Effect of congenital adrenal hyperplasia treated by glucocorticoids on plasma metabolome: a machine-learning-based analysis

Lee Nguyen, Edi Prifti, Farid Ichou, Monique Leban, Christian Funck-Brentano, Philippe Touraine, Joe-Elie Salem, Anne Bachelot

To cite this version:

Lee Nguyen, Edi Prifti, Farid Ichou, Monique Leban, Christian Funck-Brentano, et al.. Effect of congenital adrenal hyperplasia treated by glucocorticoids on plasma metabolome: a machine-learning-based analysis. Scientific Reports, Nature Publishing Group, 2020, 10 (1), pp.8859. 10.1038/s41598-020-65897-y. inserm-02872323

HAL Id: inserm-02872323
https://www.hal.inserm.fr/inserm-02872323
Submitted on 17 Jun 2020

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.
Effect of congenital adrenal hyperplasia treated by glucocorticoids on plasma metabolome: a machine-learning-based analysis

Lee S. Nguyen1,2,7 ✉, Edi Prifti3,4,7, Farid Ichou5, Monique Leban1,6, Christian Funck-Brentano1,3, Philippe Touraine6, Joe-Elie Salem1,3 & Anne Bachelot6

Background. Congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency leads to impaired cortisol biosynthesis. Treatment includes glucocorticoid supplementation. We studied the specific metabolomics signatures in CAH patients using two different algorithms. Methods. In a case-control study of CAH patients matched on sex and age with healthy control subjects, two metabolomic analyses were performed: one using MetaboDiff, a validated differential metabolomic analysis tool and the other, using Predomics, a novel machine-learning algorithm. Results. 168 participants were included (84 CAH patients). There was no correlation between plasma cortisol levels during glucocorticoid supplementation and metabolites in CAH patients. Indoleamine 2,3-dioxygenase enzyme activity was correlated with ACTH (rho coefficient = −0.25, p-value = 0.02), in CAH patients but not in controls subjects. Overall, 33 metabolites were significantly altered in CAH patients. Main changes came from: purine and pyrimidine metabolites, branched aminoacids, tricarboxylic acid cycle metabolites and associated pathways (urea, glucose, pentose phosphates). MetaboDiff identified 2 modules that were significantly different between both groups: aminosugar metabolism and purine metabolism. Predomics found several interpretable models which accurately discriminated the two groups (accuracy of 0.86 and AUROC of 0.9). Conclusion. CAH patients and healthy control subjects exhibit significant differences in plasma metabolomes, which may be explained by glucocorticoid supplementation.

Congenital adrenal hyperplasia (CAH) is a group of autosomal recessive diseases due to enzyme deficiencies leading to impaired cortisol biosynthesis and deficiency of the 21-hydroxylase enzyme is its most common form. Main treatment relies on steroid replacement therapy including glucocorticoids (GC) (i.e. hydrocortisone) alone or in association with mineralocorticoids (i.e. fludrocortisone)1. Although effective, long-term GC replacement therapy is associated with multiple adverse effects. They include, but are not limited to the appearance of cardiovascular metabolic risk factors (hypertension, insulin-resistance, obesity and dyslipidemia), immunosuppression with infections, psychological disturbances and osteoporosis. Moreover, oral GC intake does not supplement adequately cortisol’s physiological circadian rhythm, and the dosage needed to suppress androgens is usually higher than that needed for substitution only2. These elements all contribute to increased health costs and decrease of quality of life and life expectancy2–6.
In CAH patients, dosage adjustment of GC routinely consists in assessing clinical symptoms and adrenal androgen levels7. Because of a narrow therapeutic index, metabolomic analyses may be suitable to better characterize the effects of this medication and prevent its adverse effects while conserving its benefits. Indeed, pharmacometabolomic studies quantify and analyze circulating metabolites following drug administration, specifically identifying activated pathways8,9. Previous metabolomic analyses of CAH patients showed that those requiring a higher dosage of GC presented a different metabolomic profile compared with those who required a lower dosage, with altered free fatty acids, bile acids, and amino acid metabolites pathways10. However, no comparison was performed with healthy control subjects, hence, impact of CAH on metabolome remains unclear.

