The Application of Principal Component Analysis on Clinical and Biochemical Parameters Exemplified in Children With Congenital Adrenal Hyperplasia

Marie Lindhardt Ljubicic1,2, Andre Madsen1,2,3, Anders Juul1,2, Kristian Almstrup1,2 and Trine Holm Johannsen1,2*

1 Department of Growth and Reproduction, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 2 International Center for Research and Research Training in Endocrine Disruption of Male Reproduction and Child Health (EDMaRC), Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 3 Hormone Laboratory, Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Bergen, Norway

Purpose: Principal component analysis (PCA) is a mathematical model which simplifies data into new, combined variables. Optimal treatment of pediatric congenital adrenal hyperplasia (CAH) remains a challenge and requires evaluation of all biochemical and clinical markers. The aim of this study was to introduce PCA methodology as a tool to optimize management in a cohort of pediatric and adolescent patients with CAH by including adrenal steroid measurements and clinical parameters.

Methods: This retrospective, longitudinal cohort of 33 children and adolescents with CAH due to 21-hydroxylase deficiency included 406 follow-up observations. PCAs were applied to serum hormone concentrations and compared to treatment efficacy evaluated by clinical parameters.

Results: We provide and describe the first PCA models with hormone parameters denoted in sex- and age-adjusted standard deviation (SD) scores to comprehensibly describe the combined ‘endocrine profiles’ of patients with classical and non-classical CAH, respectively. Endocrine profile scores were predictive markers of treatment efficacy for classical (AUC=92%; accuracy 95%; p=1.8e-06) and non-classical CAH (AUC=80%; accuracy 91%; p=0.004). A combined PCA demonstrated clustering of patients with classical and non-classical CAH by serum 17-hydroxyprogesterone (17-OHP) and dehydroepiandrosterone-sulphate (DHEAS) concentrations.

Conclusion: As an example of the possibilities of PCA, endocrine profiles were successfully able to distinguish between patients with CAH according to treatment efficacy and to elucidate biochemical differences between classical and non-classical CAH.

Keywords: principal component analysis, congenital adrenal hyperplasia, CAH, endocrine profiling, treatment efficacy
INTRODUCTION

Congenital adrenal hyperplasia (CAH) is a genetic disorder primarily caused by 21-hydroxylase deficiency. Patients are grouped into classical and non-classical forms according to severity reflecting residual CYP21A2 enzyme activity. The pathophysiology of CAH is largely centered around alterations in the adrenal steroidogenesis, resulting in elevated serum concentrations of 17-hydroxyprogesterone (17-OHP) including the adrenally derived 21-deoxycortisol (1), androgens including 11-oxygenated 19-carbon androgens (2–4) and especially in classical CAH reduced mineralocorticoids and glucocorticoids. Clinically, CAH is associated with increased virilization (5), classical CAH reduced mineralocorticoids and glucocorticoids.

**Patient Care.**

Optimal treatment remains a challenge (1) and relies on evaluation of a growing panel of biochemical and clinical markers (5). Interpreting combinations of these markers introduces complexity, subjectivity, and bias into the patient care.

Principal component analysis (PCA) is a statistical method that can be applied to datasets to obtain a simplified model for stratifying patients or phenotypes by reducing the number of variables (10). In other words, PCA condenses the variables of a dataset into a smaller number of new ‘variables’ named principal components. As such, each principal component is an explanatory variable which represents the relationship between the original variables in the dataset. These principal components provide scores for combinations of variables that explain the most variance within a dataset. Importantly, compared to simpler statistical models and comparisons in which variables are included unweighted, PCA weights the variables according to their relative importance. PCA is therefore a method that can aid in determining key variables and phenotype clusters in large datasets. For example, principal component scores have been applied to enumerate and assign severity scores in newborn screening of CAH (11), in pediatric metabolic syndrome (12), as endocrine profile scores in female puberty (13), and in dietary patterns during childhood (14).

With the general aim of introducing PCA methodology as a method to optimize the interpretation of complex data in diagnostics and management of patients with Differences of Sex Development (DSD), we applied this method to adrenal steroid measurements and clinical parameters in a cohort of pediatric and adolescent patients with CAH.

**MATERIALS AND METHODS**

**Patient Cohort**

In this study the total CAH cohort included 33 patients aged 0.3 to 18.9 years. This cohort was described in a recent publication focusing on the use of the standard deviation (SD) scores in the management of CAH (15). However, one patient with an additional diagnosis of 45,X/46,XY mosaicism was excluded, as this condition could potentially influence the PCA, while a second patient was excluded due to lack of clinical information.

