



HAL
open science

Application de méthodes d'intelligence artificielle à l'exploration des troubles métaboliques et endocriniens dans les fonctions reproductives et l'infertilité

Guillaume Bachelot

► To cite this version:

Guillaume Bachelot. Application de méthodes d'intelligence artificielle à l'exploration des troubles métaboliques et endocriniens dans les fonctions reproductives et l'infertilité. Médecine humaine et pathologie. Sorbonne Université, 2024. Français. NNT : 2024SORUS096 . tel-04650330

HAL Id: tel-04650330

<https://theses.hal.science/tel-04650330v1>

Submitted on 16 Jul 2024

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.

Sorbonne Université

ED 394 : Physiologie, physiopathologie et thérapeutique

Centre de Recherche Saint-Antoine INSERM - UMR S-938

Application de méthodes d'intelligence artificielle à l'exploration des troubles métaboliques et endocriniens dans les fonctions reproductives et l'infertilité.

Par Guillaume BACHELOT

Thèse de doctorat de Bioinformatique et biologie des systèmes

Dirigée par Antonin LAMAZIERE et Charlotte DUPONT

Présentée et soutenue publiquement le 24 juin 2024

Devant un jury composé de :

Pr. BACHELOT Anne, PU-PH

Pr. BOITRELLE Florence, PU-PH

Dr. ROBIN Geoffroy, MCU-PH, HDR

Pr. ACHOUR-CHNEIWEISS Nelly, PU-PH

Pr. TENENHAUS Arthur, PU

Pr. CHABBERT-BUFFET Nathalie, PU-PH

Dr. DUPONT Charlotte, MCU-PH, HDR

Pr. LAMAZIERE Antonin, PU-PH

Présidente

Rapporteur

Rapporteur

Examinatrice

Examineur

Membre

Directrice

Directeur

A Camille, pour son soutien indéfectible

Remerciements

Je tiens à remercier le **Professeur Antonin Lamazière** et le **Docteur Charlotte Dupont** de m'avoir encadré pour cette thèse et de m'avoir accordé leur confiance. Votre aide, vos conseils, vos connaissances et votre enthousiasme pour la recherche m'ont beaucoup apporté tout au long de l'élaboration de ce travail.

Ce travail n'aurait pas été possible sans le soutien du **Professeur Rachel Lévy**, qui m'a permis de consacrer le temps nécessaire à la réalisation de cette thèse, à côté de mes activités hospitalières, sans jamais interférer et en m'assurant des conditions de travail parfaites.

Mes remerciements vont également aux membres du jury. Merci à mes rapporteurs, le **Professeur Florence Boitrelle** et le **Docteur Geoffroy Robin**, de me faire l'honneur d'évaluer ce travail.

Un très grand merci au **Professeur Anne Bachelot** d'avoir accepté de présider ce jury de thèse, et sans qui ce travail n'aurait pas été possible.

Merci au **Professeur Nelly Achour-Chneiweiss** pour sa bienveillance, son soutien depuis le début de l'internat, et d'avoir accepté de participer au jury de cette thèse.

Je remercie chaleureusement le **Professeur Nathalie Chabbert-Buffer** de m'avoir fait l'honneur de suivre mon travail durant les trois dernières années et de participer à ce jury de thèse.

Mes remerciements vont également au **Professeur Arthur Tenenhaus** d'avoir accepté de juger ce travail et de me faire l'honneur d'apporter un œil expert sur la partie modélisation mathématique.

Je remercie également le **Professeur Jacques Young**, pour sa collaboration dans ce travail et ses enseignements.

Une immense merci à **Dominique Farabos**, pour ses conseils et sa précieuse aide tout au long de ce travail.

J'adresse également mes remerciements aux ingénieurs et techniciens de la plateforme de spectrométrie de masse du Centre de Recherche de Saint Antoine, à l'ensemble de l'équipe du service de métabolomique clinique (METOMICS), ainsi qu'aux **Professeurs Jean Fiet. Yves Le Bouc et Sophie Christin-Maitre.**

A toute l'équipe de biologie de la reproduction du centre de fertilité Tenon, biologistes, techniciens et agents, avec qui je travaille au quotidien : recevez mes remerciements les plus chaleureux pour votre soutien et vos encouragements.

Enfin, je remercie chaleureusement ma famille et mes amis pour leur soutien depuis le début de cette longue aventure.

Sommaire

<u>Remerciements</u>	2
<u>Sommaire</u>	4
<u>Productions scientifiques liées à la thèse</u>	6
<u>Articles publiés directement liés à la thèse</u>	6
<u>Articles publiés indirectement liés à la thèse</u>	6
<u>Communications orales</u>	6
<u>Communications affichées</u>	7
<u>Conférences invitées</u>	7
<u>Autre valorisation</u>	8
Brevets.....	8
<u>Liste des principales abréviations</u>	9
<u>Avant-propos</u>	11
<u>Introduction</u>	14
<u>Contexte médical : infertilité et troubles métaboliques et endocriniens</u>	14
Généralités.....	14
Infertilité inexplicée.....	16
Hyperandrogénie et syndrome des ovaires polykystiques.....	20
Recouvrement de situations inexplicées, SOPK et implication du mode de vie.....	26
<u>Sources de données pour les modèles</u>	27
<u>Méthodes appliquées : outils analytiques et mathématiques</u>	30
Données omiques et métabolomiques.....	30
Analytique et spectrométrie de masse.....	31
Outils mathématiques.....	35
<u>Objectifs de la thèse</u>	41
<u>Résultats I : Développement d'un score de machine learning basé sur une approche couple pour identifier et stratifier les couples infertiles inexplicés</u>	42
<u>Contexte scientifique et objectifs</u>	42
<u>Résumé des résultats</u>	42
Article 1 :.....	43
Article 2 :.....	45
<u>Éléments de discussion</u>	45
Sur le modèle.....	45
Sur les variables / VIP.....	46
Fragmentation de l'ADN spermatique – score.....	47
Forces et faiblesses.....	48
<u>Conclusion et perspectives</u>	49
<u>Article I</u>	51
<u>Article II</u>	60

<u>Résultats II : Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs hyperplasie congénitale des surrénales.....</u>	67
<u>Contexte scientifique et objectifs.....</u>	67
Importance de distinguer correctement SOPK vs NC21OHD : deux maladies de physiopathologie différentes, mais à présentation clinique identique	67
Objectif	68
<u>Résumé des résultats</u>	70
Construction et évaluation du modèle	70
Stéroïdes importants dans le modèle.....	71
<u>Éléments de discussion.....</u>	71
Modèle	71
Stéroïdes les plus discriminants	72
Forces et faiblesses.....	72
<u>Conclusion et perspectives</u>	73
<u>Article III.....</u>	75
<u>Discussion et Perspectives.....</u>	86
<u>Discussion et perspectives générales</u>	86
<u>Miniaturisation des prélèvements</u>	89
<u>Explorations métaboliques du SOPK : mécanismes moléculaires.....</u>	89
<u>Conclusion</u>	91
<u>Bibliographie</u>	92
<u>Annexe</u>	99
<u>Table des illustrations</u>	109

Productions scientifiques liées à la thèse

Articles publiés directement liés à la thèse

- **Machine learning approach to assess the association between anthropometric, metabolic and nutritional status and semen parameters.** Guillaume Bachelot, A Lamazière, R Lévy, C. Dupont . Asian journal of andrology avril 2024 DOI: 10.4103/aja20247
- **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, Yves Le Bouc, Jacques Young, Antonin Lamazière. Human Reproduction, Novembre 2022, doi: <https://doi.org/10.1093/humrep/deac254>
 - ⇒ Prix de Publication SFE (Société Française d'Endocrinologie) 2023 – Marseille 5 octobre 2023.
- **Development of a couple-based machine learning model to stratify subfertile patients with idiopathic infertility: a tool for reproductive medical guidance.** Guillaume Bachelot, Rachel Lévy, Céline Faure, Sébastien Czernichow, Charlotte Dupont and Antonin Lamazière. Scientific Reports December 2021 doi: 10.1038/s41598-021-03165-3

Articles publiés indirectement liés à la thèse

- **Plasma 21-deoxycortisone : a sensitive additive tool in 21-hydroxylase deficiency in newborns.** Jean Fiet*, Guillaume Bachelot*, Coumba Sow, Dominique Farabos, Marie-Noelle Dufourg, Christine Bellanne-Chantelot, Jacques Young, Muriel Houang, Antonin Lamazière. En révision mineur dans European Journal of Endocrinology en janvier 2024
- **A proof of concept of a machine learning algorithm to predict late-onset 21-hydroxylase deficiency in children with premature pubic hair.** Héléna Agnani, Guillaume Bachelot, Thibaut Eguether, Bettina Ribault, Jean Fiet, Yves Le Bouc, Irène Netchine, Muriel Houang, Antonin Lamazière. J Steroid Biochem Mol Biol. 2022 Mar doi: 10.1016/j.jsbmb.2022.106085

Communications orales

- Session scientifique Collège Médecine et biologie de la reproduction, Paris, 1er décembre 2022. **Combinaison de métabolomique et intelligence artificielle pour simplifier le diagnostic différentiel du SOPK versus Hyperplasie congénitale des surrénales.** Guillaume Bachelot
 - ⇒ Prix du collège.
- 38ème Congrès de la Société Française d'Endocrinologie, Nantes, 12 - 15 octobre 2022. **Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs bloc en 21-hydroxylase non classique.** Guillaume Bachelot

- 27èmes Journées de la Fédération Française d'Étude de la Reproduction (FFER 2022), Paris, 14 - 16 septembre 2022. **Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs bloc en 21-hydroxylase non classique.** [Guillaume Bachelot](#)
 ⇒ [Second Prix Communication Orale « Session Jeunes Chercheurs ».](#)
- Séminaire Internes du DMU, jeudi 22 avril 2021. **Développement d'un modèle de Machine Learning basé sur une « approche couple » pour stratifier les couples atteints d'infertilité idiopathique.** [Guillaume Bachelot](#)
- 25èmes Journées de la Fédération Française d'Étude de la Reproduction (FFER 2020), Clermont-Ferrand, 23-25 septembre 2020. **Développement d'un modèle de Machine Learning basé sur une « approche couple » pour stratifier les couples atteints d'infertilité idiopathique.** [Guillaume Bachelot.](#)
 ⇒ [Second Prix de la session recherche fondamentale.](#)

Communications affichées

- 28èmes Journées de la Fédération Française d'Étude de la Reproduction (FFER 2023), Arcachon, 27 - 29 septembre 2023. **Associations entre signature métabolique et nutritionnelle et paramètres spermatiques.** [Guillaume Bachelot](#)
- Congrès BIOMED-J 2022, Paris, 9 & 10 mars 2023. **Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs bloc en 21-hydroxylase non classique.** [Guillaume Bachelot](#)
- 65èmes Journées de l'innovation en biologie (JIB), Paris, 1-2 décembre 2022. **Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs bloc en 21-hydroxylase non classique.** [Guillaume Bachelot](#)
- 36th ESHRE (*“European Society of Human Reproduction and Embryology”*) Annual Meeting 5-8 Juillet 2020. **The InfertilScore, a couple-based machine learning model to stratify subfertile patients with idiopathic infertility : a new tool for reproductive medical guidance.** [Guillaume Bachelot](#)

Conférences invitées

- Journée France-Québec sur l'IA et l'alimentation le 10 avril 2024 à l'INAF (Institut sur la nutrition et les aliments fonctionnels) de l'Université Laval (Québec, Canada). Symposium : « **Nutrition, métabolisme et infertilité : quel apport de l'IA ?** ». [Guillaume Bachelot](#)
- Journée GynFoch, 7 et 8 mars 2024, Salons Hoche, Paris 8. « **Intelligence artificielle et infertilité** » [Guillaume Bachelot.](#)
- Colloque en Calade, Lyon, Vendredi 17 novembre 2023. **Score pour différencier juste sur dosages hormonaux par spectrométrie de masse les blocs en 21 hydroxylase des SOPK.**
- « Gala de la recherche du DMU BioGeM », 4 juillet 2023. **Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs bloc en 21-hydroxylase non classique**

- Journée BLEFCO des techniciens et biologistes, Paris, Vendredi 24 mars 2023. **Intelligence artificielle au laboratoire d'AMP - Partage d'expériences.** [Guillaume Bachelot.](#)
- Dietecom : Journées de nutrition pratique, Grand Rex, Paris, Vendredi 10 mars 2023. **Alifert : nutrifertilité et intérêt de l'intelligence artificielle.** [Guillaume Bachelot.](#)
- WebMasterClass Université de Laval (Canada) – Sorbonne Université : Santé endocrinienne et cardiométabolique, distanciel les 22-23 juin 2022. **Développement d'un modèle de Machine Learning basé sur une « approche couple » pour stratifier les couples atteints d'infertilité idiopathique.** [Guillaume Bachelot.](#)
- Séminaire Intelligence Artificielle du DMU BioGeM AP-HP.SU, 4 février 2021. **Utilisation d'algorithmes de Machine Learning dans le diagnostic des hyperplasies congénitales des surrénales.** [Guillaume Bachelot.](#)

Autre valorisation

Brevets

- **Co-inventeur du brevet n° PCT/FR2022/050759** déposé le **21 avril 2022** par la Société d'Accélération de Transfert de Technologies LUTECH (SATT LUTECH) aux noms de Sorbonne Université et de l'Assistance Publique - Hôpitaux de Paris (AP-HP). « *Fourniture d'une évaluation d'un risque médical de déficience 21-hydroxylase de forme non classique* »
- **Co-inventeur du brevet n°PCT/FR2020/051185** déposé le **3 juillet 2020** (« Alifert ») par la Société d'Accélération de Transfert de Technologies LUTECH (SATT LUTECH) : « *Système, procédé et programme d'ordinateur de fourniture d'une évaluation d'un risque médical, procédé d'obtention d'un modèle pour le système, méthode d'évaluation d'un risque médical et compléments nutritionnels* » aux noms de Sorbonne Université et de l'Assistance Publique - Hôpitaux de Paris (APHP).

Liste des principales abréviations

ADN : Acide Désoxyribonucléique

ACP : Analyse en Composantes Principales

ACTH : Hormone Corticotrope

AMH : Hormone Anti-Müllérienne

AMP : Assistance Médicale à la Procréation

DHEA : Dehydroepiandrosterone

FIV : Fécondation In Vitro

FSH : Hormone Folliculo-Stimulante

HCS : Hyperplasie Congénitale des Surrénales

HCS-NC : Hyperplasie Congénitale des Surrénales forme non classique

HDL : High Density Lipoprotein Cholestérol

IA : Intelligence Artificielle

IUI : Insémination Intra-Utérine

IMC : Indice de Masse Corporelle

LC-MS/MS : Chromatographie en Phase Liquide couplée à la Spectrométrie de Masse

ML : Machine Learning

NC21OHD : Nonclassical adrenal hyperplasia due to 21-hydroxylase deficiency

NIPALS : Non-linear Iterative Partial Least Squares

OPLS-DA : Orthogonal Partial Least Squares-Discriminant Analysis

17OHP : 17-Hydroxyprogestérone

PLS : Partial Least Squares ou Projection to Latent Structure

PHRC : Programme Hospitalier de Recherche Clinique

SHBG : Sex Hormone Binding Globulin

SFE : Société Française d'Endocrinologie

SOPK : Syndrome des Ovaires Polykystiques

TSH : Thyroïdostimuline

TUNEL : Terminal dUTP Nick-End Labeling

VIP : Variable Importance in the Projection

Avant-propos

L'infertilité représente un enjeu majeur de santé publique, affectant au moins 50 millions de couples à travers le monde. Dans 10 à 25 % des cas, l'infertilité ne peut être attribuée à un défaut spécifique et aucune étiologie n'est identifiée après un bilan complet des deux partenaires. Le caractère inexplicé est la conséquence d'une absence d'identification de cause évidente, mais un certain nombre de travaux suggèrent un terrain métabolique complexe et difficile à interpréter. En effet, le mode de vie moderne, en particulier les habitudes alimentaires et le manque d'activité physique, ont des conséquences métaboliques (diabète, dyslipidémie, obésité, inflammation chronique) qui sont aujourd'hui largement reconnues pour leur impact négatif sur les fonctions reproductives. D'un point de vue évolutionniste, ces troubles métaboliques pourraient résulter d'un conflit entre un patrimoine génétique façonné par des millénaires de restriction alimentaire et d'exercice physique, et un environnement moderne caractérisé par une surabondance de nourriture et un mode de vie sédentaire. D'autres situations métaboliques complexes, telles que le syndrome des ovaires polykystiques (SOPK), peuvent également avoir des répercussions sur la fertilité. Le SOPK est une pathologie endocrinienne très fréquente, pas totalement comprise, associée à des troubles métaboliques similaires à ceux observés chez les patients présentant une infertilité inexplicée. Sa définition et ses critères de diagnostic sont régulièrement revus et débattus entre les experts. Cet éventail de pathologies métaboliques complexes, comprenant vraisemblablement l'infertilité inexplicée et le syndrome des ovaires polykystiques, et dont les contours ne sont pas encore clairement définis, semblent étroitement corrélées et dépendantes de facteurs comportementaux, de mode de vie, nutritionnels et métaboliques, qui constituent autant de leviers à exploiter en première intention dans le cadre d'une prise en charge thérapeutique.

L'objectif principal de ce travail de recherche est d'explorer ces conditions complexes de l'infertilité, difficiles à identifier et à caractériser, dans lesquelles le(s) métabolisme(s) joue(nt) un rôle essentiel dans la physiopathologie. Notre but est de développer des outils diagnostiques et de suivi en utilisant des approches omiques et des méthodes basées sur l'intelligence artificielle mise en place dans le laboratoire (Pr Lamazière Antonin), dans le but à plus long terme d'opérer un changement de sémiologie pour les troubles métaboliques et endocriniens de la fertilité. Deux applications ont été menés en parallèles. La première, qui s'appuie sur les données du programme hospitalier de recherche clinique (PHRC) ALIFERT (Pr Lévy Rachel et Dr Dupont Charlotte), porte sur l'infertilité inexplicée. Elle a conduit au

développement d'un score de couple permettant d'identifier et de stratifier les couples confrontés à cette condition. La seconde application, axée sur les aspects endocriniens, s'est penchée sur les hyperandrogénies chez les femmes, en se concentrant notamment sur le diagnostic différentiel entre le SOPK et l'hyperplasie congénitale des surrénales due à un blocage de la 21-hydroxylase (Pr Bachelot Anne).

Dans ce travail, notre hypothèse est donc que l'application d'outils innovants issus de l'intelligence artificielle pourrait améliorer le traitement de données médicales hétérogènes. Cela permettrait ainsi une meilleure appréhension de la complexité sous-jacente à ces pathologies, et faciliterait leur identification, leur stratification et leur suivi.

La rédaction de cette thèse se présente sous un format article et s'ouvre sur une introduction bibliographique visant à définir le contexte médical, à discuter des données traitées et à présenter les méthodes utilisées pour développer les outils diagnostiques. Dans le contexte médical, une attention particulière sera accordée à la nature complexe et difficilement identifiable des pathologies d'exclusion telles que l'infertilité inexplicée et le SOPK. Le choix des données traitées et modélisées, principalement axées sur les aspects métaboliques, endocriniens et nutritionnels, sera justifié par des travaux antérieurs portant sur les mécanismes moléculaires de ces pathologies. Enfin, un état des lieux sera dressé et les différentes techniques analytiques et mathématiques, incluant la spectrométrie de masse et l'intelligence artificielle, utilisées dans cette thèse seront décrites. Les résultats seront ensuite présentés en deux parties, selon les deux applications conduites en parallèle.

La première concerne le développement et l'évaluation d'un modèle d'apprentissage automatique permettant d'identifier et de mesurer l'implication des facteurs nutritionnels et métaboliques dans l'infertilité inexplicée. Ce travail, reposant sur les données rétrospectives du PHRC ALIFERT, visait à exploiter des données médicales d'origine hétérogène pour créer un profil "couple" complet, et exploiter ce profil par des méthodes d'apprentissage pour identifier les couples selon leur statut fertile/infertile et mettre en évidence l'importance de certains marqueurs métaboliques. Une validation rétrospective et prospective sur des cohortes américaines (USA et Canada) est en cours.