In the present study, we aimed to compare the metabolomic profile of CAH patients treated with GC supplementation, to healthy control subjects, using a machine-learning-based algorithm (Predomics)11. A previously published package of metabolomic analyses (MetaboDiff) was used as quality-controller12.

Results
Clinical and hormonal differences between CAH and control. Table 1 summarizes clinical and characteristic parameters from this prospective study cohort (n = 168), which included 84 CAH patients and 84 healthy control subjects. Both groups were similar in age, gender proportion, systolic and diastolic blood pressure, history of cardiovascular disease (1.2% vs 7.1%, p = 0.054) and smoking status (active or past represented 27% in both groups). Healthy control subjects were taller than CAH patients (169 vs 163 cm, p < 0.001) and had lower body-mass index (23.07 vs 25.8, p < 0.001). Clinical and biological differences between gender and between groups were analyzed in a previous paper13. Among CAH patients, 42 (50%) had classic salt-wasting form, 16 (19.0%) with simple virilizing form and 26 (31.0%) had non-classical form.

Compared to healthy control subjects, CAH patients presented significantly higher levels of insulin levels, HOMAIR, 17-OH progesterone, progesterone, total testosterone, ACTH, renin and aldosterone (see Table 1). In both groups, androstenedione levels correlated with succinate, glutamate, fatty acid metabolites, urea cycle, arginine and proline metabolites. Androstenedione also correlated with leucine, isoleucine and valine metabolism pathway in healthy control subjects.

In CAH patients, renin levels correlated with glutamate and lysine metabolites, while it only correlated with tryptophan metabolites in healthy control subjects. There was no association between plasma cortisol levels and

