ACTH stimulation, 58 patients with unknown Follicular number right, 60 patients with unknown Follicular number left

§§ : 2 patients with unknown FSH, LH and Estradiol, 1 patient with unknown 17OHP one hour after ACTH stimulation, 6 patients with unknown Follicular number left and right
Data are expressed as median (interquartile range). Unpaired two-tailed Mann-Whitney U test. *p < 0.05; **p < 0.001; n.s. not significant

1

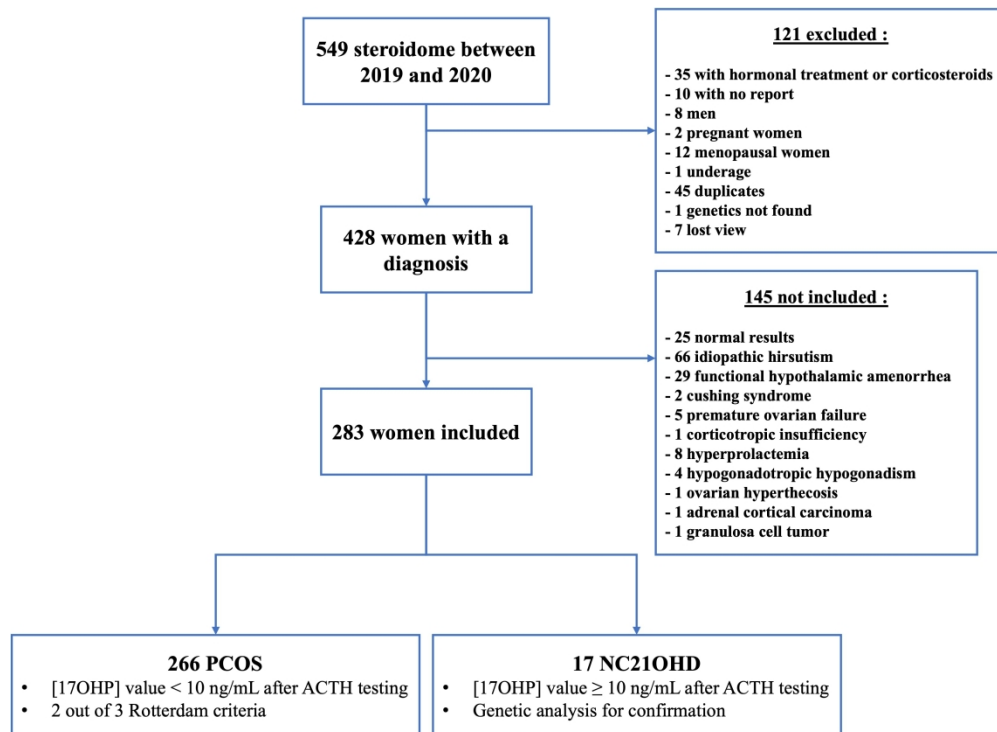
Screening value : 2ng/mL [17OHP]				
		Actual 17 NC21OHD		Total
		No	Yes	
Predicted NC21OHD	No	True Negative = 313	False Negative = 1	314
	Yes	False Positive = 25	True Positive = 16	41
Total		338	17	355
Diagnosis value : 10ng/mL [17OHP]				
		Actual 17 NC21OHD		Total
		No	Yes	
Predicted NC21OHD	No	True Negative = 338	False Negative = 12	350
	Yes	0	True Positive = 5	5
Total		338	17	355
Machine Learning model (15 steroids)				
		Actual 17 NC21OHD		Total
		No	Yes	
Predicted NC21OHD	No	True Negative = 338	0	338
	Yes	0	True Positive = 17	17
Total		338	17	355

2

3 **Table II : Comparison of the discriminating power of the basal [17OHP] threshold used for**
4 **screening (2ng/mL); the threshold used for diagnosis (10ng/mL) as well as the model in**
5 **identifying non-classic 21OH deficiency (N21HOD) cases.**

6 Contingency table using basal [17OHP] cut-off value of 2 ng/mL has a sensitivity of 94.1%
7 and a specificity of 92.6% ; Contingency table using basal [17OHP] cut-off value of 10ng/mL
8 showed a sensitivity of 29.4% and a specificity of 100%. Contingency table using the model

9 developed from the 15 steroids in prime rate. The model exhibiting sensitivity and specificity
10 of 100%.



Supplemental Figure 1 : Flow chart

261x191mm (330 x 330 DPI)

Supplementary Table SI : Age, serum 17OHP after ACTH test and CYP21A2 genotype (allele 1 and allele 2) for each patient; Training set (KB) n = 19 and Test set (PSL) n = 17

TRAINING SET				
Patient #	Age (yrs)	Serum 17OHP 1h after Synacthen® (ng/mL)	CYP21A2 genotype	
			Allele 1	Allele 2
1	36	32	p.V281L	p.V281L
2	19	25	p.V281L	p.M283V
3	29	22	p.V281L	p.V281L
4	35	22	p.V281L	p.V281L
5	40	47	p.V281L	p.I172N ; p.Q318X
6	37	32.6	p.V281L	p.V281L
7	28	69.5	p.R356W	p.G178R
8	38	48	p.V281L	p.I172N ; p.Q318X
9	40	25	p.V281L	c.293-13C>G/A
10	22	21	p.V281L	p.V281L
11	30	24.9	p.W21X	p.V281L
12	34	29.7	p.V281L	p.V281L
13	34	45.6	p.V281L	p.P453S
14	34	16.4	p.V281L	p.P453S
15	22	29.8	p.V281L	c.1-?_936+?del, p.0?
16	32	24	p.V281L	p.V281L
17	20	32.5	p.V281L	p.V281L
18	21	26.5	c.293-13C>G/A	p.V281L
19	20	74.5	p.P30L	p.R356W