Forts de ces premiers résultats, notre exploration s'est ensuite concentrée, dans une seconde partie, sur la voie de biosynthèse des stéroïdes pour l'exploration des hyperandrogénies chez la femme, dont fait partie le SOPK. Cette initiative est motivée par le fait que l'excès

d'androgènes semble être directement lié au développement de complications métaboliques telles que la résistance à l'insuline, l'intolérance au glucose, la dyslipidémie ou encore le diabète, selon un processus et des interactions complexes. De plus, les hormones stéroïdiennes, en particulier les stéroïdes sexuels, jouent un rôle important dans les fonctions reproductives. En appliquant des techniques de modélisation mathématique à une cartographie de la voie de biosynthèse des stéroïdes (acquisition des concentrations de 15/22 espèces moléculaires par chromatographie en phase liquide couplée à la spectrométrie de masse (LC-MS/MS)), la première étape de ce travail a consisté à simplifier le diagnostic différentiel du SOPK avec l'hyperplasie congénitale de forme non classique par bloc en 21-hydroxylase (NC21OHD), permettant ainsi de s'affranchir du test dynamique au Synacthène. Une validation multicentrique et une extension des capacités de stratification du modèle sont actuellement en cours. Enfin, les résultats obtenus et l'importance de cette exploration multidisciplinaire du métabolisme seront discutés, en vue de développer d'autres outils diagnostiques, de suivi et, *in fine*, de changer la sémiologie endocrinienne des infertilités pour rationaliser la prise en charge basée sur le mode de vie.

En définitive, l'infertilité ne représente pas seulement un enjeu de santé, mais également un enjeu social, économique et démographique majeur. Cette complexité inhérente à l'infertilité impose une approche transdisciplinaire combinant différentes perspectives et domaines de connaissance pour aborder pleinement toutes ses dimensions. Ces différentes perspectives peuvent être appréhendées de manière intégrative et holistique à l'aide de nouveaux outils issus de l'intelligence artificielle, offrant ainsi un espoir et une opportunité pour une compréhension et une interprétation plus fine de la physiologie, de la physiopathologie et, plus généralement, pour améliorer la prise en charge des patients à travers une médecine plus prédictive, préventive et personnalisée.

Introduction

Contexte médical : infertilité et troubles métaboliques et endocriniens

Généralités

L'infertilité, définie par une absence de grossesse après 12 mois de rapports sexuels réguliers et non protégés, ne constitue pas seulement un enjeu de santé, mais aussi un véritable défi social, économique et démographique ¹. L'infertilité est une question à la fois intime et sociétale, et rares sont les sujets qui présentent autant d'implications individuelles et collectives. En France, elle touche directement 3,3 millions de personnes, un chiffre qui ne cesse d'augmenter ².

Au moment où la fréquence de l'infertilité masculine et féminine ne cesse d'augmenter à un rythme particulièrement alarmant, en particulier au cours des vingt dernières années, il devient incontournable d'identifier, de prévenir et de traiter l'infertilité de la manière la plus adaptée et précoce possible ³. Les enjeux portent à la fois sur l'avenir de millions de couples confrontés à ces situations, mais aussi, à plus long terme, sur la préservation de l'espèce humaine. Dans cette perspective, une stratégie nationale de prévention et de recherche sur l'infertilité a été engagée en France en 2021, et plus récemment, le plan infertilité, qui prévoit de renforcer la lutte contre les perturbateurs endocriniens, de développer la recherche sur l'assistance médicale à la procréation (AMP) et d'ouvrir 100 000 places en crèches ou chez les assistantes maternelles d'ici à 2027.

L'infertilité est une maladie aux causes multiples, médicales, environnementales et sociétales (elles-mêmes souvent combinées), avec des implications collectives, sociales et politiques. En France, de même que dans tous les pays industrialisés, l'augmentation de l'infertilité est avant tout la conséquence du recul de l'âge de la première maternité, qui atteint environ 30 ans aujourd'hui. La fertilité diminuant progressivement à partir de 35 ans, la maternité tardive augmente mécaniquement le risque d'infertilité et résulte d'un certain nombre de facteurs sociétaux. Des facteurs environnementaux sont par ailleurs impliqués dans l'augmentation de l'infertilité. Une méta-analyse réalisée en 2021 a révélé une baisse de plus de 50 % de la concentration en spermatozoïdes chez les hommes du monde entier entre 1973 et 2020 ³. Ce phénomène apparaît -entre autres- lié à l'exposition régulière à des perturbateurs

endocriniens. Par ailleurs, des études récentes montrent l'impact négatif des modes de vie occidentaux sur la fertilité des hommes et des femmes, en particulier pendant la période préconceptionnelle, c'est-à-dire les 6 mois précédant la grossesse : consommation de tabac ou de cannabis, obésité, troubles du comportement alimentaire...⁴⁻⁷

Outre les causes environnementales et sociétales, l'infertilité peut résulter de causes médicales. Dans les trois quarts des cas, l'infertilité est d'origine masculine (troubles de la spermatogenèse), féminine (pathologies tubaires, troubles de l'ovulation ou endométriose), ou implique les deux sexes^{8,9}. Dans 10 à 25 % des cas, l'infertilité n'est pas imputable à un défaut spécifique de l'un ou l'autre sexe : on parle alors d'infertilité inexplicée (**Figure 1**).

RÉPARTITION DES CAUSES D'INFERTILITÉ

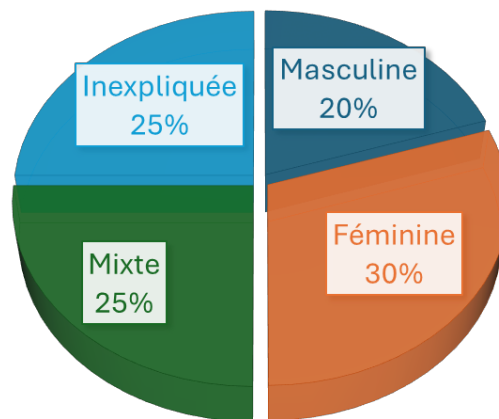


Figure 1 : Répartition schématique des origines/causes de l'infertilité dans le couple (adapté de Thonneau *et al.*⁸).

Parmi les causes d'infertilité masculine, l'insuffisance testiculaire se distingue nettement comme la plus fréquente, entraînant diverses altérations de la quantité ou de la qualité des spermatozoïdes. Ces altérations peuvent varier en intensité, allant de l'oligozoospermie (diminution du nombre de spermatozoïdes) à l'azoospermie (absence totale de spermatozoïdes dans l'éjaculat), en passant par la tératozoospermie (anomalies morphologiques). Elles peuvent être d'origine constitutionnelle, acquises à la suite d'une infection, d'un traumatisme ou secondaires à des traitements tels que la chimiothérapie, dont les effets peuvent être irréversibles. Dans de nombreux cas, aucune cause spécifique n'est identifiée pour expliquer les altérations du sperme associées à l'infertilité masculine.

Chez la femme, les principales causes sont soit anatomiques, soit hormonales. Parmi les anomalies anatomiques, on peut citer les altérations utérines telles que les polypes, les fibromes et les synéchies, ainsi que les altérations des trompes de Fallope et l'endométriose. Quant aux causes hormonales, elles comprennent le SOPK, développé de façon plus précise si dessous, qui peut entraîner des troubles de l'ovulation ou une absence de production d'ovocytes matures, ou encore l'insuffisance ovarienne.

Infertilité inexplicée

Absence de définition universelle :

Comme évoqué précédemment, dans 10 à 25 % des cas, l'infertilité n'est pas imputable à un défaut spécifique et le diagnostic est posé lorsqu'aucune anomalie des appareils reproducteurs féminin et masculin n'est identifiée. On parle alors d'infertilité inexplicée¹⁰⁻¹². Cette absence d'identification peut être la conséquence 1) d'une anomalie inexplorée (problème de l'absence de définition universelle) ou 2) d'une anomalie inconnue et/ou que les outils diagnostiques actuels ne permettent pas de déceler (problème de l'identification de nouveaux biomarqueurs).

L'infertilité inexplicée est donc inévitablement et par définition un diagnostic d'exclusion, faisant suite aux investigations standard. Cette notion d'investigations "standard" est importante et se trouve au cœur de la complexité de la définition de ce type d'infertilité¹³. En effet, une revue systématique proposée par Raperport *et al.* a montré des disparités extrêmement importantes dans les critères diagnostiques utilisés dans les études pour qualifier l'infertilité d'"inexplicée"¹⁴. Par exemple, seules 35.7% des études (soit 92/258) incluses exigeaient la présence des 4 critères suivants : 1) un spermogramme normal, 2) des cycles ovulatoires réguliers 3) une cavité utérine normale et 4) la présence d'au moins 1 trompe perméable. Point particulièrement déterminant pour la suite de ce travail, dans 79,5 % des études seulement, la régularité de l'ovulation devait être démontrée, 16,1 % d'entre elles n'ayant pas précisé la méthode d'évaluation. En ce qui concerne le profil hormonal des patientes, un profil endocrinien "normal" pour les femmes était requis dans seulement 44,2 % (114/258) des études, dont 103/258 spécifiaient les hormones évaluées. Plus intéressant encore, seules 17 études incluaient la testostérone dans l'évaluation des patientes, 10 incluaient la dehydroepiandrosterone (DHEA) et 4 incluaient la « *sex hormone binding globulin* » (SHBG). Aucune information sur la présence de signes cliniques d'hyperandrogénie n'a été mentionnée.

Le dosage de l'hormone anti-Müllérienne (AMH), désormais une alternative à l'échographie dans la définition des critères de Rotterdam lors de la mise à jour du 17 août 2023, était présent dans seulement 7 études ¹⁵.

Dans la mesure où 50 % des patientes atteintes de SOPK ont des cycles réguliers et que la recherche d'une hyperandrogénie, clinique ou biologique, apparaît peu fréquente dans l'inclusion des patientes qualifiées d'inexpliquées, la présence de SOPK demeure très difficile à écarter formellement dans les études portant sur l'infertilité inexpliquée. Ces résultats soulignent les difficultés majeures rencontrées dans l'analyse et la comparaison des études de la littérature en ce qui concerne l'infertilité inexpliquée, en raison de l'hétérogénéité des critères diagnostiques, mais aussi des méthodes d'évaluation de ces critères.

En France, Rolland *et al.* ont interrogé 195 médecins, majoritairement des gynécologues, confirmant l'absence de consensus tant sur les tests diagnostiques utilisés que sur la stratégie thérapeutique ¹⁶. L'AMH était prescrite par 87,6 % des répondants. Une valeur faible, en deçà d'un seuil spécifique, pouvait écarter le diagnostic d'infertilité inexpliquée, mais aucune information n'était fournie concernant les valeurs élevées. Récemment, un groupe d'experts européens a émis des recommandations concernant les critères diagnostiques, les méthodes d'évaluation ainsi que sur la prise en charge thérapeutique de l'infertilité inexpliquée ¹². Selon ces recommandations, le diagnostic peut être établi en présence de cycles réguliers ¹⁷, d'une perméabilité des trompes ¹⁸⁻²⁰, d'une cavité utérine normale ²¹⁻²³ et d'un spermogramme normal ²⁴.

Ainsi, comme démontré notamment par Raperport *et al.* et Rolland *et al.*, l'absence de définition universelle de l'infertilité inexpliquée, due à une variabilité des critères d'inclusion dans les études, constitue inévitablement un obstacle au développement de lignes directrices diagnostiques et thérapeutiques cohérentes. De même, l'incapacité à identifier un dysfonctionnement dans certains cas nécessite éventuellement le recours à des tests complémentaires.

Test, biomarqueurs, et critères additionnels :

Outre l'absence de bilan standardisé et d'une définition universelle, la détection d'une anomalie inconnue et/ou non identifiable par les outils diagnostiques actuels justifie la nécessité de rechercher et d'identifier de nouveaux biomarqueurs.

Au sein des études sur l'infertilité inexplicée, Raperport rapportait que parmi les 258 études examinées, 32,2% d'entre elles (soit 83 études) incluaient une limite d'âge pour les femmes, avec 82/83 établissant une limite supérieure et 39/83 établissant une limite inférieure. De plus, l'Indice de Masse Corporelle (IMC) de la femme était inclus comme critère dans 30 études, soit 11,6%, avec différents seuils utilisés pour exclure le diagnostic d'infertilité inexplicée (25, 26, ... jusqu'à 35). En ce qui concerne les tests supplémentaires chez l'homme, seule une étude incluait la mesure de la fragmentation de l'ADN (Acide Désoxyribonucléique) des spermatozoïdes comme critère d'inclusion. Dans les recommandations émises par la société européenne, l'IMC de la femme semblait jouer un rôle crucial en pré-conception, tandis que la fragmentation de l'ADN spermatique n'était pas recommandée dans la pratique courante²⁵⁻²⁷. Enfin, les tests systémiques et circulants de mesure du stress oxydatif, tant dans le sperme chez l'homme que dans la circulation sanguine chez la femme, n'étaient pas non plus recommandés à l'heure actuelle²⁸⁻³¹.

La littérature nous montre qu'il est difficile d'établir des seuils pour des paramètres standard tels que l'âge, l'IMC, l'AMH, etc., et que les quelques tests supplémentaires présentés et évalués ne sont pas non plus satisfaisants. Une de nos hypothèses est que, considérés individuellement, ces marqueurs ne fournissent pas suffisamment d'informations, ce qui explique le manque de consensus sur d'éventuels seuils. En revanche, comme nous le verrons dans les premiers résultats de cette thèse, la combinaison de certains de ces marqueurs (cliniques et biologiques) dans un score pourrait permettre une meilleure identification et/ou stratification de ces couples présentant une infertilité inexplicée.

Prise en charge thérapeutique :

L'une des conséquences directes de cette absence de consensus et de définition universelle est la difficulté d'établir des bonnes pratiques et d'harmoniser la prise en charge, qui devient alors le plus souvent empirique³². La prise en charge thérapeutique et le choix de la technique d'AMP la plus adaptée sont considérés comme l'une des 10 priorités de la recherche sur l'infertilité³³. Les sociétés savantes ont chacune émis leurs propres recommandations^{10,11}.

Pour l'*American Society of Reproductive Medicine*, qui confirme le caractère empirique de la prise en charge thérapeutique, il est recommandé de commencer par 3 ou 4 cycles de stimulation ovarienne avec Insémination intra-utérine (IIU), puis de passer à la Fécondation *in vitro* (FIV) en cas d'échec¹¹. Par ailleurs, la stimulation simple de l'ovulation avec des rapports

sexuels programmés et l'IUI sur un cycle naturel (sans simulation) sont déconseillés. Au Canada, les recommandations vont également dans le sens d'une prise en charge progressive et graduelle, en commençant par les IUI avec stimulation avant de passer à la FIV, cette dernière étant plus efficace mais plus invasive et plus coûteuse¹⁰. L'inclusion dans l'équation et dans le choix du traitement de l'âge de la femme et de la durée de l'infertilité, dans une optique d'individualisation, a également été proposée. Dans l'étude française de Rolland *et al.*, aucun consensus n'a été trouvé entre démarrer par quelques cycles d'IUI avant de passer à la FIV ou opter directement pour la FIV. Par ailleurs, une récente méta-analyse de 2023 comparant la FIV à l'IUI avec stimulation hormonale n'a pas montré que le recours direct à la FIV se traduisait par l'obtention d'une naissance vivante plus rapidement que le recours aux IUI³⁴. Enfin, les recommandations de la Société européenne confirment que la réalisation de cycles d'IUI après stimulation doit être considéré comme traitement de première ligne. Cependant, la FIV en première intention doit être envisagée en fonction de certains profils du couple (âge de la femme, durée de l'infertilité, etc.). Une prise en charge comportementale, comprenant une alimentation saine et une activité physique régulière (habitudes de vie saines), est également recommandée.

Comme pour l'établissement du diagnostic, nous sommes confrontés ici à la même difficulté d'interprétation de la littérature concernant la stratégie thérapeutique optimale dans le contexte d'infertilité inexplicée. Malgré l'hétérogénéité des populations incluses dans les études, les sociétés savantes, tant européennes qu'américaines, semblent converger vers une prise en charge graduelle, commençant par l'IUI avec stimulation hormonale avant le passage en FIV en cas d'échec. L'intégration des facteurs environnementaux et du mode de vie, comme la nutrition et l'activité physique, semble être également important dans la prise en charge des patients avec une infertilité inexplicée.

Finalement, l'infertilité inexplicée représente un immense défi, tant dans l'établissement du diagnostic que dans la prise en charge thérapeutique. Une harmonisation du bilan de fertilité permettrait de définir de manière consensuelle le caractère "inexpliqué" de la pathologie et d'établir un véritable diagnostic d'exclusion. L'exploration et la validation de tests additionnels permettraient également une meilleure compréhension et l'identification de causes actuellement inconnues. Ainsi seulement, des recommandations claires sur la stratégie thérapeutique pourront être établies.

Hyperandrogénie et syndrome des ovaires polykystiques

Hyperandrogénie chez la femme : définition

A côté des infertilités inexplicables, un autre groupe de pathologie voit sa prévalence augmenter ces dernières années et dont les contours sont là aussi mal définis. Il s'agit de certains troubles de l'ovulation comme le syndrome des ovaires polykystiques, et plus généralement des hyperandrogénies. L'hyperandrogénie chez la femme, fréquemment associée à des troubles du cycle et à une infertilité, est l'un des motifs de consultation les plus fréquents en médecine et endocrinologie de la reproduction. Elle est définie par un excès d'androgènes circulants. Ces androgènes circulants sont principalement la testostérone, la Δ 4-androstènedione, la 5 α -dihydrotestostérone et la DHEA ainsi que son sulfate DHEAS.

Cet excès d'androgènes se manifeste habituellement cliniquement par l'accentuation de la pilosité à topographie masculine, et de façon moins spécifique par l'acné et l'alopécie. Sa prévalence est comprise, selon les études, entre 6 et 12% chez les femmes en âge de procréer¹⁵, et son incidence est en augmentation et suit l'épidémie des troubles métaboliques et de l'obésité. Chez la femme, les androgènes sont sécrétés par deux glandes : les ovaires (testostérone essentiellement mais également Δ 4-androstènedione) et les corticosurrénales (DHEAS principalement). En conséquence, plusieurs situations peuvent être à l'origine d'une hyperandrogénie : une maladie ovarienne ou surrénalienne, entraînant une sécrétion excessive d'androgènes, ou, plus rarement, un apport exogène d'androgènes dans le cadre du dopage sportif par exemple.

Différentes causes d'hyperandrogénies

Devant un tableau clinique d'hyperandrogénie, deux pathologies sont couramment évoquées en premier lieu : le SOPK, pathologie multifactorielle d'origine ovarienne, et l'hyperplasie congénitale des surrénales (HCS) de forme non classique ou à révélation tardive (HCS-NC), pathologie d'origine surrénalienne, causée par un bloc enzymatique dans la voie de biosynthèse des stéroïdes. D'autres affections ovariennes plus complexes peuvent également conduire à un excès d'androgènes. Il peut s'agir d'une résistance à l'insuline, comme observé dans le HAIR-AN syndrome (« Hyperandrogenism-Insulin Resistance-Acanthosis Nigricans Syndrome ») ou encore de tumeurs virilisantes, souvent bénignes, difficiles à différencier de l'hyperthécose ovarienne. Comme nous le verrons par la suite, bien que le bloc enzymatique en

21-hydroxylase soit responsable de 95% des HCS, d'autres blocs, beaucoup plus rares, comme le bloc en 11-hydroxylase ou en 17-hydroxylase existent ³⁵. Les tumeurs virilisantes de la surrenales sont quant à elles très rares. Les hirsutismes idiopathiques et les hyperandrogénies associées à une autre endocrinopathie (comme le syndrome de Cushing, l'hyperprolactinémie ou l'acromégalie) représentent d'autres causes d'hyperandrogénies.