|                       | Control (n = 84) | CAH (n = 84) |
|-----------------------|-----------------|--------------|
| Age (years)           | Median 27.2     | Median 23.4  |
|                       | Perc. 25 23.4   | Perc. 75 35.0|
|                       | Perc. 25 22.6  | Perc. 75 36.9|
| Height (cm)           | Median 169.0    | Median 163.0|
|                       | Perc. 25 163.0  | Perc. 75 177.0|
|                       | Perc. 25 157.0  | Perc. 75 170.0|
| Weight (kg)           | Median 67.8     | Median 60.5  |
|                       | Perc. 25 60.5   | Perc. 75 74.0|
|                       | Perc. 25 57.5   | Perc. 75 78.0|
| BMI                   | Median 23.07    | Median 20.88|
|                       | Perc. 25 20.88  | Perc. 75 23.61|
|                       | Perc. 25 25.8   | Perc. 75 22.19|
|                       | Perc. 25 29.03  | Perc. 75 0.01|
| Systolic blood pressure| Median 112.0    | Median 105.0|
|                       | Perc. 25 105.0  | Perc. 75 120.0|
|                       | Perc. 25 110.0  | Perc. 75 103.0|
|                       | Perc. 25 123.0  | Perc. 75 0.64|
| Diastolic blood pressure| Median 70.0     | Median 64.0  |
|                       | Perc. 25 64.0   | Perc. 75 74.0|
|                       | Perc. 25 68.0   | Perc. 75 59.0|
|                       | Perc. 25 75.0   | Perc. 75 0.16|
| Insulin               | Median 5.4      | Median 4.0   |
|                       | Perc. 25 4.0    | Perc. 75 8.2|
|                       | Perc. 25 8.8    | Perc. 75 5.8|
|                       | Perc. 25 13.0   | Perc. 75 <0.001|
| HOMAIR                | Median 1.0      | Median 0.8   |
|                       | Perc. 25 1.7    | Perc. 75 1.8|
|                       | Perc. 25 1.1    | Perc. 75 1.1|
|                       | Perc. 25 2.5    | Perc. 75 0.01|
| Estradiol (pg/mL)     | Median 46.5     | Median 27.7  |
|                       | Perc. 25 27.7   | Perc. 75 146.0|
|                       | Perc. 25 48.5   | Perc. 75 30.5|
|                       | Perc. 25 99.0   | Perc. 75 0.88|
| Progesterone (ng/mL)  | Median 0.8      | Median 0.4   |
|                       | Perc. 25 0.4    | Perc. 75 1.3|
|                       | Perc. 25 2.0    | Perc. 75 0.8|
|                       | Perc. 25 6.4    | Perc. 75 <0.001|
| FSH (IU/mL)           | Median 4.9      | Median 3.3   |
|                       | Perc. 25 7.6    | Perc. 75 4.9|
|                       | Perc. 25 2.9    | Perc. 75 6.5|
|                       | Perc. 25 0.88   | Perc. 75 0.88|
| LH (IU/mL)            | Median 5.7      | Median 4.3   |
|                       | Perc. 25 9.2    | Perc. 75 5.2|
|                       | Perc. 25 5.2    | Perc. 75 2.7|
|                       | Perc. 25 7.5    | Perc. 75 0.22|
| Androstenedione (ng/mL)| Median 2.3      | Median 1.6   |
|                       | Perc. 25 3.0    | Perc. 75 2.8|
|                       | Perc. 25 2.8    | Perc. 75 1.4|
|                       | Perc. 25 5.3    | Perc. 75 0.17|
| 17-OH-progesterone (ng/mL)| Median 1.6    | Median 0.9   |
|                       | Perc. 25 2.4    | Perc. 75 11.9|
|                       | Perc. 25 11.9   | Perc. 75 2.8|
|                       | Perc. 25 29.5   | Perc. 75 <0.001|
| Total testosterone (ng/mL)| Median 0.4    | Median 0.3   |
|                       | Perc. 25 0.3    | Perc. 75 4.5|
|                       | Perc. 25 0.8    | Perc. 75 0.3|
|                       | Perc. 25 2.9    | Perc. 75 0.03|
| SHBG (ng/mL)          | Median 56.1     | Median 35.3  |
|                       | Perc. 25 76.2   | Perc. 75 46.7|
|                       | Perc. 25 29.9   | Perc. 75 29.9|
|                       | Perc. 25 82.3   | Perc. 75 0.28|
| ACTH (pg/mL)          | Median 19.5     | Median 12.4  |
|                       | Perc. 25 29.9   | Perc. 75 31.3|
|                       | Perc. 25 10.3   | Perc. 75 67.2|
|                       | Perc. 25 0.09   | Perc. 75 0.09|
| Renin (pg/mL)         | Median 12.8     | Median 8.2   |
|                       | Perc. 25 17.0   | Perc. 75 19.5|
|                       | Perc. 25 12.1   | Perc. 75 41.7|
|                       | Perc. 25 0.01   | Perc. 75 0.01|
| Aldosterone (pg/mL)   | Median 107.0    | Median 73.1  |
|                       | Perc. 25 148.0  | Perc. 75 137.0|
|                       | Perc. 25 114.0  | Perc. 75 264.0|
|                       | Perc. 25 0.09   | Perc. 75 0.09|
| Cortisol (µg/L)       | Median 110      | Median 77.5  |
|                       | Perc. 25 140    | Perc. 75 94.85|
|                       | Perc. 25 35.65  | Perc. 75 144|
|                       | Perc. 25 0.49   | Perc. 75 0.49|

Table 1. Clinical and hormonal features of control subjects and CAH patients. Abbreviations: ACTH: AdrenoCorticoTrophic Hormone; BMI: body-mass index; FSH: Follicle Stimulating Hormone; HOMAIR: Homeostasic model assessment of insulin resistance; LH: luteinizing hormone; Perc: percentile; SHBG: Sex hormone-binding globulin; 17-OH: 17α-Hydroxyprogesterone.
metabolic pathways in CAH patients, whereas in control patients, cortisol correlated with lysine, leucine, isoleucine, valine, steroid, nicotinate and nicotinamide metabolites.

Finally, in CAH patients, indoleamine 2,3-dioxygenase (IDO) activity correlated with ACTH (\( \rho = -0.25, p = 0.02 \)), whereas it was not in control participants. The association between IDO and ACTH in CAH patients was independent from GC supplementation (\( p < 0.001 \)), after adjustment in multivariable analysis using global linear modeling.

**Metabolic differences between CAH and control patients.** Differences in metabolite abundance between CAH and control patients are presented in Fig. 1 and Supplementary Table A. Overall, between the two groups, 50 metabolites differed significantly (denoted with a ‘+’ symbol in Fig. 1), 33 of which, resisted adjustment for multiple testing (denoted with a ‘#’ symbol in Fig. 1). Main pathways identified included: purine and pyrimidine metabolism, branched aminoacid metabolism, tricarboxylic acid (TCA) cycle and associated pathways (urea, glucose, pentose phosphates).