**Hormone Assays**

Serum 17-OHP and the adrenal androgens (DHEAS, androstenedione, and testosterone) were measured by liquid chromatography tandem-mass spectrometry-based (LC-MS/MS) with inter-assay coefficients of variation (CVs) ranging...
from 1.4 to 2.5% (15). These androgens were included as they are part of routine follow-up in our clinic. The method has previously been described in detail (16). Serum LH and FSH were analyzed by a time-resolved fluoro-immuno-assay (AutoDELFIA, Perkin Elmer, Turku, Finland) with an inter-assay coefficient of variation (CV) below 5% and a limit of detection (LoD) of 0.05 IU/L for both. Until September 2014, serum SHBG was measured by AutoDELFIA with a CV below 6% and a LoD of 0.23 nM, while it from September 2014 and onwards was measured by a chemiluminescence immunoassay (Access2, Beckman Coulter, Brea, CA, USA) with a LoD of 0.35 nmol/L and a CV below 5%. All AutoDELFIA-derived SHBG results were factored to corresponding Access2 SHBG results after internal method comparison.

Hormone References Curves
The use of the Generalized Additive Model for Location, Scale, and Shape (GAMLSS) to obtain SD scores has previously been described for height and BMI (17), blood pressure (15), 17-OHP and the adrenal androgens (15), and LH and FSH (18). Based on healthy participants aged 0.2 to 20.0 years of age (19), reference curves were established for SHBG (1068 males, 668 females). The GAMLSS algorithm was applied using R (20).

Principal Component Analysis
Dimension reduction by PCA was applied to sex- and age-adjusted SD scores for the following endocrine variables: 17-OHP, DHEAS, androstenedione, testosterone, SHBG, LH, and FSH. Patient observations with one or more missing hormone measurements were not included. Accordingly, 109 and 93 eligible visits (i.e. observations) from 18 patients with classical CAH and 15 patients with non-classical CAH, respectively, were included. PCA computation was done using the prcomp() command in R without variable scaling, as previously described (13). PCA biplots were generated using the ggbiplot package in R (21).

The PCA output typically consists of five to ten uncorrelated principal components, each explaining a proportion of the total dataset variance. The first principal component typically explains most of the variance, but successive principal components may also detect relevant variance aligning with a topic of interest (e.g. treatment efficacy). In this study, principal components were described in terms of a) proportion of explained variance and b) Eigenvalue (i.e. the principal component standard deviation squared). Only principal components with an Eigenvalue above 1.0 were considered significant (22). Each principal component contains a unique combination of correlation coefficients between the principal component and the input variables. These coefficients were deemed strong if above 0.4.

The correlation coefficients from the different principal components can be used to calculate new, numerical principal components scores (hereafter named endocrine profile scores because only the biochemical markers are included in the PCAs included in this study). In context of the total dataset variance, such endocrine profile scores represent coordinates that tend to cluster related ‘phenotypes’ within a dataset. Thus, endocrine profile scores were applied to distinguish between treatment efficacy groups. The coefficients required to calculate endocrine profile scores equivalent to our current models to detect treatment efficacy are provided in Supplementary Table S1.

Other Statistics
Receiver operating characteristic (ROC) curves were computed using the pROC package in R and used to evaluate the ability of individual biochemical markers and the PCA-based endocrine profile scores to distinguish between optimally and insufficiently treated patients. The ROC Youden index was used to determine the optimal biomarker cutoff to distinguish the two groups (23). Also based on ROC, the area under the curve (AUC) was used to evaluate the ability of a variable (i.e. endocrine profile score and individual biochemical markers to distinguish between groups: ≥90%: excellent, ≥80%: good, ≥70%: fair, ≥60%: poor, and <60%: fail, as previously detailed (24). Statistical differences between groups were evaluated by Wilcoxon rank sum test (R), corresponding to the Mann Whitney U test (GraphPad Prism).