Démarche diagnostique devant une hyperandrogénie et enjeu de santé publique

Face à un tableau clinique d'hyperandrogénie, un dosage de testostérone est dans un premier temps nécessaire pour confirmer le diagnostic. Ce prélèvement doit se faire suivant des règles strictes : en phase folliculaire du cycle menstruel, à jeun, et à distance de tout traitements hormonaux ou corticoïdes. Ce sont ensuite les recommandations de la Société Française d'Endocrinologie (SFE) et son algorithme décisionnel qui permettent d'orienter vers une étiologie, qui permettra une prise en charge adaptée ³⁶. Ces recommandations, qui datent de 2010, sont présentés dans la figure 2 :

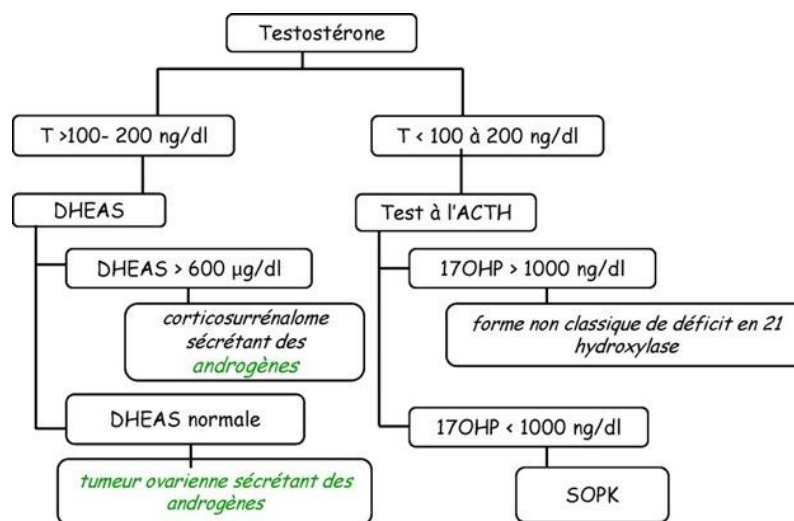


Figure 2 : Algorithme décisionnel d'après les recommandations de la SFE pour l'exploration des hyperandrogénies, à partir d'un résultat pathologique (élevé) de testostérone plasmatique (d'après M. Pugeat *et al.* ³⁶).

Les recommandations précisent par ailleurs que : « Si la testostérone est simplement au-dessus des valeurs normales, le diagnostic le plus probable est le syndrome des ovaires polykystiques (SOPK). Mais, une forme non classique de déficit en 21-hydroxylase doit être recherchée (dosage de la 17OHP) et selon le contexte clinique une maladie de Cushing doit être éliminée. »

D'après l'algorithme décisionnel du consensus de la SFE pour l'exploration des hyperandrogénies chez la femme, c'est donc le dosage de la 17-hydroxy-progesterone (17OHP) après stimulation à l'hormone corticotrope (ACTH) qui permet de discriminer et distinguer le SOPK (quand [17OHP] < 10 ng/mL) d'une forme non classique de l'HCS par déficit en 21-hydroxylase de manière formelle (quand [17OHP] > 10 ng/mL)³⁷. La 17OHP est un des substrats de la 21-hydroxylase, et est donc augmentée en cas de bloc enzymatique et donc d'HCS (accumulation des espèces moléculaires en amont du bloc, et baisse des produits). De plus, le diagnostic de SOPK, qui reste un diagnostic d'exclusion, repose sur les critères de Rotterdam (9,10). Ces derniers imposent d'écarter les diagnostics différentiels d'hyperandrogénies (et en particulier la NC21OHD) et de troubles du cycle menstruel avant de poser de manière formelle le diagnostic de SOPK. Comme nous le verrons par la suite, l'utilisation de la 17OHP comme stratégie pour distinguer SOPK / NC21OHD présente un certain nombre de limites (analytique, choix du seuil, double origine ovarienne/surrénalienne, nécessité de l'épreuve dynamique...).

De ce fait, et compte tenu de la forte prévalence du SOPK (plus de 10% des femmes de la population générale), la question du diagnostic différentiel SOPK / NC21OHD est un enjeu majeur de santé publique, que nous développerons plus en détails par la suite et qui fait l'objet d'une des publications de cette thèse.

Syndrome des ovaires polykystiques

Le SOPK est l'un des troubles endocriniens et métaboliques les plus courants chez les femmes pré-ménopausées, et touche environ 10% des femmes en âge de procréer^{15,38,39}. Hétérogène par nature, le SOPK est défini par une combinaison de signes et de symptômes d'excès d'androgènes et de dysfonctionnement ovarien en l'absence d'autres diagnostics spécifiques. L'étiologie de ce syndrome reste largement inconnue, mais des preuves de plus en plus nombreuses suggèrent que le SOPK pourrait être un trouble multigénique complexe avec de fortes influences épigénétiques et environnementales, comprenant entre autres des facteurs alimentaires et de facteurs du mode de vie. Le SOPK est fréquemment associé à une adiposité abdominale, une résistance à l'insuline, l'obésité, des troubles métaboliques et des facteurs de risque cardiovasculaires.

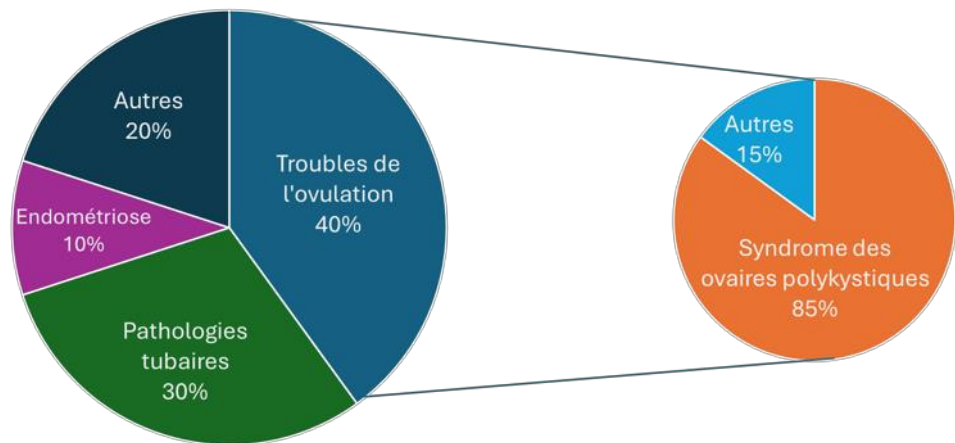


Figure 3 : Principales causes d'infertilités féminines – places des troubles ovulatoires et du SOPK, première cause d'infertilité chez la femme (adapté de ⁴⁰).

Les critères diagnostiques et la prise en charge du SOPK ont très récemment été revus par un consensus international, les principales nouveautés étant 1) l'élargissement des critères diagnostiques, avec l'inclusion de l'AMH comme alternative à l'échographie chez les adultes, 2) la meilleure prise en compte des spécificités du SOPK, notamment les facteurs de risque métaboliques, cardiovasculaires et psychologiques et 3) l'intégration et le renforcement de la notion du mode de vie dans la prise en charge du SOPK ¹⁵.

Les critères diagnostiques, basés sur les critères de Rotterdam, sont donc actuellement les suivants :

1. Hyperandrogénisme biologique ou clinique ;
2. Dysfonction ovulatoire ;
3. Ovaires polykystiques ou AMH élevée.

Par ailleurs et comme mentionné précédemment, toute autre cause d'hyperandrogénie doit être écartée, comme les tumeurs ovariennes ou surrénaliennes, le syndrome de Cushing ou encore l'hyperplasie congénitale des surrénales.

Algorithme diagnostique pour le syndrome des ovaires polykystiques (SOPK)

Étape 1 : Cycles irréguliers + hyperandrogénie clinique

→ (exclure les autres causes)* = diagnostic

Étape 2 : En l'absence d'hyperandrogénie clinique

→ Recherche d'une hyperandrogénie biochimique (exclure les autres causes)* = diagnostic

Étape 3 : Si SEULEMENT cycles irréguliers OU hyperandrogénie

→ Demander une échographie pour PCOM*, si positive (exclure d'autres causes)* = diagnostic

Figure 4 : Algorithme diagnostique pour le SOPK, d'après les recommandations internationales 2023 (adapté de Teede *et al.* ¹⁵). *Exclusion d'autres causes = thyroïdostimuline (TSH), prolactine, 17OHP, hormone folliculo-stimulante (FSH) ou, si la situation le justifie, exclusion d'autres causes (par exemple, syndrome de Cushing, tumeurs surrenaliennes).

En ce qui concerne la détection de l'hyperandrogénie biologique, les recommandations préconisent l'utilisation de la LC-MS/MS à la place des méthodes immunoanalytiques traditionnelles (nous aborderons cette partie ultérieurement dans le manuscrit). Au-delà de la testostérone et même de la 5 α -dihydrotestostérone, les implications en physiologie et pathologie humaines des androgènes 11-oxygénés d'origine surrenalienne récemment identifiés, en particulier la 11-cétotestostérone, sont également discutées ^{41,42}. Cependant, des recherches supplémentaires sont nécessaires pour établir des seuils sur des populations plus larges. L'AMH est désormais proposée en alternative à l'échographie comme critère diagnostique. Étant donné les variations possibles en fonction de l'âge, de l'IMC, de la période du cycle menstruel et de la prise de contraceptifs, aucun seuil n'a été défini par consensus international. Chaque laboratoire doit établir ses propres normes en prenant en compte tous ces facteurs ainsi que la méthode analytique utilisée.

Une fois le diagnostic établi, il est recommandé d'évaluer et de prendre en charge les risques d'infertilité, de maladies cardiovasculaires, métaboliques et psychologiques, étant donné le risque accru de développer à moyen ou long terme une résistance à l'insuline voire un diabète et des dysfonctionnements métaboliques. Cette évaluation repose notamment sur l'examen du statut glycémique, du profil lipidique et de la pression artérielle, indépendamment de l'âge ou

de l'IMC. En effet, l'insulino-résistance joue un rôle central dans la physiopathologie du SOPK. Cette évaluation est justifiée par le caractère hétérogène du SOPK en termes de phénotypes, de manifestations cliniques et de ces conséquences métaboliques, particulièrement présentes dans les formes avec hyperandrogénie. En effet, comme nous verrons dans la partie suivante, l'excès d'androgènes semble être l'un des éléments majeurs impliqués dans le risque métabolique associé au SOPK.

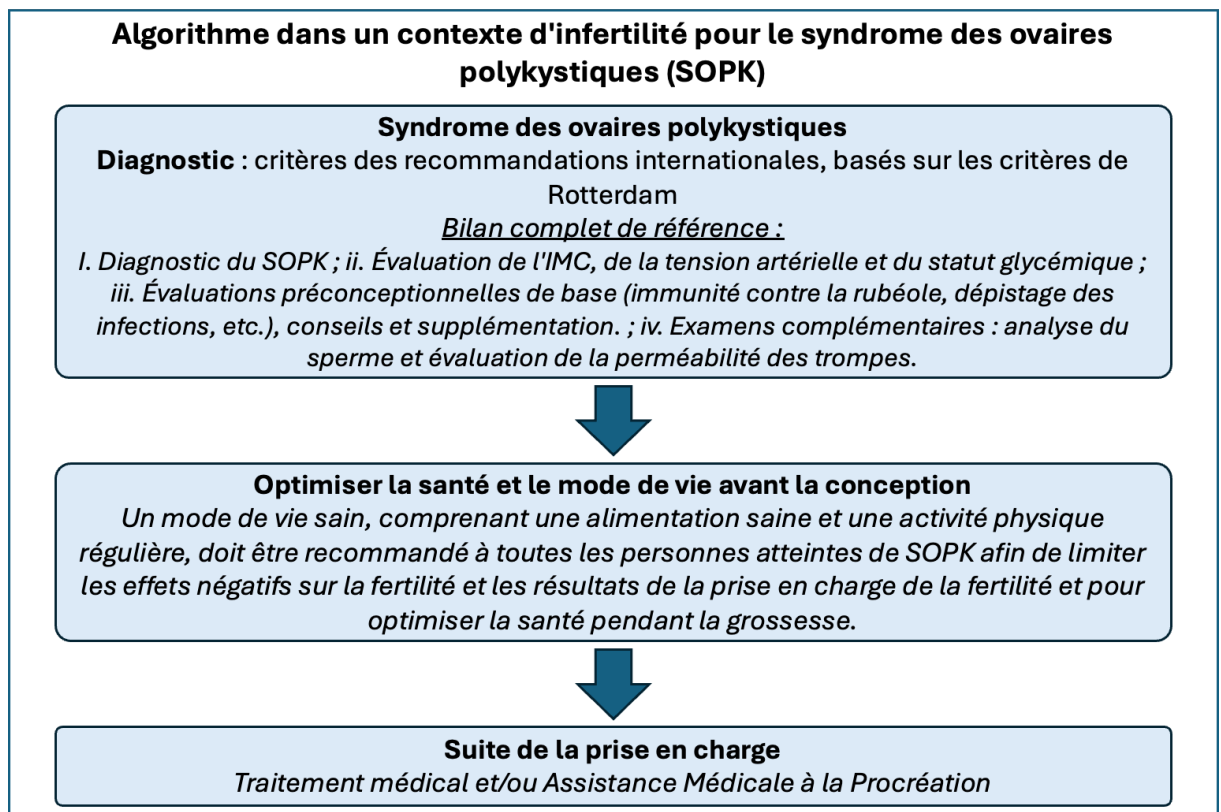


Figure 5 : Place centrale de l'optimisation pré-conceptionnelle du mode de vie dans l'algorithme de prise en charge de l'infertilité dans un contexte de SOPK, avant tout autre traitement en AMP (adapté de Teede *et al.* ¹⁵).

Du point de vue thérapeutique, en particulier dans le contexte de l'infertilité, les nouvelles recommandations internationales mettent l'accent sur l'importance du mode de vie. Une optimisation du mode de vie avant la conception, y compris une alimentation saine et une activité physique régulière, est recommandée, quelle que soit la suite de la prise en charge.

Le SOPK représente ainsi une pathologie complexe et insuffisamment définie. En effet, son diagnostic repose sur l'exclusion d'autres conditions médicales, et les critères diagnostiques

font souvent l'objet de discussions et de débats entre experts. De plus, l'évaluation et la prise en compte du mode de vie jouent un rôle primordial dans la prise en charge initiale du SOPK.

Recouvrement de situations inexpliquées, SOPK et implication du mode de vie

Comme observé et décrit plus haut, et compte tenu des difficultés de définitions et d'identification, il apparaît difficile d'exclure formellement la présence d'un trouble endocrinien (hyperandrogénie ou SOPK) chez une patiente étiquetée inexpliquée.

Par ailleurs, l'analyse de la littérature portant à la fois sur le SOPK et sur l'infertilité inexpliquée révèlent d'importantes similitudes entre ces pathologies :

1. Problèmes diagnostiques communs : errance et retard dans la prise en charge (diagnostic difficile et d'exclusion)
2. Comorbidités partagées entre ces 2 groupes (obésité, syndromes métaboliques, etc.)
3. Absence de consensus sur la stratégie thérapeutique à adopter
4. Importance d'un accompagnement hygiéno-diététique commun

Puisque nous retrouvons les mêmes éléments de langage, les mêmes incertitudes, tant dans l'établissement du diagnostic que dans le choix de la stratégie thérapeutique la plus appropriée, l'une de nos hypothèses est qu'il existerait un continuum de pathologies, entre certaines formes qualifiées d'inexpliquées et des troubles endocriniens et métaboliques de type OPK.

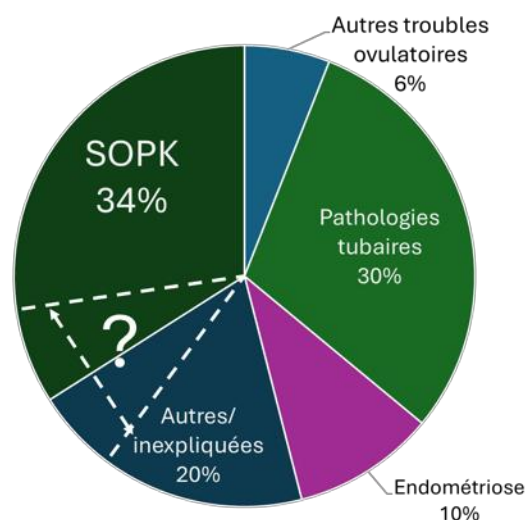


Figure 6 : Hypothèse du continuum et de recouvrement des situations entre le SOPK et l'infertilité inexplicée. Schématiquement, ce groupe représenterait une part importante des cas d'infertilité féminine, en augmentation ces dernières années. La complexité réside dans l'identification précise et la stratification de ces troubles, ainsi que dans les métabolismes impliqués.

Ce recouvrement, constitué de patientes souffrant de troubles endocriniens et métaboliques mixtes difficiles à caractériser, représente un enjeu de santé publique particulièrement important, puisqu'il serait à l'origine de plus de 50 % des cas d'infertilité chez les femmes.

Une meilleure identification et caractérisation de ces troubles, le plus en amont possible, apparaît donc indispensable, dans une perspective de prise en charge personnalisée, préventive prédictive et participative.

Sources de données pour les modèles

Ainsi, à la fois pour l'infertilité inexplicée mais également dans le SOPK, la prise en charge thérapeutique est souvent empirique. Une des raisons qui peut être avancée est la mauvaise ou insuffisante caractérisation diagnostique, en lien avec ces diagnostics d'exclusion, empêchant de pouvoir agir de manière ciblée. C'est toute la justification de ce travail de thèse, qui tente de mieux identifier les contours et mécanismes de ces pathologies dans l'objectif, *in fine*, de mieux les prendre en charge. La première étape consiste à définir des voies métaboliques et biomarqueurs d'intérêts, basées sur des hypothèses physiopathologiques et mécanistiques, et les intégrer dans des modèles.

Dans le contexte de l'infertilité inexplicée, bien que par définition aucune cause n'ait été identifiée, des facteurs liés à l'environnement et au mode de vie pourraient contribuer à la survenue de cette altération de la fonction de reproduction. En lien avec ces hypothèses métaboliques et nutritionnelles, ALIFERT, une étude cas-témoins multicentrique (4 centres) réunissant des couples fertiles et des couples présentant une infertilité primaire inexplicée (400 individus au total), visait à explorer le rôle et l'impact de l'alimentation et du mode de vie sur la fertilité des couples (NCT01093378). Cette exploration tenait compte de nombreux facteurs:

biologiques (non spécifiques - glycémie, lipides, inflammation, électrolytes - et spécifiques - vitamines E, A, C, D, B et folates), cliniques, anthropométriques (IMC, graisse viscérale) ainsi que plusieurs questionnaires (nutrition, consommation de drogues, d'alcool, de tabac, mais également concernant l'activité physique, le sommeil et le stress).

Les premiers résultats issus de cette cohorte ont montré que les hommes et les femmes présentant une infertilité primaire inexplicée présentaient plus fréquemment un comportement sédentaire, une inactivité physique, et des troubles métaboliques (surpoids/obésité et désordres métaboliques (augmentation de la glycémie, dyslipidémie))^{43,44}. Dans ces premières études, les paramètres issus des hommes et des femmes ont été considérés indépendamment. Parallèlement, les variables de l'étude ont été systématiquement appréhendées de manière univariée. Or, les mécanismes sous-jacents sont probablement complexes, multifactoriels, et les facteurs de risques étudiés pourraient avoir des effets cumulatifs ou combinés.