The comparative metabolomics package “MetaboDiff” showed significant differences between the two groups and isolated two modules: module 13 | Aminosugar metabolism and module 4 | Purine metabolism, (Hypo) Xanthine/Inosine containing (see Fig. 2).

Similarly, the machine-learning based approach identified ternary (Ter) models combining several metabolites that accurately classify CAH and control patients. Metabolites which featured the higher importance in these models included (decreasing order, with featuring importance greater than 2.5): lactose/sucrose, diethanolamine, guanosine, xanthine, creatine, 3-hydroxy-3-methylglutarate, inosine, hydroxybutanoate, hypoxanthine and hexosamine (see Fig. 3).
Moreover, two models illustrating the machine-learning approach were manually generated. A first Ter model combining 4 metabolites (guanosine, inosine, xanthine and creatine) reached an empirical accuracy of 0.82 and area under receiver operator characteristics curve (AUROC) of 0.86 and a second Ter model combining 4 metabolites (guanosine, inosine, xanthine and 3-hydroxy-3-methylglutarate), empirical accuracy was 0.81 and AUROC 0.85.

For comparison purpose, the best automatically generated 4-features Ter model found a combination of xanthine, hydroxybutanoate isomer, n-acetyl-beta-alanine and hexosamine. Its empirical accuracy was 0.86 and AUROC 0.9 (estimated generalized accuracy of the algorithm was 0.74 and AUROC was 0.86).

**Discussion**

The main findings of this study are: (i) CAH patients treated with GC express a specific metabolomic signature allowing for metabolomic analyses to accurately discriminate them from healthy control subjects matched on gender and age and (ii) the machine-learning-based approach, **Predomics**, finds concordant results with the differential metabolomics analysis tool **MetaboDiff**, and allows to generate accurate and interpretable models.

Metabolomics analyses showed that CAH patients exhibited a specific metabolic signature. Main altered pathways were those influencing the amount of purine and pyrimidine metabolites, branched amino acids, tricarboxylic acid (TCA) cycle metabolites (urea, glucose, pentose phosphates). The isolated use of GC in healthy control subjects has previously been associated with metabolomics alterations\(^1\). These alterations were important even for small dosage of GC supplementation and redisposed all the main energetic pathways, including glycolysis, TCA cycle, urea cycle and their connection with fatty acid and amino acids. In the present work, most differences between CAH patients and control subjects lied on these same pathways. A plausible explanation may be that CAH patients may not exhibit other metabolomics alterations than those related to treatment. With similar cortisol levels in both groups, and without any significant correlation between cortisol and metabolic pathway; these findings support the fact that cortisol is not an adequate marker of disease control in this well-supplemented population. Moreover, diurnal cortisol intra-individual variations is well-known factor difficult to predict or correlate.

In this cohort, ACTH was associated with the IDO enzyme activity. IDO catalyzes the reaction degrading tryptophan to kynurenine, passing the rate-limiting step in the pathway which ends to nicotinamide adenine dinucleotide (NAD) biosynthesis\(^15,16\). ACTH is overexpressed by the pituitary gland in uncontrolled CAH patients, due to a lack of negative feedback\(^1\). Incidentally, in CAH patients, the most common mutation results in a 21-hydroxylase deficiency, which in humans, may be mimicked by a microsomal reduced NAD phosphate-dependent cytochrome p450 enzyme (POR) deficiency\(^17\). While POR deficiency also involves several other metabolic pathways, the association found between ACTH and IDO enzyme activity may explain how NAD pathway is altered in both diseases. Remarkably, the association remained statistically significant after adjusting for GC supplementation in multivariable analysis.

All metabolites found by the machine-learning based approach **Predomics**, to discriminate CAH patients from healthy control subjects, were related to pathways known to be influenced by GC treatment. Specifically, purine metabolism (containing xanthine, hypoxanthine, guanosine, inosine) highly discriminated CAH patients from healthy control subjects; followed by aminosugar metabolism (lactose, sucrose, hexosamine), branched aminoacid metabolites (3-hydroxy-3-methylbutyrate) and creatine, mevalonate and ketones metabolism, all which branch into the citrate cycle. Combining purine metabolites to another identified metabolite allowed to manually generate simple models with only four features, which showed high accuracy and discrimination, as compared to automatically selected models.