RESULTS
In total, 33 children and adolescents with CAH were included (six boys and 27 girls), representing 406 successive visits to the outpatient clinic. Glucocorticoid treatment was administered at all visits by classical CAH patients and in 137 of 187 visits by non-classical CAH patients. Treatment efficacy was evaluated at each visit by SD scores and grouped accordingly into optimally, sub-optimally, or insufficiently treated, as described in the Methods section. Baseline characteristics of the optimally and insufficiently treated patients are provided in Table 1. To minimize misclassification across treatment efficacy groups, we focused on the two treatment extremes and assessed the performance of the listed hormones as biochemical markers (Table 1). Notably, optimally treated patients were significantly shorter and leaner than insufficiently treated patients. There was no difference in glucocorticoid doses between the treatment groups (Table 1). Optimally treated patients exhibited significantly lower SD scores of serum DHEAS, testosterone, SHBG, and LH, and significantly higher SD scores of serum FSH, while SD scores of serum 17-OHP did not differ between the patient groups. As evident from ROC analyses, no single hormone represented a viable biochemical marker of treatment efficacy.

In order to investigate if any combination of these hormones could discriminate between optimally and insufficiently treated patients, we applied PCA to all included hormones. Due to the marked clinical differences between classical and non-classical CAH, a PCA was performed for each CAH group separately. The PCAs produced four principal components for each CAH group, each accounting for substantial variance (Eigenvalues > 1). Table 2. The listed correlation coefficients describe the relative contribution of each hormone to a given principal component. In classical CAH, principal component 1 (representing 76% of the total variance) was primarily dominated by 17-OHP with a correlation coefficient of 0.93. In contrast, principal component 1...
In this pilot study, we introduce the simultaneous and combined evaluation of several hormones by use of PCAs in a cohort of children and adolescents with CAH and demonstrate the immediate benefits of this method of simplification. Namely, we show that an endocrine profile score produced by PCA was able to distinguish between optimally and insufficiently treated patients with classical CAH, although conclusions are limited by the small dataset. Thus, further studies applying PCA in CAH cohorts are warranted.

### DISCUSSION

In this study, we developed a new endocrine profile score that can be applied in both classical and non-classical CAH patients. The score is calculated using principal component analysis (PCA) and includes variables such as 17-OHP, DHEAS, androgens, and gonadotropins. PCA allowed for a different overview of the data not aimed at distinguishing between treatment efficacy groups. Specifically, we observed a strong correlation between the 17-OHP and DHEAS concentrations in classical CAH (0.95) and a weaker correlation in non-classical CAH (0.69). This relationship is depicted in the biplot (Figure 1). The biplot revealed a perpendicular relationship between the 17-OHP and DHEAS axes, which is a characteristic of the PCA method.

Combining the classical and non-classical CAH datasets in a single PCA allowed for a different overview of the data not aimed at distinguishing between treatment efficacy groups. The new PCA produced an endocrine profile score that was able to distinguish between optimally and insufficiently treated patients. With the aim of investigating the underlying biochemistry of CAH, a so-called biplot featuring the two principal components accounting for the most explained variance was created (Figure 1). The accuracies of the endocrine profile score in classical CAH (95%), non-classical CAH (91%), and classical CAH (97%) were obtained for each patient. By applying the scores directly in individual patients, we were able to distinguish between optimally and insufficiently treated patients. With the aim of investigating the underlying biochemistry of CAH, a so-called biplot featuring the two principal components accounting for the most explained variance was created (Figure 1). The accuracies of the endocrine profile score in classical CAH (95%), non-classical CAH (91%), and classical CAH (97%) were obtained for each patient. By applying the scores directly in individual patients, we were able to distinguish between optimally and insufficiently treated patients. With the aim of investigating the underlying biochemistry of CAH, a so-called biplot featuring the two principal components accounting for the most explained variance was created (Figure 1). The accuracies of the endocrine profile score in classical CAH (95%), non-classical CAH (91%), and classical CAH (97%) were obtained for each patient. By applying the scores directly in individual patients, we were able to distinguish between optimally and insufficiently treated patients.
Within the field of Differences of Sex Development, new clinical and biochemical markers continue to emerge, making it increasingly difficult for clinicians to reconcile the many markers at the same time. The superiority of PCA compared to simpler, traditional statistical methods is best described by its ability to reduce complex data by identifying redundant variables and uncover patient clusters that may correspond to clinically relevant phenotypes. PCA has previously been successfully used in endocrinology, for example to enhance detection of CAH in newborns (11) and pubertal onset in girls (13) and to distinguish endocrine profile clusters with implications for pubertal timing in girls (25). Moreover, PCA enabled the conclusion that the metabolic profile of polycystic ovarian syndrome is neither intrinsic nor specific (26).