Il a donc semblé pertinent de combiner les différentes variables de l'étude, afin d'explorer plus en profondeur les informations contenues dans ces données. Notre hypothèse était qu'en associant les paramètres des deux membres du couple, et donc en considérant le couple comme une seule et unique observation, nous pourrions identifier les couples infertiles au travers d'une sélection des biomarqueurs les plus pertinents et d'un score. Au total, 80 paramètres (somme des paramètres provenant à la fois de l'homme et de la femme) ont pu être inclus dans l'étude : des paramètres anthropométriques et démographiques (âge, IMC, tour de hanche et taille, graisse viscérale et pressions artérielles) ; des paramètres du métabolisme lipido-glucidique (glycémie, bilan lipidique), mais également ferritine et créatinine) ; des paramètres vitaminique et du statu anti-oxydatif (Vitamines A, C, E, B9 et B12, caroténoïdes, ainsi que des éléments Zinc, Sélénium) ; et un panel de stéroïdes (chez les femmes uniquement).

La seconde partie de ce travail de thèse avait quant à elle pour objectif de simplifier le diagnostic différentiel du SOPK avec le NC21OHD. La 21-hydroxylase étant une enzyme impliquée dans la voie de biosynthèse des stéroïdes, nous nous sommes naturellement intéressés à cette même voie et avons choisi de mesurer un panel de stéroïdes pour identifier cette pathologie due à un bloc enzymatique⁴⁵. Les données d'entrée de nos modèles ont donc consisté en une cartographie de la voie de biosynthèse des stéroïdes, correspondant aux concentrations de 15 espèces moléculaires (figure 7) : 21-désoxycortisol (21-DF), 11b-hydroxyandrostènedione (11b-OHA), androstènedione (4A), testostérone (T), 17-OH-progesterone (17OHP), 17-OH-prégnénolone (17-OHPreg), déhydroépiandrostérone (DHEA),

(E), cortisol (F), corticostérone (B), 11-désoxcorticostérone (DOC), 21-désoxcorticostérone (21-DB), progesterone (P) et prégnénolone (Prog)⁴⁶.

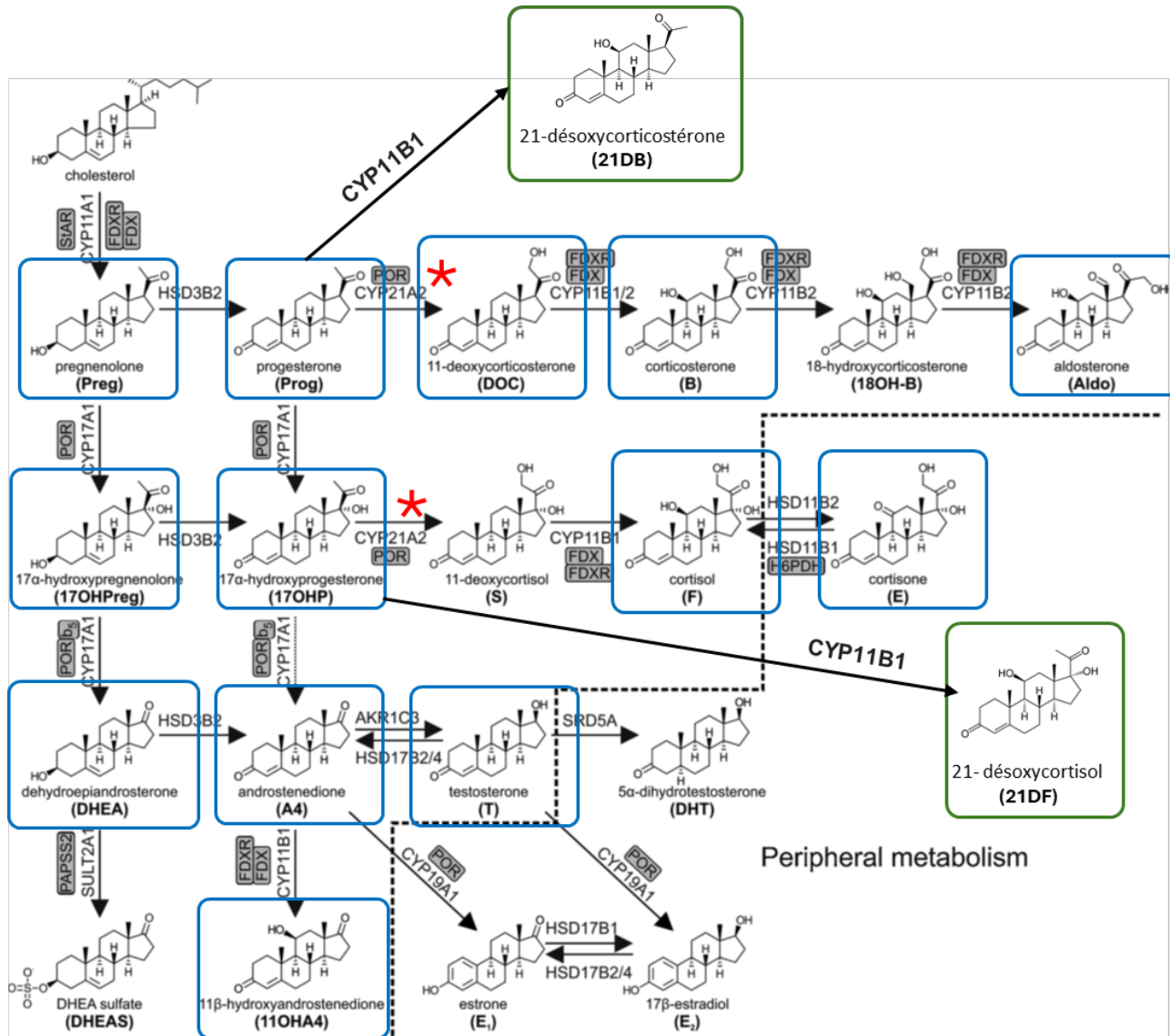


Figure 7 : Vue d'ensemble de la stéroïdogénèse surrénalienne et la modulation périphérique. Les stéroïdes inclus dans le modèle sont entourés en bleu. La 21-hydroxylase est identifiée avec une étoile rouge. En cas d'accumulation, la 11βhydroxylase (CYP11B1) peut exercer directement son activité sur 17OHP et P, donnant lieu à la production de 21DF et 21DB, respectivement (entourés en vert). Les protéines accessoires impliquées : cytochrome b5 (b5) ; cytochrome P450 oxydoréductase (POR) ; ferrédoxine (FDX) ; ferrédoxine réductase (FDXR) ; hexose-6-phosphate déshydrogénase (H6PDH) ; PAPS synthase 2 (PAPSS2) ; steroidogenic acute regulatory protein(StAR) (d'après Schiffer *et al.*⁴⁵).

Toute la démarche de cette thèse est fondée sur une approche globale et intégrative, afin, *in fine*, de mieux appréhender de manière rationnelle les relations entre les différentes variables. Cette approche globale et intégrative n'est possible que par le biais d'une exploration exhaustive des voies d'intérêt et d'un traitement mathématique et informatique puissant. La description plus technique de ce couplage des sciences omiques et des méthodes issues de l'intelligence artificielle fera l'objet du chapitre suivant.

Méthodes appliquées : outils analytiques et mathématiques

L'objectif de cette partie est de présenter succinctement la méthodologie et les différents outils utilisés dans ce travail. On distingue 2 types d'outils : analytiques et post-analytiques.

En accord avec notre approche, l'exploration du métabolisme -au sens large- dans ces conditions complexes et inexplicables nécessite l'utilisation conjointe de deux outils. Le premier est d'ordre analytique, visant à produire des profils et des empreintes métaboliques, qui constituent les données d'entrée des modèles. Le second consiste en la modélisation ou l'analyse à l'aide de méthodes d'intelligence artificielle pour interpréter les profils ainsi générés. Ces méthodes permettent également d'éventuellement les intégrer à d'autres données cliniques, biologiques et diverses afin de construire des scores diagnostiques.

Données omiques et métabolomiques

Les sciences dites omiques (ou -omiques) sont constituées de diverses disciplines de la biologie, telles que la génomique (étude du génome), la protéomique (étude à grande échelle des protéines), la transcriptomique (étude des transcriptomes, *i.e.* de toutes les molécules d'ARN), et la métabolomique, qui nous intéressera plus particulièrement dans cette thèse. Les "omiques" s'intéressent à la caractérisation et à la quantification systématiques de pools de molécules biologiques qui reflètent la structure, la fonction et la dynamique d'un organisme, et qui sont souvent utilisées pour décrire ou comprendre un phénotype.

La métabolomique offre une vue d'ensemble biochimique d'un système biologique et de l'impact physiologique d'une maladie, d'une alimentation ou d'un traitement, par exemple. La métabolomique est une approche puissante car les métabolites et leurs concentrations, contrairement à d'autres mesures "omiques", reflètent directement l'activité biochimique sous-jacente et l'état des cellules et tissus. Ainsi, en biologie systémique, la métabolomique

représente au mieux le phénotype moléculaire et l'état physiologique actuel d'un organisme ⁴⁷⁻
50.

Une analyse métabolomique implique une approche intégrée de la préparation des échantillons, de la standardisation, de l'étalonnage, des méthodes de séparation, de la détection des métabolites et de l'analyse des données. Les types d'échantillons couramment explorés sont le plasma, l'urine, la salive, les tissus ou les cultures cellulaires. Les interactions des métabolites au sein d'un système biologique sont appelées "métabolome". Le métabolome est l'ensemble des métabolites présents dans un organisme ou un échantillon biologique. Les métabolites sont des composés de faible poids moléculaire, généralement inférieur à 1,5 kDa, qui sont les intermédiaires ou les produits des voies de biosynthèse/catabolisme. Il s'agit, par exemple, des acides aminés, des nucléotides, des glucides et des lipides, qui sont souvent étudiés séparément dans le cadre de la recherche en lipidomique (incluant les stéroïdes).

La métabolomique vise à analyser de manière plus ou moins globale et/ou quantitative la composition chimique des métabolites présents dans un échantillon biologique, souvent par le biais de techniques analytiques avancées telles que la chromatographie liquide couplée à la spectrométrie de masse (LC-MS/MS) ou la spectroscopie de résonance magnétique nucléaire. Deux stratégies de profilage métabolomique courantes sont l'analyse métabolomique ciblée et non ciblée. L'analyse ciblée quantifie de manière absolue certains métabolites connus (en termes de structure), tandis que l'analyse non ciblée fournit un profil métabolique global des métabolites connus et inconnus. L'approche non ciblée est plus adaptée à la découverte de nouveaux biomarqueurs tandis que l'approche ciblée, quantitative est privilégiée en biologie médicale pour le diagnostic. C'est donc naturellement que nos analyses en masse se concentrent sur des voies métaboliques connues pour une utilisation en routine clinique.

Dans cette thèse, les voies métaboliques étudiées et incluses dans les modèles (telles que la voie de biosynthèse des stéroïdes, les panels d'antioxydants et de vitamines, etc.) ont été analysées par spectrométrie de masse. Les principes de fonctionnement de base de cette méthode seront exposés dans le chapitre suivant.

Analytique et spectrométrie de masse

La spectrométrie de masse a permis un apport technologique et analytique important, et est maintenant disponible dans de nombreux laboratoires de biologie médicale en remplacement

des approches historiques d'immunoanalyse (moins sensibles et surtout moins spécifiques). Contrairement aux immunodosages, la MS permet de proposer une quantification absolue et surtout simultanée de profils moléculaires de plusieurs stéroïdes. D'un point de vue diagnostic, ces panels (ou signatures biochimiques) ont profondément modifié la sémiologie des pathologies endocriniennes. Grâce à un accès direct à une cartographie métabolique, il est maintenant possible d'étudier un grand nombre de stéroïdes d'origine surrénalienne, gonadique ou mixte. Son intérêt est renforcé par le fait que l'analyse est très sensible et peut s'effectuer, par exemple, sur des prélèvements de sérum de quelques dizaines de microlitres et même sur des tâches de sang (« *dry blood spot* »).

Principe général de fonctionnement :

La LC-MS/MS est une technique analytique puissante combinant deux méthodes de séparation et de détection pour fournir des informations précises sur la composition des échantillons. Le principe de fonctionnement de la LC-MS/MS repose sur deux étapes principales : la chromatographie liquide haute performance (HPLC) et la spectrométrie de masse en tandem (MS/MS). Dans la première étape, la HPLC est utilisée pour séparer les composants d'un échantillon en fonction de leurs propriétés physico-chimiques telles que la taille, la polarité et l'interaction avec la phase stationnaire. Cette séparation permet d'isoler les analytes d'intérêt des autres composants de l'échantillon, améliorant ainsi la sensibilité et la sélectivité de l'analyse. Une fois les composants séparés par la HPLC, ils sont introduits dans le spectromètre de masse en tandem pour l'analyse.

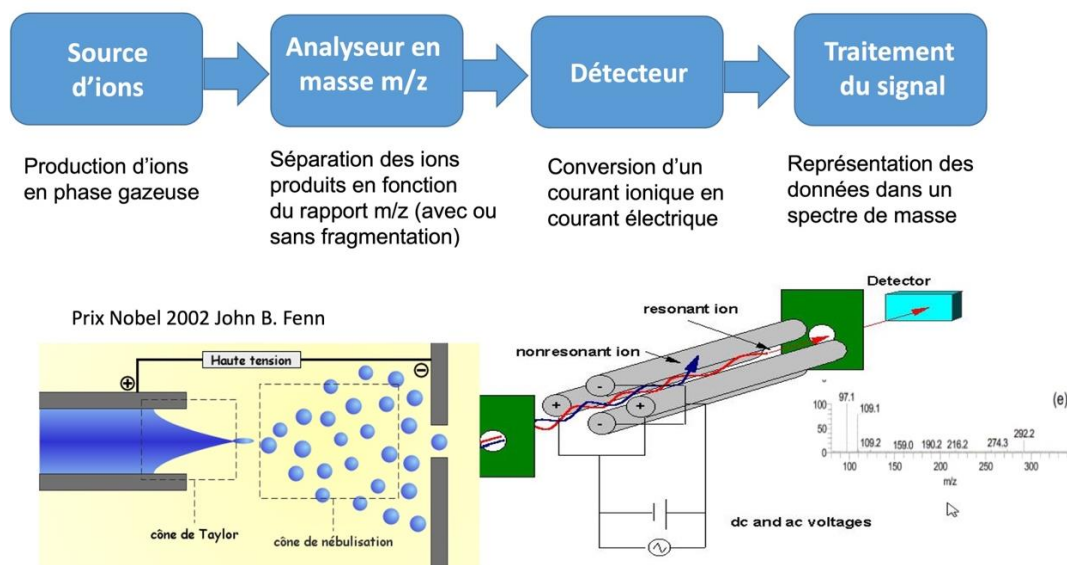


Figure 8 : Principe général du fonctionnement d'un spectromètre de masse avec une source d'ionisation par electrospray (adapté du cours de spectrométrie de masse du Pr. Olivier Laprévotte).

La première étape consiste à l'ionisation des métabolites d'intérêt. L'ionisation se fait à pression atmosphérique, soit par électro-nébulisation (ESI « *electrospray ionisation* »), soit par ionisation chimique à pression atmosphérique (APCI) (figure 8) ⁵¹. Ces ions sont ensuite accélérés dans le spectromètre de masse où ils sont séparés en fonction de leur rapport masse/charge (m/z) grâce à des quadripôles. Dans la configuration d'un spectromètre en tandem MS/MS, les ions sélectionnés sont soumis à une fragmentation supplémentaire dans une cellule de collision, ce qui permet d'obtenir des informations structurales spécifiques sur les composants de l'échantillon (figure 9). En pratique, les ions formés sont séparés dans un premier quadripôle Q1, qui conduit les ions parents jusqu'à la chambre de collision dans laquelle a lieu une fragmentation en ions fils, qui sont triés par le deuxième quadripôle Q3 en fonction de leur m/z et véhiculés jusqu'au système de détection ^{52,53}.



Figure 9 : Structure du système de chromatographie liquide couplé à la spectrométrie de masse en tandem (LC-MS/MS). Cette technique combine la séparation par chromatographie liquide (LC) et l'analyse du rapport m/z de la spectrométrie de masse en tandem (MS/MS) afin de générer des spectres. (<https://sciex.com/products/mass-spectrometers/triple-quad-systems/triple-quad-7500-system>)

Les données générées par l'HPLC-MS/MS se présentent sous forme de profils omiques complexes, caractérisés par une multitude de pics chromatographiques (figure 10). Ces pics représentent les différents métabolites présents dans l'échantillon analysé. Cependant, en raison de la nature volumineuse et de la complexité de ces données, leur traitement nécessite l'utilisation de méthodes mathématiques avancées. Ces méthodes sophistiquées sont essentielles pour extraire efficacement les informations pertinentes, réduire le bruit de fond et détecter les signaux significatifs au milieu du bruit, permettant ainsi une interprétation précise et exhaustive des profils omiques obtenus.

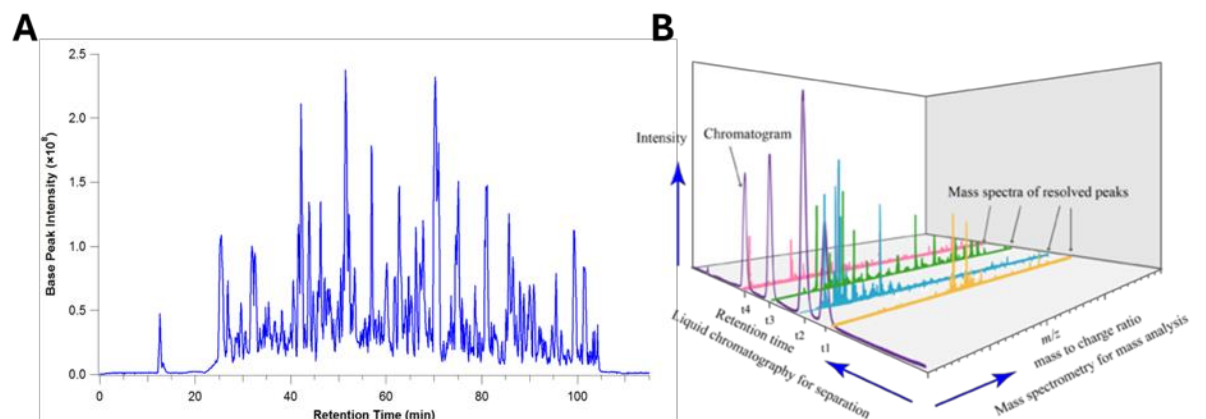


Figure 10 : Spectres générés par la LC-MS/MS. Dans la chromatographie en phase liquide, les métabolites sont envoyés dans une colonne et sont séparés en fonction de diverses propriétés chimiques. À chaque instant, l'abondance d'un groupe particulier de métabolites est mesurée (l'intensité) = c'est ce qu'on appelle un chromatogramme (A). À chaque temps du chromatogramme, le groupe de métabolites est alors ionisé (chargé) et transmis au spectromètre de masse. Cela ajoute une nouvelle dimension au graphique : un axe avec le rapport masse/charge des métabolites détectés à ce point dans le temps, et leurs intensités (B).

De plus, une des conséquences des approches multiplexes par MS est le profond changement de paradigme qu'elle induit ; le raisonnement ne se développe plus en considérant

les analytes indépendamment les uns des autres (comme l'imposent les approches par immunoanalyse) mais en considérant des empreintes métaboliques. Cette nouvelle sémiologie a nécessité de développer des modèles mathématiques multivariés permettant d'appréhender les systèmes biologiques complexes en condensant l'information au travers de projections et/ou de combinaisons de variables moléculaires. Cette approche peut également être élargie avec l'intégration d'autres données hétérogènes (cliniques et biologiques).

Outils mathématiques

Après avoir réalisé la première étape de chimie analytique pure mentionnée précédemment, consistant en l'acquisition des empreintes métaboliques, et après avoir procédé à un traitement du signal visant à extraire les informations analytiquement pertinentes de ces empreintes générées, intervient la phase d'analyse statistique et l'utilisation d'outils mathématiques. En effet, les données produites par les approches omiques sont complexes, et divers outils ont été développés dans le but de résumer, visualiser et comparer les informations biologiquement pertinentes.