**Limitations.** We acknowledge several limitations in this study. Although the number of patients was relatively high for this rare disease, the sample size still was not very large and results might differ in a larger population of CAH patients. Second, correlations were moderate which may be due to confounding factors altering...
levels of circulating metabolites. The external validity of these results relies in part, on the fact that they comfort biochemical mechanisms proven experimentally decades ago for some of them (i.e. glucocorticoid-induced purine metabolism)\(^\text{18}\), or in the last five years, using advanced metabolomic tools, for others (i.e. metabolomic profiling of healthy subjects supplemented in GC)\(^\text{14}\). Regarding statistical analyses, we performed state-of-art data-processing techniques, p-value correction methods using Bonferroni-Hochberg procedure and stringent data filtering; we also omitted variables that were associated with medication intake to avoid false-positives and applied non-parametric statistics which are robust to normality distributions but at the expense of power. The machine-learning-based approach used here minimizes the sum of a cost function and regularizes L1 norm for sparsity, under a constraint on the unary value of the linear model that predicts classes. Besides being relatively simple and interpretable, the generated models are robust to generalization and offer importance scoring of the features.

Conclusion

CAH patients and healthy control subjects display significantly different plasma metabolomes signatures. Main differences may be explained by GC supplementation of CAH patients only. Predomics identified such simple signatures allowing for accurate discrimination between CAH and healthy control subjects.

Methods

Study design. This study is ancillary to the Cardiovascular Risk Profile in Patients With Congenital Adrenal Hyperplasia (CARDIOHCS) study (clinicaltrials.gov identifier NCT01807364), a multicenter prospective observational case-control study comparing early cardiovascular damage in adult men and women with CAH due to 21α-hydroxylase deficiency and healthy control subjects\(^\text{13,19,20}\). All participants provided written informed consent to participate, and the study was approved by each hospital ethics committee including Pitie-Salpetriere University Hospital.

Study population. Patients with CAH and healthy subjects were all assessed at the Clinical Investigation Center Paris-Est (CIC-1901, Pitie-Salpetriere University Hospital, Paris, France). Eighty-four young adults (58

Figure 3. Feature importance of metabolites in models created by Predomics. Left: The feature importance (mean decrease accuracy) of the Ter models. Features (rows) are ordered by the average MDA in the three experiments. Right: boxplots indicating the distribution of the same metabolites in the two study groups. The blue color indicates enrichment in the CAH group while red in the controls (i.e. in the boxplots, lactose is enriched in the control group). On the left panel, the colours indicate the same concepts also associated with the sign of the features in the TER model (-1 is blue, and 1 is red).
women and 26 men) with CAH and 84 control subjects matched for sex, age, and smoking status were prospectively included in the CARDIOHCS study between May 2011 and August 2015\textsuperscript{13,16,20}. The adults with CAH included in the study, were diagnosed during childhood and confirmed by genetic testing of 21-\(\alpha\)-hydroxylase deficiency. Exclusion criteria for CAH and healthy subjects were: previous history of cardiovascular disease, use of combined contraceptives within 1 month of data collection, and current pregnancy. Patients with CAH were treated, as needed by their standard of care, with hydrocortisone or dexamethasone, and fludrocortisone. All patients required to be fasting for at least 12 hours prior to blood sampling, and to refrain from coffee-containing products for at least 24 hours.

**Study objectives.** The main objective of this ancillary study was to characterize the metabolomic profile of CAH patients compared to that of matched healthy controls using untargeted approaches. To this aim, two analytic methods were performed.