TEST SET				
Patient #	Age (yrs)	Serum 17OHP 1h after Synacthen® (ng/mL)	CYP21A2 genotype	
			Allele 1	Allele 2
1	22	24.2	p.V281L	P30L
2	21	41.1	p.V281L	p.V281L
3	30	25.4	p.V281L	p.V281L
4	28	28.3	p.V281L	p.V281L
5	43	53.8	p.V281L	p.V281L
6	37	41	p.V281L	p.V281L
7	21	18.6	126C>T, 113G>A, 110T>C	126C>T, 13G>A

8	47	28.2	p.V281L	p.V281L
9	49	20	IVS2	P30L
10	28	24	p.V281L	p.V281L
11	36	30	p.V281L	p.V281L
12	26	29.5	p.V281L	p.V281L
13	30	27.8	p.V281L	p.V281L
14	35	49.8	intron 2 (IVS2-13A)	p.V281L
15	31	20.1	p.V281L	p.V281L
16	34	53.6	p.V281L	p.V281L
17	38	52.1	p.V281L	p.V281L

Supplementary Table SII : basal steroidomic profiles in the different control (KB vs. PSL) and NC21OHD (KB vs. PSL) groups.

(21-deoxycortisol (21-DF), 11 β -hydroxyandrostenedione (11 β -OHA), androstenedione (D4A), testosterone (T), 17-OH-progesterone (17OHP), 17-OH-pregnenolone (17-OHPreg), dehydroepiandrosterone (DHEA), aldosterone (ALDO), cortisone (E), cortisol (F), corticosterone (B), 11-deoxycorticosterone (DOC), 21-deoxycorticosterone (21-DB), progesterone (P) and pregnenolone (Preg))

Data are expressed as median (interquartile range). Unpaired two-tailed Mann-Whitney U test. *p < 0.05; **p < 0.001; n.s. not significant

Controls			
	Training set (n=19)	Test set (n=72)	p-value
17 OH progesterone	1.00 [0.51, 1.49]	0.93 [0.69, 1.16]	n.s.
21 deoxycortisol	0.02 [0.01, 0.03]	0.01 [0.01, 0.02]	**
progesterone	4.01 [1.13, 6.90]	3.15 [1.60, 4.70]	n.s.
21 deoxycorticosterone	5.00 [4.00, 6.00]	5.00 [4.00, 6.00]	/
11-deoxycorticosterone	0.04 [0.03, 0.05]	0.05 [0.04, 0.05]	n.s.
corticosterone	2.89 [2.01, 3.77]	3.94 [3.13, 4.76]	n.s.
aldosterone	0.12 [0.07, 0.17]	0.09 [0.07, 0.11]	n.s.
delta4-androstenedione	1.07 [0.85, 1.29]	1.46 [1.27, 1.65]	*
testosterone	0.25 [0.20, 0.29]	0.36 [0.31, 0.40]	**
11beta-hydroxy Delta4-androstenedione	0.85 [0.63, 1.08]	1.18 [1.01, 1.36]	*
dehydroepiandrosterone	4.96 [3.83, 6.10]	7.80 [6.41, 9.19]	*
17 OH pregnenolone	0.97 [0.61, 1.34]	2.55 [1.91, 3.19]	**
pregnenolone	1.87 [1.51, 2.22]	1.34 [1.13, 1.55]	*
cortisol	83.69 [70.97, 96.41]	104.74 [94.34, 115.14]	*
cortisone	18.41 [16.06, 20.76]	19.42 [18.41, 20.43]	n.s.
NC21OHD patients			
	Training set (n=19)	Test set (n=17)	p-value
17 OH progesterone	12.94 [6.27, 19.60]	12.14 [3.69, 20.58]	n.s.
21 deoxycortisol	2.04 [0.94, 3.14]	2.61 [1.18, 4.04]	n.s.
progesterone	2.13 [0.11, 4.16]	1.92 [-0.38, 4.23]	n.s.
21 deoxycorticosterone	50.21 [37.28, 63.14]	69.54 [-26.38, 165.45]	*
11-deoxycorticosterone	0.03 [0.02, 0.05]	0.03 [0.01, 0.05]	n.s.
corticosterone	1.23 [0.99, 1.47]	1.21 [0.81, 1.61]	n.s.
aldosterone	0.10 [0.07, 0.13]	0.12 [0.06, 0.19]	n.s.
delta4-androstenedione	2.86 [1.98, 3.75]	3.23 [2.56, 3.90]	n.s.
testosterone	0.53 [0.40, 0.66]	0.79 [0.65, 0.94]	*
11beta-hydroxy Delta4-androstenedione	3.57 [2.39, 4.75]	3.78 [2.31, 5.26]	n.s.
dehydroepiandrosterone	9.72 [6.41, 13.04]	10.73 [7.49, 13.98]	n.s.
17 OH pregnenolone	2.03 [0.40, 3.66]	3.69 [2.14, 5.25]	*

pregnenolone	3.61 [2.57, 4.65]	2.65 [0.83, 4.48]	*
cortisol	107.48 [99.06, 115.90]	123.05 [102.37, 143.72]	*
cortisone	23.20 [21.07, 25.33]	22.31 [20.05, 24.57]	n.s.