Généralités sur l'intelligence artificielle et l'apprentissage automatique

L'intelligence artificielle (IA) représente un ensemble de théories et de techniques visant à développer des machines capables de simuler l'intelligence humaine. Bien que le concept d'IA existe depuis plus d'un demi-siècle, son essor véritable a eu lieu depuis les années 2010. Deux facteurs peuvent expliquer cette croissance : d'une part, une augmentation exponentielle du volume de données collectées et disponibles, et d'autre part, l'accroissement de la puissance de calcul des ordinateurs, facilitant ainsi l'exploitation pratique et technique de ces données. Dans la plupart des cas, notamment dans les applications d'IA en santé, il est sous-entendu que l'on fait référence à l'apprentissage automatique. L'apprentissage automatique (ou machine learning (ML) en anglais) est un domaine de recherche en IA qui cherche à doter les machines de la capacité d'apprendre à partir de données, en utilisant des modèles mathématiques. Cette approche diffère de la programmation traditionnelle, où un programme est créé manuellement par un programmeur, utilisant des données en entrée pour produire une sortie (Figure 11).

Comment fonctionne-t-il concrètement ? Fondamentalement, il s'agit d'acquérir et d'estimer des modèles généraux à partir d'exemples disponibles. Dans le cadre de l'apprentissage supervisé, ces exemples se composent de patients définis par un ensemble de

caractéristiques X , telles que des informations cliniques, biologiques et/ou d'imagerie, ainsi que par un diagnostic ou un état de santé Y . Les données sont dites étiquetées, ce qui signifie que chaque exemple (patient) a été annoté par un expert, tel qu'un médecin ayant établi un diagnostic. Dans le domaine de l'apprentissage supervisé, on distingue les problèmes de classification et de régression. La classification est utilisée lorsque Y est qualitatif (par exemple, malade/sain), tandis que la régression est appliquée lorsque Y est quantitatif (par exemple, la durée de survie).

La mise en place du modèle, également connue sous le nom d'entraînement du modèle, requiert un ensemble significatif de couples de données « caractéristiques X » → « diagnostic Y » afin d'estimer de manière automatique la fonction $f(X) = Y$, permettant ainsi de relier les caractéristiques du patient (X) à son diagnostic (Y) (voir Figure 12). La cohorte d'entraînement désigne généralement l'ensemble des patients utilisés pour construire le modèle. Par la suite, la fonction f , soit le modèle, sera en mesure de fournir un diagnostic Y , ou du moins de l'estimer avec une certaine précision, en utilisant les mêmes caractéristiques X utilisées pendant l'entraînement. Afin de garantir sa capacité à effectuer de nouveaux diagnostics de manière fiable, il est essentiel de valider le modèle sur un ensemble distinct de patients (appelé cohorte de validation ou de test). Il suffit alors de comparer le diagnostic réel des patients de cette nouvelle cohorte à celui prédit par le modèle de manière aveugle. Cette procédure de validation, pour évaluer la performance d'un modèle, conduit à la création d'une matrice de confusion pour les modèles de classification (permettant de calculer la sensibilité, la spécificité, etc.) ou à un coefficient de détermination dans le cas d'une régression.

Divers types de modèles sont disponibles et peuvent être utilisés comme fonction f . De la régression logistique simple aux réseaux de neurones complexes, chacun de ces modèles aborde mathématiquement le problème à résoudre de manière spécifique, avec ses propres avantages et limites. Il est cependant crucial de noter qu'un modèle complexe n'est pas toujours le plus adapté à la tâche à accomplir. En effet, outre la performance, un aspect essentiel à considérer est l'intelligibilité ou l'explicabilité du modèle. L'intelligibilité d'un modèle se réfère à sa capacité à expliquer ou à présenter son fonctionnement de manière compréhensible pour un être humain, et constitue un domaine de recherche actif. Dans le domaine de la santé en particulier, cette intelligibilité permet notamment de comprendre les prédictions formulées par le modèle et d'accroître la transparence. Elle facilite également l'extraction d'informations à

partir du modèle élaboré, ainsi que l'identification de biomarqueurs pertinents parmi un large éventail de caractéristiques X.

Plusieurs éléments cruciaux entrent en jeu lors de la construction d'un modèle de ML :

1. L'utilisation d'un échantillon représentatif de la population cible en tant que cohorte d'entraînement. Il est important de se poser des questions sur les patients sur lesquels le modèle développé sera appliqué par la suite : avons-nous choisi les bons exemples pour entraîner notre modèle ?
2. La pertinence du choix des variables X pour caractériser chaque patient.
3. La qualité de l'annotation (étiquette Y).

En conséquence, l'intelligence artificielle, en particulier le machine learning, représente des outils puissants, mais une collaboration étroite entre les experts du domaine d'application (tels que les médecins ou les biologistes dans le cas des applications en santé) et les experts en analyse de données et en informatique est indispensable. En effet, seul l'expert du domaine sera capable de sélectionner des caractéristiques ou variables X pertinentes en s'appuyant sur des hypothèses physiopathologiques connues, et également d'attribuer correctement les étiquettes aux exemples.

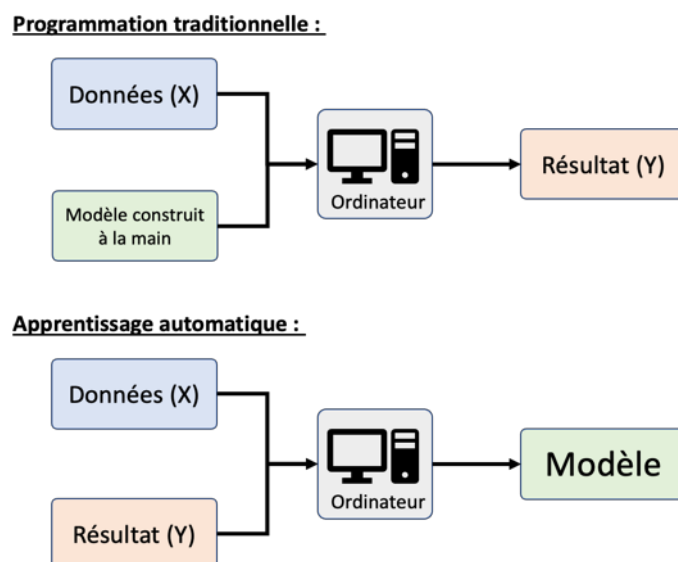


Figure 11 : Deux approches différentes : programmation traditionnelle et apprentissage automatique. La programmation traditionnelle désigne tout programme créé manuellement qui utilise des données d'entrée (X) et s'exécute sur un ordinateur pour produire la sortie (Y). Dans

l'apprentissage automatique, les données d'entrée (X) et les données de sortie (Y) sont fournies dans la machine et celle-ci élabore ses règles pour construire de manière automatique un modèle, et donc sans être programmée explicitement (d'après Bachelot *et al.* ⁵⁴).

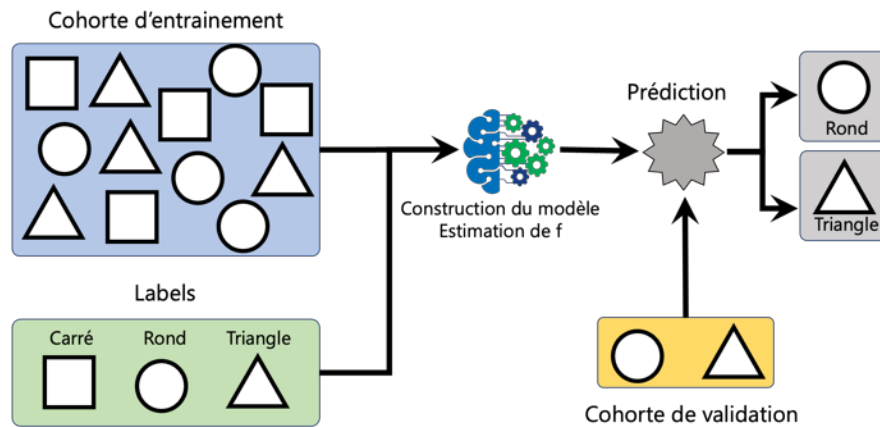


Figure 12 : Fonctionnement d'un modèle de classification. Étape 1 : le modèle est construit et entraîné avec un lot d'individus labélisé (dont on connaît le label ou l'identité Y : carré, rond ou triangle) ; la fonction f est donc estimée et va servir à l'étape suivante ; Étape 2 : prédiction du label Y d'un nouveau lot d'individus, d'identité inconnue, par le modèle précédemment entraîné (d'après Bachelot *et al.* ⁵⁴).

Approche par régression des moindres carrés partiels ou PLS (« *Partial Least Squares* » et/ou « *Projection to Latent Structure* »)

La méthode PLS, développée par Herman Wold (1966) et Michel Tenenhaus, est une technique d'analyse multivariée qui vise à modéliser les relations linéaires entre un ensemble de variables explicatives (X) et une variable cible (Y) ^{55,56}. Historiquement, la méthode PLS a été très utilisée en chimiométrie, un domaine de la chimie qui se concentre sur l'analyse de données chimiques complexes, pour résoudre des problèmes tels que la calibration de spectres, la prédiction de propriétés chimiques, la classification d'échantillons, ou encore la déconvolution de signaux complexes en spectroscopie. L'approche PLS repose sur le concept de décomposition des matrices X et Y en une série de composantes principales, également appelées facteurs latents, qui capturent les structures sous-jacentes et maximisent la covariance entre X et Y. Cette décomposition est réalisée de manière itérative, en ajustant successivement les composantes principales pour expliquer la variation résiduelle non expliquée.

En effet, un algorithme clé largement utilisé est l'algorithme NIPALS (« *Non-linear Iterative Partial Least Squares* »), développé par Herman Wold et Svante Wold⁵⁷. L'algorithme NIPALS est utilisé pour calculer les composantes principales dans la décomposition des matrices X et Y. L'algorithme NIPALS est itératif et repose sur une approche de recherche d'une composante à la fois. Il commence par une estimation initiale des composantes principales, puis ajuste successivement ces estimations pour maximiser la covariance entre les variables explicatives (X) et la variable cible (Y). Cette optimisation itérative est réalisée de manière efficace et converge généralement rapidement vers des solutions précises.

L'approche PLS combine donc des aspects de l'analyse en composantes principales (ACP) et de la régression multiple pour résoudre des problèmes de modélisation et de prédiction dans des ensembles de données multivariées. D'une part, l'approche PLS partage des similitudes avec l'ACP dans le sens où elle cherche à identifier des combinaisons linéaires des variables explicatives (X) qui capturent le maximum de variance dans les données. Ces combinaisons linéaires, appelées composantes principales, sont déterminées de manière à expliquer à la fois la variance des variables explicatives et leur relation avec la variable cible (Y). D'autre part, l'approche PLS est également prédictive, car elle vise à établir une relation linéaire entre les variables explicatives (X) et la variable cible (Y) afin de réaliser des prédictions précises sur de nouvelles observations. Cette relation est établie en utilisant les composantes principales identifiées pour maximiser la covariance entre X et Y, tout en réduisant le risque de surajustement par rapport à la régression multiple traditionnelle. Ainsi, l'approche PLS combine les avantages de l'ACP pour la réduction de la dimensionnalité des données et de la régression multiple pour la modélisation et la prédiction précises des relations entre les variables explicatives et la variable cible. Cette combinaison en fait une méthode puissante et polyvalente pour traiter efficacement des ensembles de données multivariées complexes dans divers domaines d'application.

En termes pratiques, l'approche PLS offre plusieurs avantages qui se révèlent très utiles dans nos problèmes de modélisation de données médicales et omiques : la capacité à traiter la multicolinéarité, à gérer les ensembles de données à haute dimensionnalité et à traiter les valeurs manquantes. En effet, dans l'exploration des voies métaboliques, en particulier, les espèces moléculaires impliquées sont souvent intégrées dans des voies et interconnectées par des systèmes enzymatiques, ce qui entraîne une forte multicollinéarité. L'algorithme PLS offre l'avantage de gérer efficacement cette multicollinéarité en utilisant des composantes

orthogonales issues des variables explicatives. De plus, lors de la génération de profils omiques, il est courant d'obtenir un grand nombre de molécules, alors que dans le cas de pathologies rares, par exemple, le nombre de patients est limité. Cela entraîne la création d'un tableau de données plus large que haut, ou "écrasé", que l'algorithme PLS peut efficacement traiter grâce à la réduction de dimension induite par l'utilisation de composantes. Enfin, la méthode PLS repose sur un algorithme itératif capable de traiter des données incomplètes. Les données disponibles sont utilisées pour construire les composantes, ce qui évite le besoin de compléter les données manquantes. En revanche, la principale limitation de l'approche PLS réside dans sa linéarité : les composantes sont des combinaisons linéaires des variables, et les relations entre les composantes de X et de Y sont également linéaires. Dans certains cas, il peut être opportun d'explorer des approches non linéaires, telles que les arbres de décision et leurs dérivés, ainsi que les réseaux neuronaux.

L'analyse discriminante par moindres carrés partiels (PLS-DA) est une méthode qui utilise la régression par moindres carrés partiels pour prédire un ensemble de variables binaires Y représentant les catégories d'une variable catégorielle à partir d'un ensemble de variables prédictives X. Elle se situe à mi-chemin entre l'analyse discriminante classique et l'analyse discriminante sur les composantes principales significatives des variables prédictives ⁵⁸.

Pour finir, la méthode PLS possède une caractéristique importante qui la rend particulièrement pertinente pour la recherche en santé et métabolisme : sa capacité à être interprétable. En effet, le VIP (« *Variable Importance in Projection* ») est une mesure utilisée pour évaluer l'importance des variables explicatives dans la prédiction de la variable cible. C'est une mesure de l'impact de chaque variable sur les composantes principales. Comme nous le verrons dans les résultats de cette thèse, nous serons en mesure d'identifier les biomarqueurs potentiels liés à une condition ou à un processus métabolique, et d'analyser leur interaction ainsi que leur contribution aux résultats observés. Cela nous permettra d'interpréter plus en détail les modèles PLS et de mieux comprendre les relations complexes entre les variables explicatives et la variable cible.

Objectifs de la thèse

Ce projet de thèse, à l'interface entre la biologie de la reproduction, l'exploration du métabolisme et le traitement informatique de données, a pour objectif de développer un ou plusieurs outils diagnostics innovants dans les domaines de la fertilité et du métabolisme. Il s'agira de développer des outils pour identifier et suivre des pathologies dans des contextes d'infertilité où le mode de vie moderne et l'environnement ont probablement un impact important dans leur développement et leur pronostic. Ces éléments pourront également être exploités et devenir des leviers d'action sur lesquels jouer ou à actionner pour la prise en charge thérapeutique.

Au-delà de l'analyse de données, et en s'appuyant sur l'expertise et les outils disponibles au sein du département de métabolomique clinique (Pr Lamazière Antonin), il s'agira de développer des approches omiques ciblées (par spectrométrie de masse) pour générer des profils métaboliques quantitatifs. Ainsi les voies biochimiques d'intérêts (voie de biosynthèse des stéroïdes, signature microbiote, statut oxydatif) pourront être cartographiées en vue de l'identification et la sélection de biomarqueurs.

Deux projets parallèles ont été menés. Le premier, qui s'appuie sur les données du PHRC ALIFERT (Pr Lévy Rachel et Dr Dupont Charlotte), porte sur l'infertilité inexplicée. Il a conduit au développement d'un score de couple permettant d'identifier et de stratifier les couples confrontés à cette condition. Le second projet, axé sur les aspects endocriniens, s'est penché sur les hyperandrogénies chez les femmes, en se concentrant notamment sur le diagnostic différentiel entre le SOPK et l'hyperplasie congénitale des surrénales due à un blocage de la 21-hydroxylase (Pr Bachelot Anne/Pr. Jacques Young).

Résultats I : Développement d'un score de machine learning basé sur une approche couple pour identifier et stratifier les couples infertiles inexplicés

Contexte scientifique et objectifs

Comme décrit dans l'introduction du manuscrit, dans le contexte de l'infertilité inexplicée, bien que par définition aucune cause n'ait été identifiée, des facteurs liés à l'environnement et au mode de vie pourraient contribuer à la survenue de l'altération des fonctions de reproduction. Les premiers résultats de la cohorte ALIFERT ont montré que les hommes et les femmes atteints d'infertilité primaire inexplicée étaient plus susceptibles de présenter un comportement sédentaire, une inactivité physique et des troubles métaboliques (surpoids/obésité et désordres métaboliques (augmentation de la glycémie, dyslipidémie))^{43,44}. Cependant, les paramètres étaient jusqu'à présent toujours considérés individuellement, empêchant la prise en compte de leur interconnexion et limitant la compréhension du ou des phénomènes en jeu. Il est donc apparu intéressant de combiner les différentes variables de l'étude, afin d'explorer plus en profondeur les informations contenues dans ces données. Notre hypothèse était qu'en combinant les paramètres des deux membres du couple, et donc en considérant le couple comme une seule observation, nous pourrions identifier les couples infertiles grâce à une sélection des biomarqueurs les plus pertinents.

L'objectif de ce premier travail était donc 1) de construire et d'évaluer un modèle d'apprentissage automatique capable d'identifier et de stratifier les couples présentant une infertilité inexplicée, et 2) d'identifier les biomarqueurs les plus pertinents dans le modèle à partir du large panel initial de la cohorte ALIFERT (n=80 variables). Pour ce faire, nous avons exploité les données cliniques et biologiques recueillies au cours de l'étude comme données d'entrée pour les modèles d'apprentissage automatique. La nature multicentrique de l'étude nous a également permis de valider la performance des modèles de manière externe et indépendante.

Résumé des résultats

Les résultats ont ainsi été présentés en deux étapes : 1) évaluation de la pertinence de l'approche par couple, de l'identification des biomarqueurs importants et construction d'un

modèle réduit (**article 1**), et 2) mise en perspective du score avec les paramètres spermatiques des hommes infertiles de la cohorte (**article 2**).

Article 1 :

Supériorité du modèle du couple par rapport aux modèles hommes et femmes indépendants

Dans le cadre d'un travail d'apprentissage automatique ou d'analyse multivariée, la première étape consiste en une analyse exploratoire des données. Nous avons donc utilisé ACP comme méthode de réduction de dimensions et d'exploration multidimensionnelle, à la fois pour déceler les premières tendances et surtout pour repérer les valeurs aberrantes ou extrêmes. Les ACP avec les scores pour les hommes, les femmes et les couples (A, B et C respectivement) projetés sur le premier plan factoriel n'ont pas montré de séparation entre les groupes fertiles et infertiles. De plus, les graphiques des scores pour les hommes et les couples révélaient une valeur aberrante : un homme avec un IMC de 45, soit plus de 8 écarts-types au-dessus de la moyenne. Ce patient et donc ce couple ont été exclus des étapes ultérieures, car ils n'étaient pas représentatifs de la population cible pour la validation et l'utilisation du modèle. Le maintien de ce patient dans le modèle aurait artificiellement donné de la puissance à la variable "IMC" et faussé le modèle. Pour tester l'avantage d'une approche basée sur le couple, des modèles "hommes", "femmes" et "couples" ont été construits séparément et leur performance ont été comparés. Seuls les résultats de l'algorithme OPLS-DA seront présentés dans l'article, les 4 autres modèles d'apprentissage automatique (Support Vector Machine, Nearest Neighbours Classifier, Random Forest et Logistic Regression) construits et évalués n'ayant pas montré de meilleures performances.

Après comparaison des résultats internes et externes des modèles "hommes", "femmes" et "couples", nous avons observé que l'approche du couple a abouti à un gain significatif en puissance discriminatoire. Le modèle "femme" était par ailleurs plus informatif que le modèle "homme", et proche des performances du modèle "couple".

Sélection des biomarqueurs pertinents et construction d'un score couple réduit de 13 variables et performances

L'objectif de cette deuxième étape était de réduire la dimensionnalité du système en ne conservant que les variables les plus pertinentes parmi les 80 variables initiales. Le score

“couple” s’est montré plus efficace que les scores des membres du couple indépendants. En revanche, la collecte et l’intégration des 80 paramètres nécessaires pour obtenir un résultat semblaient, dans la pratique, difficilement envisageables.