**Hormone analyses.** Blood samples for the determination of serum concentrations of 17-OH progesterone, progesterone, estradiol, total testosterone, aldosterone, renin, androstenedione, sex hormone-binding globulin (SHBG), adrenocorticotropic hormone (ACTH), follicle-stimulating hormone (FSH), luteinizing hormone (LH) were collected in dry tubes and further assayed in the immunology laboratory of Pitié-Salpêtrière University Hospital. Estradiol, progesterone, SHBG and testosterone were assayed using chemiluminescence (Modular-E170; Roche, Mannheim, Germany) and 17-OH progesterone by radioimmunoassay (KIP1409; DIAsource ImmunoAssays, Louvain-la-Neuve, Belgium). Cortisol was sampled concomitantly with the other hormones, more than 8 hours after last intake of corticosteroids in CAH patients and after 12 hours of fasting in healthy subjects. Baseline hormonal measurements were assessed in all subjects, under their regular treatment for CAH patients.

**Metabolomic analyses preparations.** Eight volumes of frozen acetonitrile (\(-20^\circ\text{C}\)) containing internal standard (labelled mix of amino acids at 12.5\(\mu\text{g/mL}\)) were added to 50\(\mu\text{L}\) serum samples and vortexed. Samples were sonicated for 15 minutes and centrifuged for 2 minutes at 10,000 \(\times\) g and at 4\(^\circ\text{C}\). Then, centrifuged samples were incubated at 4\(^\circ\text{C}\) during one hour for slow protein precipitation. Samples were then centrifuged at 20,000 \(\times\) g at 4\(^\circ\text{C}\) and supernatants were transferred to another series of tubes and then dried out and frozen at \(-80^\circ\text{C}\) until the liquid chromatography-mass spectrometry (LC-MS) analyses. Samples were reconstituted based on starting mobile phase composition of the chromatographic method (water/acetonitrile (99:1; v-v) containing 0.1% of formic acid). Reconstituted samples were, then, centrifuged and transferred to vials before LC-MS analyses.

**Ultra-performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) analyses of serum samples.** Metabolite profiling analysis was carried out on a UPLC Waters Acquity (Waters Corp, Saint-Quentin-en-Yvelines, France) coupled to a hybrid Orbitrap based instrument, a Q-Exactive (Thermo Fisher Scientific, Illkirch, France). LC-MS analyses was made in full scan positive and negative mode with a resolution of 70,000 (FWHM) and a scan range of m/z 50–750. Mass spectrometer was systematically calibrated before LC-MS analyses in both ion polarity modes with the Pierce calibration solution.

LC-MS analyses were performed using a modification of the method of Garali et al. to screen microbiota-derived metabolites\textsuperscript{21}. Briefly, a ‘PFPP column’ Discovery HS F5-PFPP column, 370 000 (FWHM) and a scan range of m/z 50–750. Mass spectrometer was systematically calibrated before LC-MS analyses in both ion polarity modes with the Pierce calibration solution.

LC-MS analyses were performed after last intake of corticosteroids in CAH patients and after 12 hours of fasting in healthy subjects. Baseline hormonal measurements were assessed in all subjects, under their regular treatment for CAH patients.

**Untargeted metabolomics data analyses.** Processing steps of MS data, including peak picking, peak grouping, retention time correction, and annotation of isotopes and adducts, were performed using XCMS R package with CentWave algorithm and CAMERA tools implemented in R software and the galaxy workflow.

Processing of LC-MS data were analyzed based on standard protocols and resulted in a datamatrix in which each metabolomic feature was characterized by a retention time (RT), mass to charge ratio (m/z), and its corresponding intensities for each sample and the isotope-adduct annotation from CAMERA tool.

Metabolomics data matrix was filtered, normalized, curated and log-10 transformed based on quality assurance (QA) strategy\textsuperscript{27,28}. Peaks with more than 30% of missing values were discarded. Metabolites from the Xenobiotics category were also discarded. Principal component analysis (PCA) was also plotted to assess the absence of technical drift during data acquisition process (data not shown). Robust features were annotated based on their mass to charge ratio (m/z) and retention time (RT) using public, and ‘in-house’ databases then confirmed based on MS/MS experiments. Following these processing and annotation steps, 140 features met the acceptance criteria. Moreover, indoleamine 2,3 dioxygenase (IDO) enzyme activity was computed as the abundance ratio of kynurenine over tryptophan.