Applying PCA to our dataset revealed possible new approaches to the management of patients with CAH, although conclusions are limited due to the small dataset. Nonetheless, PCA allowed for principal components to be created that included and weighted all hormone SD scores that were included in the model. The corresponding endocrine profile scores (calculated from principal component 2 in classical CAH and from principal component 3 in non-classical CAH) were able to distinguish between optimally and insufficiently treated patients. The fact that these scores outperformed the individual biochemical markers highlights the strength of combining hormones into a single variable by PCA. Theoretically, the endocrine profile scores could be applied in any outpatient clinic, enabling the physician to determine whether e.g. adjustment of glucocorticoid treatment in a CAH patient is needed at any given time. By simple calculation of endocrine profile scores, subsequent evaluation of treatment efficacy by relevant cutoffs is theoretically possible in other centers as well. Ideally, each center should calculate its own cutoffs based on local data, but the cutoffs presented in this study are based on sex and age-adjusted SD scores, which allows for comparison across sexes, ages, and centers. Thus, in the real world, the current cutoffs may aid clinicians from other centers as well if applied with caution. Despite this, we also recognize the need for a larger and more homogenous dataset to fully optimize management of CAH patients.

In line with this, our dataset was small, and the outcome of treatment efficacy was limited partly by the simplistic and somewhat artificial groups and partly by the availability of clinical information for each visit, making misclassification a possibility. Furthermore, we were unable to include clinical information such as virilization, bone ages, and other biochemical markers (renin, 21-deoxycortisol, and 11-oxygenated 19-carbon steroids), which would arguably have improved stratification and optimized our PCA results.

Examining the two PCA outputs from the patients with classical and non-classical CAH, respectively, revealed integral differences in terms of endocrine constellations. In classical CAH, 17-OHP was strongly correlated to the primary principal component, whereas in non-classical CAH both 17-OHP and DHEAS were strongly correlated to the primary principal components.
component. Thus, 17-OHP was shown to dominate both primary principal components, but it did not discriminate between optimally and insufficiently treated patients. In line with current guidelines (5), this highlights that 17-OHP is not a viable single marker of treatment efficacy. Similarly, DHEAS played a larger role in the variation between patients with non-classical CAH than between patients with classical CAH. These differences in 17-OHP and DHEAS were also visible when combining all patients (classical as well as non-classical) in a PCA-derived biplot that clustered the two CAH groups. The differences in 17-OHP and DHEAS were further highlighted by the ability of their ratio to distinguish between patients with classical and non-classical CAH. However, due to the majority of the patients undergoing treatment in our cohort, this ratio cannot be directly applied in CAH diagnostics. Though the ratio may not be clinically relevant, the process of deriving the ratio and the finding altogether demonstrates that PCA methodology can be used to illuminate the underlying biochemistry. This use of PCA has also previously been reported in patients with polycystic ovaries (26). By example, we therefore demonstrate how PCA can introduce new markers and ratios of possible clinical interest in future research projects on patients with CAH.

The methodological strengths of this pilot study included 1) the introduction of a powerful statistical method in the context of CAH; 2) the use of SD scores for all clinical and biochemical markers which allowed for PCA models and comparisons across sex and age and without consequent loss of sample size; and 3) the use of liquid chromatography tandem-mass spectrometry in the analysis of 17-OHP and adrenal androgens. The limitations included 1) the dataset was small and each observation in the PCA model equated a follow-up visit introducing possible biases, i.e., insufficiently treated patients at higher risk of complications may have been overrepresented due to more visits as a direct consequence; 2) in 73% of visits by patients with non-classical CAH, patients received glucocorticoid treatment which may be an overrepresentation of those needing treatment; 3) stratification of CAH patients according to treatment efficacy based on equally weighted clinical parameters (anthropometry and blood pressure) was simplistic and may not
accurately reflect true treatment efficacy; 4) misclassification between treatment efficacy groups was also a possibility, and the largest group (suboptimal) was not analyzed in ROC analyses; 5) treatment efficacy classification did not include virilization status, pubertal status or bone age due to lack of uniformly collected data; 6) measurements of C-19 oxygenated steroids and 21-deoxycorticisol were not included as they were not measured; 7) measurements of renin were not included due to lack of reference ranges and thereby lack of SD scores; and 8) blood sampling was done during the opening hours of the outpatient clinic without regard to medicine intake.