Afin de réduire le nombre de variables, nous avons procédé à une sélection itérative de ces dernières en fonction du VIP : plus le VIP est élevé, plus la variable est discriminante entre les groupes fertiles/infertiles. La réduction des variables les moins significatives a également abouti à la réduction de la colinéarité entre les variables redondantes, à la réduction du "bruit" généré par la variance des variables les moins discriminantes et à l’obtention d’un modèle plus ergonomique. En pratique, nous avons successivement sélectionné les variables avec un $VIP > 1$ pour obtenir un premier modèle réductionniste à 24 puis un second à 13 variables, correspondant au meilleur compromis performance / nombre de variables. La performance (exactitude) du modèle à 13 variables sur la cohorte de test externe était de 73,8 %. Ce résultat est supérieur à celui observé avec le modèle utilisant les 80 variables (68,8 %).

Dans le modèle à 13 variables, on retrouvait une contribution importante de paramètres anthropométriques, métaboliques et antioxydants (pouvant être de ce fait divisés en trois groupes de variables). Dans ce modèle, la glycémie s’est avérée être la variable la plus importante. Les variations des paramètres anthropométriques tels que la graisse viscérale, l’IMC et le tour de taille chez les femmes et les hommes étaient également parmi les paramètres les plus significatifs dans le modèle. Le rétinol circulant a également joué un rôle significatif et, dans une moindre mesure, les autres espèces antioxydantes telles que la lutéine, le bêta-carotène et l’alpha-carotène ont montré un score VIP élevé. En outre, la proportion de variables provenant de l’homme et de la femme était équivalente, ce qui renforçait la pertinence de l’approche couple puisque le modèle extrayait spontanément des informations issues des deux membres du couple.

Ce premier travail sur la cohorte ALIFERT nous a permis de montrer qu’il était possible, en intégrant des données hétérogènes (anthropométrique, métaboliques, nutritionnelles) à l’aide de méthodes de modélisation, d’identifier les couples présentant une infertilité inexplicée des couples contrôles, sur la base de facteurs métaboliques et nutritionnels.

Article 2 :

Association entre le statut anthropométrique, métabolique et nutritionnel et les paramètres du sperme

Au cours de l'étude clinique ALIFERT, les paramètres spermatiques des hommes infertiles inclus dans la cohorte ont été également mesurés. Nous avons évalué la relation entre les paramètres spermatiques conventionnels et fragmentation de l'ADN spermatique d'une part (les paramètres spermatiques conventionnels étaient normaux ou subnormaux pour tous les hommes de la cohorte), et le résultat du score calculé pour chaque homme de la cohorte d'autre part.

Nous avons pu observer une corrélation significative entre la valeur de score calculé et le taux de fragmentation de l'ADN spermatique ($r = -0.263$, $p < 0.001$), marqueur qualitatif de la spermatogenèse. Dans notre cohorte, le taux moyen de fragmentation de l'ADN, mesuré avec la méthode TUNEL (« *Terminal dUTP nick-end labeling* »), était de 28.6% +/- 18.4.

Ces résultats doivent être confirmés sur un échantillon plus large, incluant des patients présentant des paramètres spermatiques anormaux ainsi que des témoins, mais suggèrent qu'il pourrait exister un lien complexe entre le mode de vie, l'alimentation et la fragmentation de l'ADN spermatique. Un mode de vie et une alimentation déséquilibrée ont des répercussions sur les aspects anthropométriques, métaboliques et nutritionnels, pouvant entraîner un déséquilibre du stress oxydatif. Ces conséquences sont multiples, complexes et intriquées, et l'utilisation de l'intelligence artificielle offre une opportunité unique pour fournir une évaluation plus holistique de l'influence des marqueurs considérés.

Éléments de discussion

Sur le modèle

Dans la première partie de ce travail de thèse, nous avons donc développé et évalué un outil diagnostique innovant, permettant d'identifier et de stratifier les couples atteints d'infertilité inexplicée. L'approche couple, consistant à combiner des paramètres issus à la fois des deux membres du couple est inédite dans la littérature. Cette approche suppose que, dans une perspective de complémentarité voire de compatibilité au sein du couple, le "degré" de risque d'infertilité du premier membre du couple pourrait être compensé ou rattrapé par le "degré" de

risque d'infertilité du second membre du couple. Nous avons démontré que l'approche de modélisation du couple était effectivement plus discriminante qu'un modèle dans lequel les hommes et les femmes étaient considérés séparément et que les variables les plus importantes pour la prédiction étaient liées aux troubles métaboliques, au statut oxydatif et à l'anthropométrie (et ce malgré l'absence de diabète ou de syndrome métabolique chez les patients inclus). Le modèle réduit, construit à partir de 13 variables, a démontré sa capacité prédictive au sein d'une cohorte externe avec un léger surajustement ou surapprentissage (83,8% sur la cohorte d'entraînement vs. 73,8% sur la cohorte de test). Le modèle final, entraîné sur l'ensemble des données (196 couples), nécessitera par la suite une validation supplémentaire sur une cohorte de données indépendante.

Sur les variables / VIP

Une de nos hypothèses de départ était que chez ces couples infertiles, dont aucune cause n'a été identifiée, le mode de vie et la nutrition ont un impact significatif qui participe et contribue à leur infertilité via un ou plusieurs mécanismes complexes, faisant notamment intervenir le métabolisme. Les résultats obtenus ci-dessus, présentant le classement des variables en fonction de leur importance dans le modèle, corroborent nos hypothèses et suggèrent que ces conséquences métaboliques, plus difficiles à identifier individuellement, seraient probablement impliquées dans l'infertilité. Ces facteurs du mode de vie sont souvent étudiés séparément, mais leurs effets peuvent être cumulatifs lorsqu'ils sont associés. De plus, plusieurs travaux suggèrent que les effets de ces facteurs seraient réversibles si des mesures appropriées sont prises : ce dernier point reste encore à démontrer avec des études interventionnelles prospectives. D'un point de vue mécanistique et physiopathologique, ces variables identifiées sont raisonnables : une glycémie élevée, un faible taux de cholestérol HDL (High Density Lipoprotein Cholestérol) et des paramètres anthropométriques associés à l'obésité abdominale sont évocateurs du syndrome métabolique, qui a un impact avéré sur les fonctions reproductives des hommes et des femmes. Il est légitime de supposer que, malgré l'absence systématique de syndrome métabolique, les couples infertiles de la cohorte présentaient un état métabolique peu favorable, identifiable par une certaine combinaison de ces différents marqueurs métaboliques et anthropométriques. Le statut oxydatif est également cohérent avec l'inflammation persistante associée au syndrome métabolique et, en particulier, l'alpha-carotène et le bêta-carotène chez les hommes et l'alpha-carotène, le rétinol et la lutéine chez les femmes sont inclus dans le modèle restreint. Ces caroténoïdes ont des effets

antioxydants et sont présents dans divers fruits et légumes frais, présents en grande quantité au cours d'un régime alimentaire adéquat (alpha-carotène, bêta-carotène et lutéine) ou dans des aliments d'origine animale (rétinol).

Une modification des facteurs liés à l'alimentation et au mode de vie avant la conception devrait être envisagée en premier lieu en cas d'infertilité inexplicée. Des études interventionnelles prospectives sont désormais nécessaires et en cours pour appuyer le modèle et déterminer la corrélation entre la modification du mode de vie, les résultats calculés par le modèle et les chances de grossesse, à la fois naturellement et après AMP. Les résultats de ce travail montrent également tout l'intérêt de la modélisation multivariée dans la possibilité de faire des allers-retours entre la physiopathologie et les modèles, afin d'identifier des biomarqueurs pertinents et de mieux comprendre les mécanismes biologiques impliqués. Elle permet ainsi de confirmer ou d'infirmer les hypothèses initialement émises.

Fragmentation de l'ADN spermatique – score

La fragmentation de l'ADN correspond à des cassures ou lésions dans le matériel génétique des spermatozoïdes. Ces altérations sont susceptibles de jouer un rôle dans l'infertilité masculine et les issues de grossesses après AMP^{59,60}. Les hommes infertiles présenteraient un taux de fragmentation plus élevé que les hommes fertiles (-1,67 ; IC à 95 % - 2,12 à -1,21 ; 28 études), avec un seuil proposé pour discriminer les hommes infertiles des hommes fertiles de 20 % (AUC 0,844, P < 0,001)⁶¹. La fragmentation de l'ADN peut être la conséquence de causes intrinsèques (surpoids, varicocèle, âge, cryptorchidie, ...) ou extrinsèques (chaleur, tabac, alimentation, alcool, ...), avec comme contributeur majeur le stress oxydatif⁶²⁻⁶⁴. Le lien entre le stress oxydatif et le taux de fragmentation de l'ADN a été reporté à plusieurs reprises⁶⁵⁻⁶⁷. Les résultats présentés ici suggèrent qu'un mode de vie inadéquat, se manifestant par un état métabolique et nutritionnel moins favorable, et objectivé par un score plus bas, entraînerait des conséquences sur la qualité des gamètes. Ce critère qualitatif a été évalué par le taux de fragmentation de l'ADN, dont les patients présentaient un taux moyen au-dessus du seuil retenu pour cette technique. Un lien entre le statut métabolique, le surpoids et la fragmentation de l'ADN des spermatozoïdes a déjà été montré⁶⁸. De plus, Oleszczuk *et al.* ont constaté qu'environ 1 homme sur 5 souffrant d'une infertilité inexplicée avait un taux de fragmentation de l'ADN supérieur à 30%⁶⁹. Les mécanismes impliqués sont nombreux et complexes, soit directs, soit par l'intermédiaire du stress oxydatif. L'évaluation métabolique

globale à l'aide d'un score pourrait potentiellement servir d'indicateur indirect de la fragmentation de l'ADN des spermatozoïdes.

Forces et faiblesses

Ce travail de traitement des données a été effectué rétrospectivement sur les données du PHRC ALIFERT, une étude clinique unique mise en place il y a plus de 10 ans. Une des forces majeures de cette étude est sa nature multicentrique, impliquant des patients de 4 centres en France, réduisant ainsi les biais de recrutement. Malgré la multiplicité de zones géographiques, d'expositions environnementales et de prises en charge, des tendances communes et des facteurs liés au mode de vie et à l'alimentation ont été identifiés comme étant clairement impliqués dans cette pathologie. De plus, l'intégration de marqueurs sanguins moins conventionnels, tels que les vitamines ou les marqueurs du stress oxydatif, qui ne sont pas mesurés en pratique courante, a permis d'explorer de nouvelles pistes et de faire le lien entre différents systèmes et voies biochimiques.

Dans le domaine de l'apprentissage automatique, le phénotypage et l'étiquetage des couples utilisés comme échantillons d'apprentissage sont fondamentaux. La difficulté inhérente à l'infertilité, en particulier l'infertilité inexplicée, est qu'il existe un degré de subjectivité dans le diagnostic, à la différence d'autres maladies pour lesquelles une mutation, par exemple, est parfaitement identifiée. Bien que les modèles d'apprentissage automatique donnent rarement des résultats parfaits, cette difficulté de phénotypage pourrait en partie expliquer les résultats perfectibles obtenus. La nature rétrospective de cette étude, malgré l'existence d'une base de données bien construite et comportant peu de valeurs manquantes, ne nous a toutefois pas permis de remonter dans le temps pour compléter ou corriger des variables ou des patients.

Outre les facteurs liés au mode de vie, les paramètres environnementaux, et en particulier les perturbateurs endocriniens, intéressent de plus en plus la communauté médicale. Nous ne disposons pas de ces informations dans le cadre de cette étude, mais elles auraient sans doute été tout aussi pertinentes. De plus, les dosages par spectrométrie de masse des stéroïdes ont été réalisés ultérieurement : le manque de matériel biologique (sérum) n'a pas permis de réaliser ces dosages chez les femmes. Compte tenu de nos hypothèses concernant l'importance de cette voie métabolique dans la fertilité féminine, de la complexité du diagnostic du SOPK et de son épidémiologie, il semblait possible que certaines patientes de la cohorte ALIFERT présentent un trouble endocrinien apparenté à cette pathologie. Par ailleurs, si l'on

considère les critères d'exclusion de la cohorte ALIFERT, les femmes ne devaient pas être anovulatoires et, compte tenu de l'hétérogénéité des phénotypes du SOPK, nous ne pouvions pas exclure avec certitude la présence de telles patientes dans notre cohorte.

Conclusion et perspectives

La définition de l'infertilité et son caractère inexpliqué sont liés et conditionnés par les résultats du bilan d'infertilité plus ou moins normalisé. Bien que ces examens permettent d'identifier une cause dans une majorité de cas, une part importante des couples infertiles (entre 20 et 30 % selon les études) n'ont pas de réponse et d'explication tangible à leur situation.

Nos résultats et les travaux précédemment publiés sur l'impact des facteurs liés au mode de vie sur l'infertilité suggèrent que d'autres explorations peuvent être pertinentes en vue d'une compréhension plus globale du statut de fertilité des individus et des couples. Ces facteurs étant fortement liés et dépendants les uns des autres, le recours à des méthodes de traitement des données fondées sur l'IA permet un traitement holistique et intégratif de ces informations et pourrait, à terme, conduire à une redéfinition de la notion d'infertilité inexplicée.

Un travail important reste à accomplir, notamment dans la détermination plus précise des mécanismes sous-jacents impliqués, de la temporalité et surtout des effets des actions et interventions nutritionnelles ou sur le mode de vie sur les chances de conception naturelle et sur les résultats en AMP. Un travail en ce sens est en cours de déploiement, avec l'application du score Alifert dans une étude clinique interventionnelle prospective Canadienne menée par le Pr. Baillargeon (Université de Sherbrooke). Dans cette étude multicentrique menée dans 6 centres, nommée "Ferti-Santé", l'objectif était de comparer l'efficacité de l'intervention "Ferti-Santé" par rapport à un groupe témoin recevant un traitement standard. L'intervention "Ferti-Santé" combinait un suivi nutritionnel personnalisé avec un programme d'exercices physiques. L'objectif était de démontrer une augmentation du nombre de grossesses cliniques et une réduction des complications pendant la grossesse chez les femmes bénéficiant de l'intervention "Ferti-Santé". Un certain nombre de données cliniques et biologiques ont été récoltées, lors de plusieurs visites (dont une visite d'inclusion, des visites à 6 et 12 mois puis, si grossesse, deux visites, une en début et une en fin de grossesse). Cette collaboration permettra peut-être de répondre à 2 questions, à savoir 1) l'intervention "Ferti-Santé" permet-elle de modifier et améliorer le score ? et 2) existe-t-il un lien entre la valeur du score et les chances de grossesses ?

Ce travail a donc démontré l'intérêt d'utiliser des méthodes issues de l'IA pour identifier ces troubles métaboliques et nutritionnels à bas bruits. La combinaison de biomarqueurs nouveaux et originaux sous forme de scores pourrait à terme, dans certains cas et pour certains patients, requalifier et redéfinir ces infertilités dites inexplicées, en mettant en cause ces désordres métaboliques complexes secondaires à un mode de vie inadapté. Les concepts de quantification et de seuils, ainsi que le choix des variables à inclure dans ces scores, doivent faire l'objet d'un travail multidisciplinaire approfondi, rigoureux et fondé sur des preuves.

Enfin, l'infertilité inexplicée est, par définition, un défaut de diagnostic. Aux côtés d'autres troubles endocriniens et métaboliques à bas bruit, tels que le SOPK, ils représentent un ensemble voire un continuum de pathologies hautement prévalentes, complexes et multifactorielles, difficiles à caractériser à l'aide d'outils conventionnels. Comme mentionné ci-dessus, il n'a pas été possible d'effectuer des dosages de stéroïdes par spectrométrie de masse chez les femmes de la cohorte ALIFERT. Compte tenu du risque que certaines patientes de la cohorte ALIFERT présentent un trouble endocrinien apparenté au SOPK, la suite de mon travail s'est orientée vers cette pathologie endocrinienne et métabolique extrêmement complexe et prévalente.



OPEN

Proof of concept and development of a couple-based machine learning model to stratify infertile patients with idiopathic infertility

Guillaume Bachelot^{1,16}, Rachel Lévy^{1,2,16}, Anne Bachelot³, Céline Faure¹, Sébastien Czernichow⁴, Charlotte Dupont^{1,2,16}, Antonin Lamazière^{2,5,16}✉ & Alifert Collaborative Group*

We aimed to develop and evaluate a machine learning model that can stratify infertile/fertile couples on the basis of their bioclinical signature helping the management of couples with unexplained infertility. Fertile and infertile couples were recruited in the ALIFERT cross-sectional case–control multicentric study between September 2009 and December 2013 (NCT01093378). The study group consisted of 97 infertile couples presenting a primary idiopathic infertility (> 12 months) from 4 French infertility centers compared with 100 fertile couples (with a spontaneously conceived child (< 2 years of age) and with time to pregnancy < 12 months) recruited from the healthy population of the areas around the infertility centers. The study group is comprised of 2 independent sets: a development set (n = 136 from 3 centers) serving to train the model and a test set (n = 61 from 1 center) used to provide an unbiased validation of the model. Our results have shown that: (i) a couple-modeling approach was more discriminant than models in which men's and women's parameters are considered separately; (ii) the most discriminating variables were anthropometric, or related to the metabolic and oxidative status; (iii) a refined model capable to stratify fertile vs. infertile couples with accuracy 73.8% was proposed after the variables selection (from 80 to 13). These influential factors (anthropometric, antioxidative, and metabolic signatures) are all modifiable by the couple lifestyle. The model proposed takes place in the management of couples with idiopathic infertility, for whom the decision-making tools are scarce. Prospective interventional studies are now needed to validate the model clinical use.

Trial registration: NCT01093378 ALIFERT <https://clinicaltrials.gov/ct2/show/NCT01093378?term=ALIFERT&rank=1>. Registered: March 25, 2010.

Abbreviations

BMI	Body mass index
CI	Confidence interval
PCA	Principal component analysis
OPLS-DA	Orthogonal partial least squares discriminant analysis
ART	Assisted reproductive technology
IUI	Intra-uterine insemination

¹Service de Biologie de La Reproduction-CECOS, Hôpital Tenon, AP-HP/Sorbonne Université, 75020 Paris, France. ²Sorbonne Université, Saint Antoine Research Center, INSERM UMR 938, 75012 Paris, France. ³Service d'Endocrinologie et Médecine de La Reproduction, Centre de Référence Des Maladies Endocriniennes Rares de La Croissance et du Développement, Centre Des Pathologies Gynécologiques Rares, Hôpital Pitié Salpêtrière (APHP), Sorbonne Université, 75013 Paris, France. ⁴Université de Paris, INSERM, UMR1153, Epidemiology and Biostatistics Sorbonne Paris Cité Center (CRESS), METHODS team, Service de Nutrition, Hôpital Européen Georges Pompidou, AP-HP, Paris, France. ⁵Département de Métabolomique Clinique, Hôpital Saint Antoine, AP-HP/Sorbonne Université, 27 Rue Chaligny, 75012 Paris, France. ¹⁶These authors contributed equally: Guillaume Bachelot, Rachel Lévy, Charlotte Dupont and Antonin Lamazière. *A list of authors and their affiliations appears at the end of the paper. ✉email: antonin.lamaziere@aphp.fr

IVF	In vitro fertilization
VIP	Variable Importance for the projection

The World Health Organization (WHO) recognizes infertility as a public health issue defined as the lack of pregnancy after > 12 months of regular unprotected sexual intercourse¹. Infertility affects 12.5% of women and 10% of men².