**Statistical analysis.** Continuous variables are presented as median (interquartile range) and categorical variables as number (proportions). For enrichment and statistical analyses of the metabolomics data the MetaboDiff R package (v 0.9.3) was used\textsuperscript{29}. Briefly, the MetaboDiff analysis consists on different steps, including annotation, imputing missing data by k-nearest neighbor approach and removal of metabolites with at least 40% of missing data. Data are then renormalized using variance stabilizing normalization to ensure that variance remains nearly constant over the measured spectrum\textsuperscript{30}. To derive meaningful subpathways that are enriched between groups, MetaboDiff generates a metabolic correlation network, which offers the possibility to integrate external information such as pathway information\textsuperscript{12}. Module significance can be determined as the average absolute metabolite
The Tryptophan Kynurenine Pathway and Its Importance to Neurotherapeutics

The Tryptophan Kynurenine Pathway and Its Importance to Neurotherapeutics

What is the Tryptophan Kynurenine Pathway and Why is it Important to Neurotherapeutics?

1. Salem, J. E.
2. Bachelot, A., Grouthier, V., Courtillot, C., Dulon, J. & Touraine, P. MANAGEMENT OF ENDOCRINE DISEASE: Congenital adrenal hyperplasia.
3. Fardet, L. & Fève, B. Systemic glucocorticoid therapy: a review of its metabolic and cardiovascular adverse events.
4. Oray, M., Abu Samra, K., Ebrahimiadib, N., Meese, H. & Foster, C. S. Long-term side effects of glucocorticoids.
5. Falhammar, H.
6. Falhammar, H. Increased Mortality in Patients With Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency. The Journal of Clinical Endocrinology & Metabolism 99, E2715–E2721, https://doi.org/10.1210/jc.2014-2957 (2014).
7. Falhammar, H. et al. Increased Cardiovascular and Metabolic Morbidity in Patients With 21-Hydroxylase Deficiency: A Swedish Population-Based National Cohort Study. The Journal of clinical endocrinology and metabolism 100, 3520–3528, https://doi.org/10.1202/jc.2015-2093 (2015).
8. Reisch, N. Substitution therapy in adult patients with congenital adrenal hyperplasia. Best Pract Res Clin Endocrinol Metab 29, 33–45, https://doi.org/10.1016/j.bjendem.2014.11.002 (2015).
9. Kaddurah-Daouk, R. & Weinshilboum, R. Metabolomic Signatures for Drug Response Phenotypes: Pharmacometabolomics Enables Precision Medicine. Clinical pharmacology and therapeutics 98, 71–75, https://doi.org/10.1016/j.cpt.2013 (2014).
10. Alwesh, M. A. et al. Plasma metabolic profile varies with glucocorticoid dose in patients with congenital adrenal hyperplasia. Sci Rep 7, 17092, https://doi.org/10.1038/s41598-017-17220-5 (2017).
11. Piffli, E. et al. Interpretable and accurate prediction models for metagenomics data. GigaScience 9, https://doi.org/10.1093/gigascience/giaa010 (2020).
12. Mock, A. et al. MetaboDiff: an R package for differential metabolic analysis. Bioinformatics (Oxford, England) 34, 3417–3418, https://doi.org/10.1093/bioinformatics/btt344 (2018).
13. Salem, J. E. et al. Complex Association of Sex Hormones on Left Ventricular Systolic Function: Insight into Sexual Dimorphism. Journal of the American Society of Echocardiography: official publication of the American Society of Echocardiography 31,231–240, e231, https://doi.org/10.1016/j.echo.2017.10.017 (2018).
14. Bordag, N. et al. Glucocorticoid (dexamethasone)-induced metabolism changes in healthy males suggest prediction of response and side effects. Sci Rep 5, 15954, https://doi.org/10.1038/srep15954 (2015).
15. Geng, J. & Liu, A. Heme-dependent dioxygenases in tryptophan oxidation. Archives of biochemistry and biophysics 544, 18–26, https://doi.org/10.1016/j.abb.2013.11.009 (2014).
16. Davis, I. & Liu, A. What is the tryptophan kynurenine pathway and why is it important to neurotherapeutics? Expert review of neurotherapeutics 15, 719–721, https://doi.org/10.1586/14737175.2015.1049999 (2015).
17. Fluck, C. E. & Miller, W. L. P450 oxidoreductase deficiency: a new form of congenital adrenal hyperplasia. Current opinion in pediatrics 28, 435–441, https://doi.org/10.1097/MOP.0000226369571956.5x (2016).
18. Pizzichini, M., Di Stefano, A. & Marinello, E. The regulation of purine ribonucleotide biosynthesis by glucocorticoid hormones. The Italian journal of biochemistry 34, 305–312 (1985).
19. Nguyen, L. S. et al. Influence of hormones on the immunotolerogenic molecule HLA-G: a cross-sectional study in patients with congenital adrenal hyperplasia. European journal of endocrinology 181, 481–488, https://doi.org/10.1530/ije-19-0379 (2019).
20. Rosenbaum, D. et al. Early central blood pressure elevation in adult patients with 21-hydroxylase deficiency. J Hypertens 37, 175–181, https://doi.org/10.1093/ijh/ihx007 (2019).
21. Garali, I. et al. A strategy for multimodal data integration: application to biomarkers identification in spinocerebellar ataxia. Briefings in bioinformatics 19, 1356–1369, https://doi.org/10.1093/bib/bby060 (2018).
22. Giacomoni, F. et al. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. Bioinformatics (Oxford, England) 31, 1493–1495, https://doi.org/10.1093/bioinformatics/btu813 (2015).
23. Smith, C. A., Want, E. J., O’Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical chemistry 78, 779–787, https://doi.org/10.1021/ac051437y (2006).
24. Tautenhahn, R., Patti, G. J., Rinnehart, D. & Suzuki, G. XCMS Online: a web-based platform to process untargeted metabolic data. Analytical chemistry 84, 5035–5039, https://doi.org/10.1021/ac300606x (2012).
25. Want, E. J. et al. Global metabolic profiling procedures for urine using UPLC-MS. Nature protocols 5, 1005–1018, https://doi.org/10.1038/nprot.2010.50 (2010).
26. Dunn, W. B. et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nature protocols 6, 1060–1083, https://doi.org/10.1038/nprot.2011.335 (2011).
27. Dunn, W. B., Wilson, J. D., Nicholls, A. W. & Broadhurst, D. The importance of experimental design and Qc samples in large-scale and Ms-driven untargeted metabolic studies of humans. Bioanalysis 4, 2249–2264, https://doi.org/10.4155/bio.12.204 (2012).
28. Veselkov, K. A. et al. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. Analytical chemistry 83, 5864–5872, https://doi.org/10.1021/ac201065j (2011).