In conclusion, with the aim of introducing a combined and simultaneous evaluation of several hormones, we described and applied PCA in a cohort of children and adolescents with CAH. Via multiple hormones reduced to endocrine pro
cutoffs, this approach is directly adaptable into a clinical setting. Moreover, the PCA model enabled us to elucidate the relative importance of individual hormones in patients with CAH. We therefore propose that PCA be applied as a tool in future research on larger DSD cohorts.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of data protection laws. However, the codes used in R are available upon request. Requests to access the datasets should be directed to trine.holm.johannsen@regionh.dk.

REFERENCES

1. Merke DP, Auchus RJ, Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency. N Engl J Med (2020) 383(13):1248–61. doi: 10.1056/NEJMra1909786
2. Turcu AF, Nanba AT, Chomic R, Upadhyay SK, Giordano TJ, Shields JJ, et al. Adrenal-Derived 11-Oxygenated 19-Carbon Steroids Are the Dominant Androgens in Classic 21-Hydroxylase Deficiency. Eur J Endocrinol (2016) 174(5):601–9. doi: 10.1530/EJE-15-11181
3. Kamrath C, Wettstaedt L, Boettcher C, Hartmann MF, Wudy SA. Androgen Excess Is Due to Elevated 11-Oxygenated Androgens in Treated Children With Congenital Adrenal Hyperplasia. J Steroid Biochem Mol Biol (2018) 178:221–8. doi: 10.1016/j.jsbmb.2017.12.016
4. Jha S, Turcu AF, Sinait N, Brooker B, Auchs RJ, Merke DP, 11-Oxygenated Androgens Useful in the Setting of Discrepant Conventional Biomarkers in 21-Hydroxylase Deficiency. J Endocr Soc (2020) 5(2):1–9. doi: 10.1210/jendo/bva192
5. Speiser PW, Arlt W, Auchs RJ, Baskin LS, Conway GS, Merke DP, et al. Congenital Adrenal Hyperplasia Due to Steroid 21-Hydroxylase Deficiency: An Endocrine Society® Clinical Practice Guideline. J Clin Endocrinol Metab (2018) 103(11):4043–88. doi: 10.1210/jc.2018-01865
6. Mathusamy K, Elamin MB, Smushkin G, Murad MH, Lampropulos JF, Elamin KR, et al. Adult Height in Patients With Congenital Adrenal Hyperplasia: A Systematic Review and Metaanalysis. J Clin Endocrinol Metab (2010) 95(9):4161–72. doi: 10.1210/jc.2009-2616
7. Volčk TMK, Simm D, Beier C, Dörr HG. Obesity Among Children and Adolescents With Classic Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency. Pediatr (2006) 117(1):e98–105. doi: 10.1542/peds.2005-1005
8. Mooij CF, van Herwaarden AE, Swee FCG, Roeleveld N, de Korte CL, Kapusta L, et al. Cardiovascular and Metabolic Risk in Pediatric Patients With Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency. J Pediatr Endocrinol Metab (2020) 33(9):957–66. doi: 10.1515/pem-2020-0068
9. Roche EF, Charmandari E, Dattani MT, Hindmarsh PC. Blood Pressure in Children and Adolescents With Congenital Adrenal Hyperplasia (21-Hydroxylase Deficiency): A Preliminary Report. Clin Endocrinol (Oxf) (2003) 58(5):589–96. doi: 10.1046/j.1365-2265.2003.01757.x
10. Ringnér M. What Is Principal Component Analysis? Nat Biotechnol (2008) 26(3):303–4. doi: 10.1038/nbt0308-303
11. Lasarev MR, Bialk ER, Allen DB, Held PK. Application of Principal Component Analysis to Newborn Screening for Congenital Adrenal Hyperplasia. J Clin Endocrinol Metab (2020) 105(8):e2930–40. doi: 10.1210/clinem/dgaa371
12. Katzmarnyk PT, Pérusse L, Malina RM, Bergeron J, Després J-P, Bouchard C. Stability of Indicators of the Metabolic Syndrome From Childhood and Adolescence to Young Adulthood: The Québec Family Study. J Clin Epidemiol (2001) 54(2):190–5. doi: 10.1016/S0895-4356(00)00315-2
13. Madsen A, Bruserud IS, Berteisen B-E, Roelants M, Oehme NH, Viste K, et al. Hormone References for Ultrasound Breast Staging and Endocrine Profiling to Detect Female Onset of Puberty. J Clin Endocrinol Metab (2020) 105(12):e4886–95. doi: 10.1210/clinem/dgaa679
14. Günther ALB, Schulze MB, Kroke A, Diethelm K, Joslowski G, Krupp D, et al. Early Diet and Later Cancer Risk: Prospective Associations of Dietary Patterns During Critical Periods of Childhood With the GH-IGF Axis, Insulin Resistance and Body Fatness in Younger Adulthood. Nutr Cancer (2015) 67(6):877–92. doi: 10.1080/01635581.2015.1056313
15. Clausen CS, Ljubicic ML, Main KM, Andersson A-M, Petersen JH, Frederiksen H, et al. Congenital Adrenal Hyperplasia in Children: A Pilot Study of Steroid Hormones Expressed as Sex- and Age-Related Standard Deviation Scores. Horm Res Paediatr (2020) 93(4):226–38. doi: 10.1159/000509079