The diagnosis and infertility management are a painful and long journey, especially for idiopathic or unexplained infertility. The latter condition is defined by a lack of diagnosis for couples that have failed to conceive after one or two years of non-protected sexual intercourse³ and affects 30 to 40% of infertile couples⁴. Standard investigation involving the tests of ovulation, tubal patency, and semen analysis. In addition to generating anxiety in the couple, the diagnostic wavering of idiopathic infertility and empirical treatments represent a massive cost to health care systems⁴. Even if no cause is identified, environment and lifestyle have been suspected as potential idiopathic infertility causes⁵. For both men and women, overweight and metabolic syndrome are recognized to negatively affect fertility^{6,7}, but also on achieving pregnancy after assisted reproductive technology (ART)^{8–10}. Studies showed increased ovulation disorders and miscarriages in obese or overweight women^{11,12}. Likewise, overweight or obesity in men affected sperm parameters but also sperm DNA fragmentation^{13,14}.

Consequently, it is important to focus again on modifiable risk factors among infertile couples using the resources of statistical modeling.

In a previous study of idiopathic infertile couples and fertile couples (the ALIFERT clinical study, NCT01093378), we observed that the sedentary behavior, physical inactivity, body composition, and metabolic disorders are significant factors of idiopathic infertility among men and women^{15–18}. In this previous ALIFERT study, parameters for men and women were considered independently in order to tackle specific infertility issues. However, the mechanisms leading to idiopathic infertility are likely complex and multifactorial and the couple man + woman is also considered presently.

In the recent years, alongside deep learning and neuronal networks, machine learning computational models such as the Bayesian network, random forest, and partial least square (PLS) have driven substantial advances in the integration and treatment of complex bio-clinical signatures^{19,20}. In this study, we propose to evaluate the performance of a couple-based Orthogonal Partial Least Square (OPLS) model facilitating the medical decision-making process in the context of idiopathic infertility. Our hypothesis is such a multivariate model, in which the man + woman couple formed a single observation and, in which the biochemical, clinical, and anthropometric variables of both men and women are merged together can discriminate efficiently between fertility and idiopathic infertile.

To our knowledge, this study in the field of idiopathic infertility classification is the first to evaluate comprehensively a machine-learning algorithm and to assess the generalizability of this model across various diagnostic centers. The proven possibility prompts for a future interventional study.

Materials and methods

Subjects. Fertile and infertile couples were recruited in the ALIFERT cross-sectional case–control study between September 2009 and December 2013¹⁸ (National biomedical research Id. P071224, ethics committee approval [“Comité de Protection des Personnes”] AOM 2009-A00256-51, NEudra CT 08180, clinicaltrials.gov NCT01093378). The clinical study was performed in accordance with relevant guidelines/regulations. Informed consent was obtained from all participants. The research involving human research participants was performed in accordance with the Declaration of Helsinki.

ALIFERT was designed as a multicentric study. The study group consisted of 97 infertile couples (presenting a primary idiopathic infertility > 12 months) recruited from four infertility centers in France (Hôpital Jean-Verdier ART center, Bondy (JV); Hôpital Nord ART center, Saint-Étienne (SE); Polyclinique de Navarre ART center, Pau (PAU); Cochin Hospital ART center, Paris (CCH)) and 100 fertile couples (with a spontaneously conceived child under 2 years of age with time to pregnancy less than 12 months) recruited from the general population in areas of the participating centers. Men were under 45 years of age and women under 38 years of age.

Infertile couples presented a infertility > 12 months, of which the primary idiopathic character is based on the following investigations: Men did not present severe sperm alteration nor urogenital pathology, Women did not present neither anovulation, ovarian failure on the basis of follicle count nor hormone balance at day 3 (FSH, LH, and estradiol) nor uterotubal pathology assessed by hysterosalpingography. Patients with current or previous proven metabolic or digestive disease were excluded.

The control group had a spontaneously conceived child under 2 years of age with time to pregnancy less than 12 months. Eligibility criteria have been detailed previously¹⁸.

The study group was comprised of two independent cohorts: a development set serving to train the model (n = 136), including 73 infertile and 63 fertile couples from JV, SE, and PAU, and a test set (n = 61), including 24 infertile and 37 fertile couples from CCH. The development set was used to tune the machine-learning model, and the external validation set was used to evaluate the model performance. Further model refinement was also conducted on the overall cohort (n = 196) couples.

Input data acquisition. Written informed consent was collected for both fertile and infertile couples. To note, all variables collected from women were encoded with a prefix of “w” before the variable name (i.e. *w_variable*). Variables without prefix were collected from men.

Assessments. Study and control subjects have been assessed by the same trained investigators using the same calibrated devices.

Anthropometric assessments. The investigator measured height, weight (Tanita BC-420MA analyzer), and waist circumference measured at the narrowest point between the lower border of the ribs and the iliac crest.

Blood pressure assessment. Systolic and diastolic blood pressures have been measured using a sphygmomanometer cuff after 5 min rest around patient forearm in a supine position. The systolic and diastolic pressures were the mean of right and left values.

Carbon monoxide status. Exhaled carbon monoxide (CO) has been measured in parts per million (ppm), as a supportive indicator with the underlying assumption that exhaled CO in smokers²¹ is higher than in non-smokers. Exhaled CO measurement was performed by having subjects exhale completely then inhale fully in open air, withhold their breath for 10 s, and then exhale completely into the portable CO monitor (Tabataba analyser-FIM medical, Villeurbanne 69625 France)²¹.

Blood samples and analyses. Blood samples were collected after 12 h fasting for measurement of fresh plasma total cholesterol, high-density lipoprotein (HDL-cholesterol), low-density lipoprotein (LDL-cholesterol), triglycerides and glucose. Serum and plasma were stored at -80°C until further analyses.

LC-MS/MS steroid profiling. Steroid profiles were measured in serum by liquid chromatography–mass spectrometry (LC-MS/MS) as described elsewhere²².

Antioxidants and micronutrients. Serum vitamin D (ng/mL), vitamin B9 (folic acid-erythro) (nmol), vitamin B9 (nmol/L), vitamin B12 (pM), alpha-Tocopherol (vitamin E) (mmol/L), Zinc (mmol/L), Selenium (mmol/L), Vitamin C (mg/mL), alpha-Carotene (mmol/L), beta-carotene (mmol/L), Lycopene, Lutein, Zeaxanthin, beta-Kryptoxanthin and Retinol (vitamin A) were assayed at Département de Biologie intégrée – Biologie nutritionnelle et stress oxydatif (Grenoble hospital avenue du Maquis du Grésivaudan BP 217—38043 Grenoble Cedex 09).

Serum ascorbic acid (vitamin C) was determined by using an automated method based on the principle of continuous flow. Serum retinol, tocopherol and carotenoids (lutein, zeaxanthin, beta-cryptoxanthin, lycopene, alpha-carotene and beta-carotene) were measured by HPLC (Biotek-Kontron, Montigny-le-Bretonneux, France). Serum zinc concentrations were measured by using flame atomic absorption spectrometry (model 3110; Perkin Elmer, Norwalk, CT) and selenium by atomic absorption spectrometry (4100 ZL; Perkin Elmer)^{23,24}.

In total 50 variables from men and 30 from women were collected and input in the models.

Statistical analysis and data processing. The data sets were tested with Shapiro–Wilks in order to evaluate their distribution (Supplementary Tables 1a–1d)²⁵. When the distribution was not normal, a non-parametric Wilcoxon–Mann–Whitney test was applied²⁶. If the distribution was normal, a t-test was performed. These tests were performed using the Rstudio® software (version 3.6.1). The data were loaded into SIMCA 15° software (version 15, Umetrics, Västerbotten, Sweden) in order to perform the Principal Component Analysis (PCA) and Orthogonal Partial Least Square–Discriminant Analysis (OPLS-DA). (For more details see Supplementary material section).

We calculated the model starting from 50 variables in men, and 30 variables in women. Furtheron we have matched the 2 gender sets to generate a so-called “couple” matrix comprised of 80 variables. To our knowledge, the implementation of an algorithm taking into account the combined parameters of the matched man-woman couple has not been yet tested for infertility.

PCA, an unsupervised method allowing the data dimension reduction and the exclusion of outliers was applied to the development set. After the removal of 1 outlying couple the discriminant analysis OPLS-DA has served to calculate the regression of the multiple factors *versus* the couple fertility/intertility status. The supervised multivariate machine learning algorithm^{27,28}, was trained using the development set ($n = 136$), the model’s performance being evaluated using the external validation set ($n = 31$). Based on various studies, the OPLS approach is considered particularly appropriate to circumvent difficulties of multiple variables colinearity and missing data, and has already been used in healthcare^{29–31}. Eventually, we also trained and evaluated 4 other machine-learning models (Support vector machine, Nearest Neighbors Classifier, Decision Tree and logistic regression using Python 3.8.2, Scikit-Learn library 0.22.2, Numpy library 1.18.1, and Pandas library 1.0.1) to confirm the discriminative power of the data, independently of the chosen method.

As described in Fig. 1, the OPLS model was evaluated for the goodness of fit (R^2) and the capability to predict using an internal cross-validation (Q^2)³². R^2 and Q^2 values are between 0 indicating a poor fit or poor capability to predict and 1 (highest fit or prediction by the model). An independent data set of 31 couples (*i.e.* from the CCH institution) was used to evaluate the model performance (*i.e.* accuracy) by external validation.

In order to reduce the dimensionality of our system and to generate a more “user-friendly” model, the least relevant variables were eliminated in the model using their Variable Importance for Projection (VIP).

Ethical approval. The ethics committee (“Comité de Protection des Personnes”, Paris France) approved the study. ALIFERT study (national biomedical research P071224/AOM 08,180:NEudra CT 2009-A00256-51/clinical trials NCT01093378). All the participants signed a written informed consent.

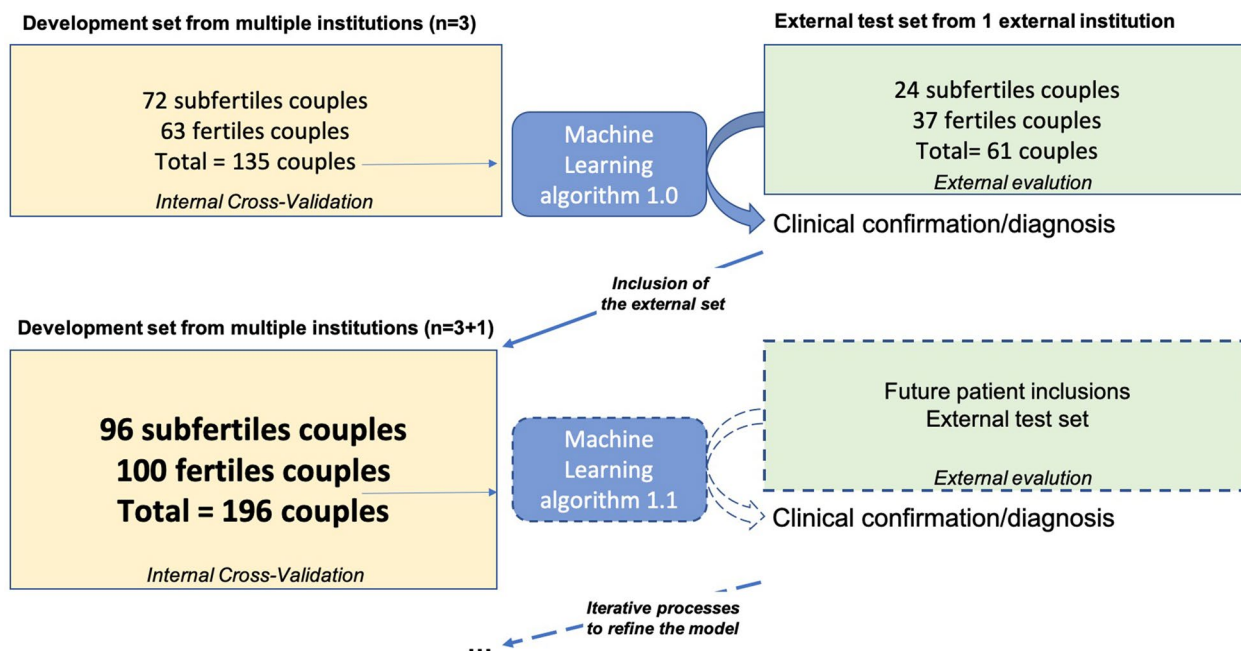


Figure 1. Data flow diagram for the study: model training iteration. The model was trained and tuned after variable selection on a development set originating from patients of multiple specialized centers. Firstly the model was internally evaluated using cross-validation. Then, a test set of couples originating from a different institution was used for an external validation. The model predictions were matched with the expected clinical output. Model training iteration (i.e. introduced externally) process was eventually performed to build a refined model with the selected features.

Results

Population description. Idiopathic infertile couples men and women showed significant differences in the anthropometric and biological/metabolic parameters compared to fertile couples. For example, BMI, waist circumference, visceral fat, and glycaemia, were higher in infertile men and women (Supplementary Table 1a, 1b) but HDL lower.

Out of 13 micronutrients and vitamins with an antioxidative activity, retinol, alpha-carotene, lutein, and beta-carotene levels were lower in infertile men compared to fertile (Supplementary Table 1c). In infertile women, alpha-carotene, beta-carotene, lycopene, and lutein levels were also lower than in fertile (Supplementary Table 1c). Following the usual recommendations, infertile women are supplemented with folate during the pre-conception period³³, which explained why their plasma folate levels were higher compared to fertile. The 2 groups were comparable in terms of steroids except for testosterone lower in infertile men (Supplementary Table 1d).

Overview of the ALIFERT dataset with principal component analysis. A principal component analysis (PCA) was first performed to reduce the large variety of biological differences along the prominent variation trends (Supplementary Fig. 1). A PCA showing the scores of the men, women and couples (respectively A, B and C) projected on the first factorial plan (t1, t2) does not show the separation between fertile and infertile groups. The percentages of explained variance along the two most prominent components (i.e. the two most influential combinations of variables) were < 0.5 (R^2 men = 0.264, R^2 women = 0.307, R^2 couples = 0.192), which suggested weak correlations between the variables. The score plots for men and couples showed an outlier (the patient (man) with BMI 45, was above average by more than 8 standard deviations). The corresponding couple was excluded from the further model calculations.

Relevance of the couple approach in the model versus single men or women. In order to test the advantage of a couple-based approach, OPLS models were separately calculated from men, women and the merged data from the two matched genders (i.e. the couple-based approach) (Supplementary Fig. 2A–C). The discrimination of the alternative fertility status is clearly revealed by the separation of red and green (fertile) symbols. The overlapping scores between the discriminated fertility status was only 7 subfertile and 4 fertile couples out of 135 observations from the development set.

Each of the three models described previously was tested on the independent validation set: the corresponding patients were plotted (black stars) regardless of their fertility (Supplementary Fig. 2–F) and secondly, their symbols were colored according to their fertile (green stars) or infertile status (red stars) (Fig. 2A–C). When the predicted status of the validation observations as based on the left- or right-quadrants position in the score plots are checked out, the model accuracy can be estimated to 59.0%, 57.4%, and 68.8% for the men, women, and couple-based models, respectively. Accordingly, the couple-based data set led to a more accurate and powerful

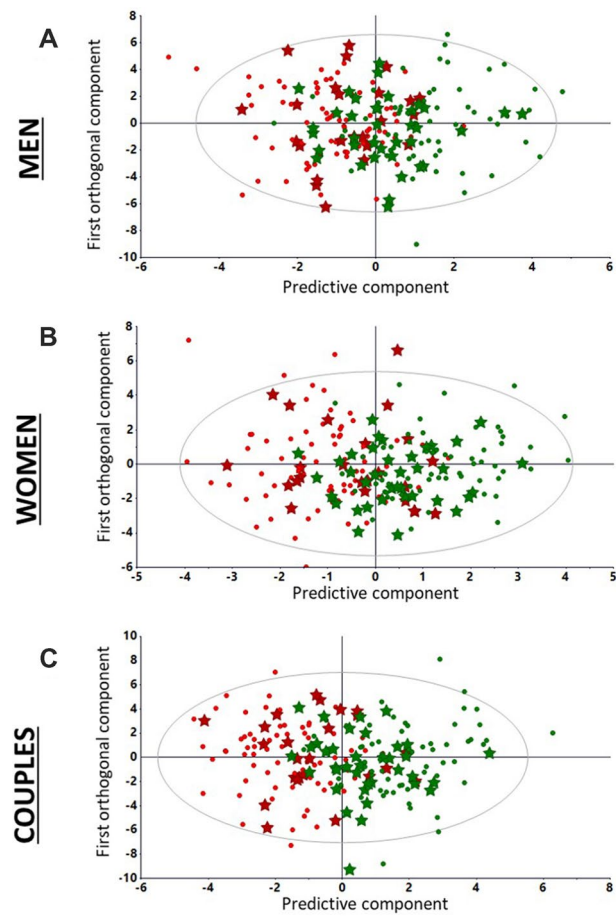


Figure 2. Orthogonal partial least squares discriminant analysis models (training and evaluation) calculated from men, women and couples development sets. **(A)** External evaluation of men fertility status: labelled external test set observations (fertile and infertile as green and red stars, respectively) are projected on the score plots generated from fertile (green dots) and infertile (red dots) men (135 × 50) set. R^2 (fit) and Q^2 (prediction capability) are 0.471 and 0.322, respectively. **(B)** External evaluation of women fertility status: labelled external test set observations (fertile and infertile as green and red stars, respectively) are projected on the score plots generated from fertile (green dots) and infertile (red dots) women (135 × 30) development set. R^2 and Q^2 were 0.560 and 0.462. **(C)** External evaluation of couples: labelled external test set observations (fertile and infertile as green and red stars, respectively) are projected on the score plots generated from fertile (green dots) and infertile (red dots) couples (135 × 80) development set. R^2 and Q^2 were 0.624 and 0.487. Accuracy scores are 0.590, 0.574 and 0.688 for models based on men, women and couples data, respectively.

model than the separate men and women models. We trained and evaluated 4 other machine-learning models (Support vector machine, Nearest Neighbors Classifier, Decision Tree and logistic regression) to confirm the discriminative power of the data, independently of the chosen method : as reported in supplemental Table 2, their performances were either worse or equivalent with OPLS-DA. These results confirmed the decision to adopt the OPLS-DA algorithm.

The Variable Importance for the Projection (VIP) histogram corresponding to the model summarized the relative influence of the 80 variables (Supplementary Fig. 3). In order to reduce the number of non-influential features regarding the discrimination of the fertility status, to limit the colinearity between redundant variables, to reduce the “noise” created the variance of less discriminant variables and to create a more parcimonious “user-friendly” model, we have proceeded to an iterative selection of the features as a function of the VIP (the higher the VIP, the more discriminatory the variable between the fertile/infertile groups).

Therefore, we successively selected features with a VIP score > 1 to obtain 24- (intermediate) and then 13-feature reductionist models.