Significance Measure. The Predomics approach, based on specific machine-learning models and genetic algorithm, searches a few sparse models in a very large combinatorial space. The best ternary models is then selected by applying a penalty on the number of features used – for each feature added the accuracy is decreased by a value of 1/100. A feature importance score is computed during the cross validation. The resulting data were used to assess differential abundance between patients and controls and using unpaired non-parametric tests (Mann-Whitney). Associations with clinical numerical variables were tested using Spearman correlations. P-values were adjusted for multiple testing using the Benjamini-Hochberg method. Prediction analyses were performed using the Predomics package searching for ternary models with the beam-search algorithm using default parameters. Here we used the MetaBoxDiff imputed data prior to normalization for interpretability reasons. Cross-validation was performed using a 20-times 5-fold configuration and the sparsity penalization coefficient was set to 1%.

Statement. All described methods were carried out with relevant guidelines and regulations.
Acknowledgements
We are extremely grateful to the study participants. The study sponsor was Assistance Publique – Hôpitaux de Paris (Département de la Recherche Clinique et du Développement). *CARDIOHCS study group: Philippe Chanson, Sylvie Salenave, Sophie Christin Maitre, Nathalie Bourcigaux, Michel Polak, Juliane Leger, Jérôme Dulon, Nora Dahmoune, Eric Bruckert.

Author contributions
L.S.N. and E.P. equally contributed to the manuscript. Both wrote the manuscript, participated to its design and performed analyses on the database. F.I. performed all metabolomics analyses. M.L. performed all hormone analyses. C.F.-B. and P.T. provided critical reviewing. J-E.S. participated to the study design, analyses and provided critical reviewing. A.B. designed the study, participated to patients enrollment and provided critical overview of the project.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-65897-y.

Correspondence and requests for materials should be addressed to L.S.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020