ETHICS STATEMENT

This study involving human participants were reviewed and approved by the Danish Patient Safety Authority (no. 3-3013-1376/1) and the Danish Data Protection Agency (no. 2015-235, I-Suite no. 04204). Written informed consent from the participants’ legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ML and TJ took part in the study design, execution, analysis, manuscript drafting and critical discussion. AM, AJ, and KA took part in the execution, analysis, manuscript drafting and critical discussion. All authors contributed to the article and approved the submitted version.

FUNDING

ML was funded by the Absalon Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2021.652888/full#supplementary-material
16. Søeborg T, Frederiksen H, Johannsen TH, Andersson A-M, Juul A. Isotope-Dilution TurboFlow-LC-MS/MS Method for Simultaneous Quantification of Ten Steroid Metabolites in Serum. Clin Chim Acta (2017) 468:180–6. doi: 10.1016/j.cca.2017.03.002

17. Tinggaard J, Aksglaede L, Sørensen K, Mouritsen A, Wohlfahrt-Veje C, Hagen CP, et al. The 2014 Danish References From Birth to 20 Years for Height, Weight and Body Mass Index. Acta Paediatr (2014) 103(2):214–24. doi: 10.1111/apa.12468

18. Ljubicic ML, Jespersen K, Aksglaede L, Hagen CP, Petersen JH, Andersen HR, et al. The LH/FSH Ratio Is Not a Sex-Dimorphic Marker After Infancy: Data From 6417 Healthy Individuals and 125 Patients With Differences of Sex Development. Hum Reprod (2020) 35(10):2323–35. doi: 10.1093/humrep/deaa182

19. Søeborg T, Frederiksen H, Mouritsen A, Johannsen TH, Main KM, Jørgensen N, et al. Sex, Age, Pubertal Development and Use of Oral Contraceptives in Relation to Serum Concentrations of DHEA, DHEAS, 17α-Hydroxyprogesterone, Δ4-Androstenedione, Testosterone and Their Ratios in Children, Adolescents and Young Adults. Clin Chim Acta (2014) 437:6–13. doi: 10.1016/j.cca.2014.06.018

20. Rigby RA, Stasinopoulos DM. Generalized Additive Models for Location, Scale and Shape. J R Stat Soc Ser C (2005) 54(3):507–54. doi: 10.1111/j.1467-9876.2005.00510.x

21. Wickham H. Ggplot2: Elegant Graphics for Data Analysis. 2nd ed. Cham: Springer International Publishing (2016).

22. Jackson DA. Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. Ecology (1993) 74(8):2204–14. 10.2307/1939574. doi: 10.2307/1939574

23. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Casp J Intern Med (2013) 4(2):627–35.

24. Safari S, Baratloo A, Elfi M, Negida A. Evidence Based Emergency Medicine; Part 5 Receiver Operating Curve and Area Under the Curve. Emerg (Tehran Iran) (2016) 4(2):111–3.

25. Faulker CS, Gutmark-Little I, Xie C, Giannini CM, Chandler DW, Biro FM, et al. Sex Hormone Phenotypes in Young Girls and the Age at Pubertal Milestones. J Clin Endocrinol Metab (2019) 104(12):6079–89. doi: 10.1210/jc.2019-00889

26. Dewailly D, Pigny P, Soudan B, Catteau-Jonard S, Decanter C, Poncelet E, et al. Reconciling the Definitions of Polycystic Ovary Syndrome: The Ovarian Follicle Number and Serum Anti-Müllerian Hormone Concentrations Aggregate With the Markers of Hyperandrogenism. J Clin Endocrinol Metab (2010) 95(9):4399–405. doi: 10.1210/jc.2010-0334

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Ljubicic, Madsen, Juul, Almstrup and Johannsen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.