Refined models after selection of the fertility discriminating variables. The iterative procedure has consisted to suppress in a stepwise maner the less influential variables with weak VIP and to calculate the accuracy of the resulting discriminant model for the couple fertility status prediction in the validation set. An intermediate model was built with 24 variables, of which the score plot exhibited a segregation between infertile and fertile groups (Supplementary Fig. 4A). The summary of the variable importance of the projection (VIP) histogram was calculated for the 24 variables (Supplementary Fig. 4B). Following iterative variable reduction/

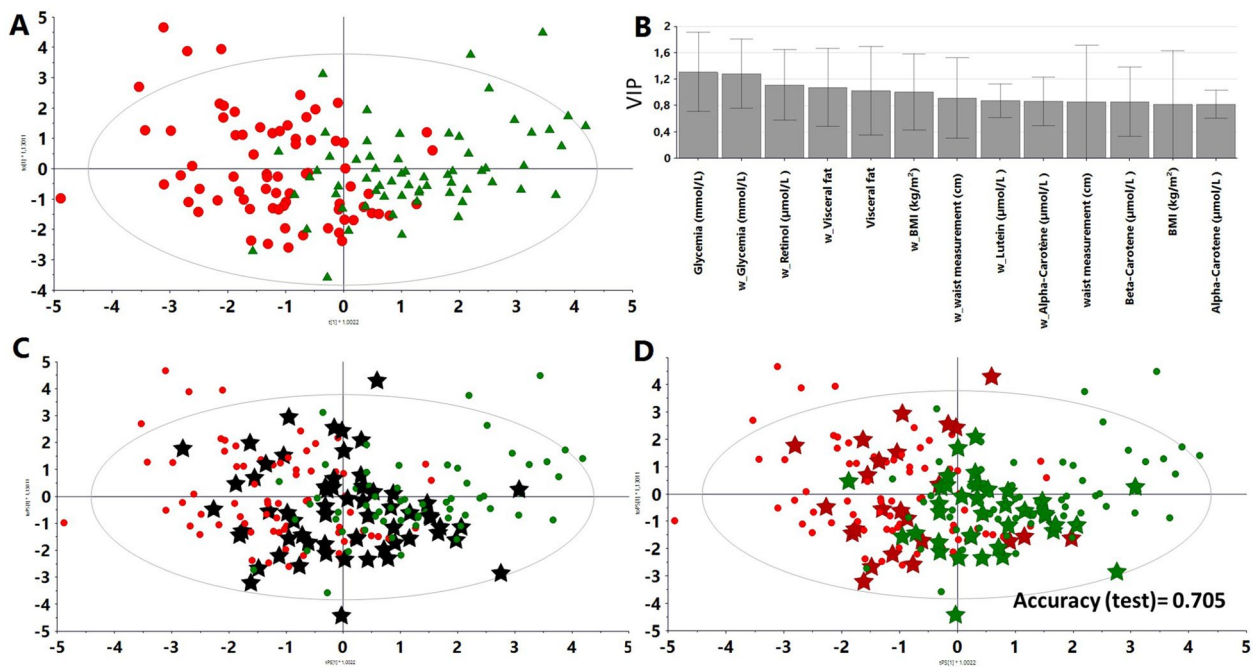


Figure 3. Supervised Orthogonal partial least squares discriminant (OPLS-DA) model with 13 variables selected from the 24 variables model with variable importance for projections (VIP) > 1: reductionist model. (A) Score plot generated from fertile (labeled as green triangles) and infertile couples (labeled as red dots). R^2 and Q^2 were 0.467 and 0.410, respectively. Accuracy (internal validation) = 0.838 and calculated on the test set = 0.705. (B) VIP (Variable Importance for the Projection) the plot summarized the influence of the 13 variables on the model. (C) Score plot reductionist model from the merged development and test sets (61 couples) (black stars). (D) Score plot reductionist model from the combined development and test sets. Scores are labeled according to their fertility status (green or red).

model validation, a discriminant model was calculated with 13 variables (Fig. 3A), of which the relative importance in the projection (VIP) was shown (Fig. 3B).

This model was then tested on the independent external test set: the corresponding couples were plotted (black stars) regardless of their label (Fig. 3C). Secondly, these couples were colored according to their fertile (green stars) or infertile (red stars) status (Fig. 3D). The accuracy score on the independent test set was 73.8%, and was therefore superior to that of the full-feature development set (68.8%). Based on these criteria, the model comprised of 13 variables was judged the most suitable to classify couples.

High contribution of anthropometric, metabolic, and antioxidative parameters. Glycaemia, for both women and men, appeared to be the most important variable for scoring fertility. Circulating retinol also played a significant role as did to a lesser extent, the other antioxidative species such as lutein, beta-carotene, and alpha-carotene exhibiting a high VIP score.

Variations of anthropometric parameters such visceral fat, body mass index, and waist measurement in both women and men were among the most significant parameters in the model. Surprisingly steroid hormones and various antioxidative species, including lutein and selenium, which showed less relevance in the intermediate discriminant analysis with 24 variables were excluded from the refined model.

Discussion

In this study, we proposed a proof of concept of a machine-learning model that classify fertile and infertile couples according to anthropometric, antioxidative, and metabolic signatures. Besides of thoroughly documented causes of infertility such as genetic, congenital, and acquired defects, increasing interest should also be focused on the environmental and lifestyle factors²¹. Indeed, addictions to tobacco, alcohol, and drugs, as well as diet, physical activity, obesity, metabolic disorders, sleep, and stress, also compromise male and female fertility^{5,18,34–37}. Interestingly, the impact of these factors is reversible if appropriate corrective measures are taken. These risk factors are commonly studied separately, while their effects may be cumulative in combination³⁸. Therefore, it seemed helpful to build a model including the several lifestyle-related factors as well as the two members of the couple to quantitatively predict the infertility degrees. Furthermore, this model could facilitate and refine the medical decision-making process on the basis of a limited number of variables in the complex context of idiopathic infertility.

With the present model, a couple is considered a single entity rather than as two independent individuals. Our validation on an independent cohort showed that: (i) the couple-modeling approach was indeed more discriminant than a model in which men's and women's are considered separately; (ii) the most important variables

for the projection were related to metabolic disorders, oxidative status, and anthropometry; and (iii) after reduction of our system dimension (from 80 down to 13 variables), we proposed a new algorithm to discriminate fertile vs. infertile patients. Among the 13 variables of interest retained at the end of the model, anthropometric, antioxidative, and metabolic parameters are the main variables that can correlate with idiopathic infertility. Elevated blood glucose, lowered HDL cholesterol, anthropometric parameters related to abdominal obesity (but no blood pressure) are suggestive of metabolic syndrome and discriminating of both men and women fertility status. The oxidative status is also consistent with the prolonged mild inflammation associated with the metabolic syndrome and importantly alpha-carotene and beta-carotene in men, as well as alpha-carotene, retinol, and lutein in women, are included in the refined model comprised of the most influential variables. These carotenoids have antioxidant actions and are present associated in a variety of fresh fruits and vegetables abundant in an appropriate diet (alpha-carotene, beta-carotene, and lutein) or in foods of animal origin (retinol).

Therefore, changing pre-conception nutritional and lifestyle factors should be considered a first-choice for unexplained infertility³⁹. We observed a small overlap between the scores of fertile and infertile couples—and assume that patients will benefit of improvements to lifestyle factors to obtain live birth¹⁶. Prospective interventional studies are now needed to support the model and determine how lifestyle change the model calculated scores and correlate with the chances of pregnancy.

The definition of idiopathic infertility being related to the results of a standard protocol these observations lead potentially to a renewal of idiopathic infertility definition by including lifestyle parameters.

The reductionist model built from 135 couples has already shown its predictive capacity within an external cohort with a slight overfitting (83.8% on the training set vs. 73.8% on the test set). In addition, upon increasing the number of couples (from 136 to 196), a moderate decrease in accuracy on the training set (83.8–80.3%) was observed, likely reflecting an overfitting reduction. Afterwards, the final machine-learning model that is trained using the entirety of the data set (196 couples) will require further validation using an independent data set. Furthermore, this study focused only on couples with idiopathic primary infertility.

Conclusion

The proposed model is a multivariate supervised statistical model built from data drawn from infertile men and women who are considered as a unique couple in observation. The couple approach yielded a significant gain in discriminatory power compared to models in which analysis of men or women are implemented separately.

The proposed model could find its place in the standard care management of couples with parental project. This approach is interesting since personalized lifestyle intervention should be considered as a first-choice treatment for idiopathic infertility and should be systematically suggested before ART. Prospective interventional studies will be needed to test the hypothesis demonstrated by this algorithm and to validate this model for clinical use. Moreover, it would also be interesting to assess whether this model would be effective in tracking the efficiency of potential lifestyle modifications.

Received: 8 July 2021; Accepted: 18 November 2021

Published online: 14 December 2021

References

- Zegers-Hochschild, F. *et al.* International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertil. Steril.* **92**, 1520–1524 (2009).
- Datta, J. *et al.* Prevalence of infertility and help seeking among 15 000 women and men. *Hum. Reprod.* **31**, 2108–2118 (2016).
- Ray, A., Shah, A., Gudi, A. & Homburg, R. Unexplained infertility: An update and review of practice. *Reprod. Biomed. Online* **24**, 591–602 (2012).
- Smith, S., Pfeifer, S. M. & Collins, J. A. Diagnosis and management of female infertility. *JAMA* **290**, 1767–1770 (2003).
- Piché, M. L., Babineau, V., Robitaille, J., Lachance, É. & Ruchat, S. M. Lifestyle-related factors associated with reproductive health in couples seeking fertility treatments: Results of a pilot study. *Int. J. Fertil. Steril.* **12**, 19–26 (2018).
- Broughton, D. E. & Moley, K. H. Obesity and female infertility: Potential mediators of obesity's impact. *Fertil. Steril.* **107**, 840–847 (2017).
- Michalakakis, K., Mintziori, G., Kaprara, A., Tarlatzis, B. C. & Goulis, D. G. The complex interaction between obesity, metabolic syndrome and reproductive axis: A narrative review. *Metabolism* **62**, 457–478 (2013).
- Mushtaq, R. *et al.* Effect of male body mass index on assisted reproduction treatment outcome: An updated systematic review and meta-analysis. *Reprod. Biomed. Online* **36**, 459–471 (2018).
- Campbell, J. M., Lane, M., Owens, J. A. & Bakos, H. W. Paternal obesity negatively affects male fertility and assisted reproduction outcomes: A systematic review and meta-analysis. *Reprod. Biomed. Online* **31**, 593–604 (2015).
- Rittenberg, V. *et al.* Effect of body mass index on IVF treatment outcome: An updated systematic review and meta-analysis. *Reprod. Biomed. Online* **23**, 421–439 (2011).
- Rich-Edwards, J. W. *et al.* Physical activity, body mass index, and ovulatory disorder infertility. *Epidemiology* **13**, 184–190 (2002).
- Ramlau-Hansen, C. H. *et al.* Subfecundity in overweight and obese couples. *Hum. Reprod.* **22**, 1634–1637 (2007).
- Sermondade, N. *et al.* Obesity and increased risk for oligozoospermia and azoospermia. *Arch. Intern. Med.* **172**, 440–442 (2012).
- Dupont, C. *et al.* Obesity leads to higher risk of sperm DNA damage in infertile patients. *Asian J. Androl.* **15**, 622–625 (2013).
- Dupont, C. *et al.* Metabolic syndrome and smoking are independent risk factors of male idiopathic infertility. *Basic Clin. Androl.* **29**, 9 (2019).
- Faure, C. *et al.* In subfertile couple, abdominal fat loss in men is associated with improvement of sperm quality and pregnancy: A case-series. *PLoS ONE* **9**, e86300 (2014).
- Faure, C. *et al.* Are superoxide dismutase 2 and nitric oxide synthase polymorphisms associated with idiopathic infertility?. *Antioxid. Redox Signal* **21**, 565–569 (2014).
- Foucaut, A. M. *et al.* Sedentary behavior, physical inactivity and body composition in relation to idiopathic infertility among men and women. *PLoS ONE* **14**, e0210770 (2019).
- Peterson, E. D. Machine learning, predictive analytics, and clinical practice: Can the past inform the present?. *JAMA* **322**, 2283–2284 (2019).

20. Shah, N. H., Milstein, A. & Bagley, P. Making machine learning models clinically useful. *JAMA* **322**, 1351–1352 (2019).
21. Deveci, S. E., Deveci, F., Açıık, Y. & Ozan, A. T. The measurement of exhaled carbon monoxide in healthy smokers and non-smokers. *Respir. Med.* **98**, 551–556 (2004).
22. Fiet, J. *et al.* A liquid chromatography/tandem mass spectrometry profile of 16 serum steroids, including 21-deoxycortisol and 21-deoxycorticosterone, for management of congenital adrenal hyperplasia. *J. Endocr. Soc.* **1**, 186–201 (2017).
23. Czernichow, S. *et al.* Effects of long-term antioxidant supplementation and association of serum antioxidant concentrations with risk of metabolic syndrome in adults. *Am. J. Clin. Nutr.* **90**, 329–335 (2009).
24. Akbaraly, N. T., Faure, H., Gourlet, V., Favier, A. & Berr, C. Plasma carotenoid levels and cognitive performance in an elderly population: Results of the EVA Study. *J. Gerontol. A Biol. Sci. Med. Sci.* **62**, 308–316 (2007).
25. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
26. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947).
27. Garali, I. *et al.* A strategy for multimodal data integration: Application to biomarkers identification in spinocerebellar ataxia. *Brief Bioinform.* **19**, 1356–1369 (2018).
28. Pérez-Enciso, M. & Tenenhaus, M. Prediction of clinical outcome with microarray data: A partial least squares discriminant analysis (PLS-DA) approach. *Hum. Genet.* **112**, 581–592 (2003).
29. Tenenhaus, M. L. *régession PLS: théorie et pratique* (Editions Technip, 1998).
30. Skorupa, A. *et al.* Grading of endometrial cancer using 1H HR-MAS NMR-based metabolomics. *Sci. Rep.* **11**, 18160 (2021).
31. Thévenot, E. A., Roux, A., Xu, Y., Ezan, E. & Junot, C. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *J. Proteome Res.* **14**, 3322–3335 (2015).
32. Eriksson, L., Trygg, J. & Wold, S. CV-ANOVA for significance testing of PLS and OPLS* models. *J. Chemom.* **22**, 594–600 (2008).
33. US Preventive Services Task Force recommendation statement. U.S. Preventive Services Task Force. Folic acid for the prevention of neural tube defects. *Ann. Intern. Med.* **150**, 626–631 (2009).
34. Sharpe, R. M. & Irvine, D. S. How strong is the evidence of a link between environmental chemicals and adverse effects on human reproductive health?. *BMJ* **328**, 447–451 (2004).
35. Mínguez-Alarcón, L., Chavarro, J. E. & Gaskins, A. J. Caffeine, alcohol, smoking, and reproductive outcomes among couples undergoing assisted reproductive technology treatments. *Fertil. Steril.* **110**, 587–592 (2018).
36. Salas-Huetos, A., Bulló, M. & Salas-Salvadó, J. Dietary patterns, foods and nutrients in male fertility parameters and fecundability: A systematic review of observational studies. *Hum. Reprod. Update* **23**, 371–389 (2017).
37. Chiu, Y.-H., Chavarro, J. E. & Souter, I. Diet and female fertility: Doctor, what should I eat?. *Fertil. Steril.* **110**, 560–569 (2018).
38. Hassan, M. A. & Killick, S. R. Negative lifestyle is associated with a significant reduction in fecundity. *Fertil. Steril.* **81**, 384–392 (2004).
39. Dupont, C. *et al.* Effectiveness of a therapeutic multiple-lifestyle intervention taking into account the periconceptual environment in the management of infertile couples: Study design of a randomized controlled trial—the PEPCI study. *BMC Preg. Childbirth* **20**, 322 (2020).

Acknowledgements

The authors would like to thank Clinical Research Unit Paris-centre coordinator Christelle Auger and CRA Deborah Rechar. We also express special thanks to Dr Claude Wolf for his attentive proofreading and technical support.

Author contributions

G.B. performed statistical analysis, participated in the study conception and design and drafting the manuscript. R.L. supervised the study, participated in study conception and design, patients recruitment, interpretation of data, critical revision of the manuscript for intellectual content. C.F. participated in study conception and design, patients recruitment, interpretation of data, critical revision of the manuscript for intellectual content. S.C. participated in study conception and design and critical revisions of the manuscript for intellectual content. C.D. participated in study conception and design, patients recruitment, interpretation of data, and drafting the manuscript. A.L. participated in study conception and design, statistical analysis, interpretation of data, and drafting the manuscript. The collaborators of the ALIFERT collaborative group participated in study design and were involved in patients' recruitment or analyses. All authors read and approved the final manuscript.

Funding

This study was supported by national biomedical research P071224 ALIFERT.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03165-3>.

Correspondence and requests for materials should be addressed to A.L.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Alifert Collaborative Group

Isabelle Aknin⁶, Isabelle Cedrin-Durnerin⁷, Steven Cens⁸, Nathalie di Clemente², Jean-Louis Guéant⁹, Serge Hercberg¹⁰, Yoann Lalatonne¹¹, Chrystèle Racine², Nathalie Sermondade¹, Angela Sutton¹², Claude Uthurriague¹³, Jean-Philippe Wolf¹⁴ & Alain Favier¹⁵

⁶Unité Fonctionnelle de Biologie de La Reproduction, Histologie—Embryologie—Cytogénétique, Hôpital Nord, Saint-Étienne, France. ⁷Service de Médecine de La Reproduction, Hôpital Jean Verdier, APHP, Bondy, France. ⁸Centre d'AMP de PAU, Polyclinique de Navarre, Pau, France. ⁹Laboratoire de Biochimie, CHU Brabois de Nancy, Université de Lorraine, Lorraine, France. ¹⁰EREN, INSERM U557; INRA; CNAM; Université Paris 13, CRNH IdF, 93017 Bobigny, France. ¹¹Laboratoire de Médecine Nucléaire, 125 Route de Stalingrad, 93009 Bobigny, France. ¹²Laboratoire de Biochimie, Hôpital Jean Verdier, APHP, Paris, France. ¹³Centre d'AMP de PAU, Polyclinique de Navarre, Pau, France. ¹⁴Service d'Histologie-Embryologie-Biologie de La Reproduction, Hôpital Cochin, APHP, Paris, France. ¹⁵Département de Biologie Intégrée, Grenoble Hospital, 38043 Grenoble Cedex 09, France.



Open Access

ORIGINAL ARTICLE

Machine learning approach to assess the association between anthropometric, metabolic, and nutritional status and semen parameters

Guillaume Bachelot^{1,2,3}, Antonin Lamaziere^{1,3}, Sebastien Czernichow⁴, Celine Faure², Chrystelle Racine¹, Rachel Levy^{1,2}, Charlotte Dupont^{1,2}, Nutrition and Fertility (ALIFERT) Group*

Many lifestyle factors, such as nutritional imbalance leading to obesity, metabolic disorders, and nutritional deficiency, have been identified as potential risk factors for male infertility. The aim of this study was to evaluate the relationship between semen parameters and anthropometric, metabolic and nutritional parameters. Relationship was first assessed individually, then after the application of a previously constructed and validated machine learning score that allows their combination. Anthropometric, metabolic, antioxidant, micronutrient, and sperm parameters from 75 men suffering from idiopathic infertility from four infertility centers in France (Jean-Verdier ART Center Hospital, Bondy; North Hospital ART Center, Saint-Étienne; Navarre Polyclinic ART Center, Pau; and Cochin Hospital ART Center, Paris) between September 2009 and December 2013 were collected. After assessing standard correlation analysis, a previously built machine learning model, providing a score ranging from 0 (the poorest) to 1 (the most favorable), was calculated for each man in the study cohort. This machine learning model, which separates infertile/fertile men with unexplained infertility on the basis of their biochemical signature, provides a more holistic evaluation of the influence of the considered markers (anthropometric, metabolic, and oxidative status). We observed a significant correlation of some anthropometric, metabolic, and nutritional disorders with some sperm characteristics. Moreover, an unfavorable machine learning score was associated with a high level of sperm DNA fragmentation. Favorable anthropometric, metabolic, and oxidative patterns, which may reflect an appropriate lifestyle, appear to positively impact overall health, in particular reproductive function. This study, consistent with previous publications, suggests that beyond semen quality parameters, in an essential assessment of male fertility, other key factors should be taken into account. In this regard, the application of emerging artificial intelligence techniques may provide a unique opportunity to integrate all these parameters and deliver personalized care.

Asian Journal of Andrology (2024) 26, 1–7; doi: 10.4103/aja20247; published online: 12 April 2024

Keywords: lifestyle; machine learning; metabolism; nutrition; sperm DNA fragmentation

INTRODUCTION

As recently published by Levine *et al.*,¹ fertility has been declining worldwide for the past half century, especially sperm quality. Many lifestyle factors have been identified as potential risk factors for infertility in men. Among them, overweight and obesity have been particularly studied.² A global increase in the prevalence of obesity has been observed over the last three decades, and now, more than half of men in reproductive age are currently considered to be overweight or obese. It has been well established that overweight and obesity are associated with male infertility³ and have a negative impact on semen parameters, including sperm count^{2,4} and sperm DNA fragmentation.^{5,6} In addition, metabolic disorders that are often the consequence of an unbalanced diet and/or insufficient physical activity also contribute to impaired reproductive function.⁷

Recently, our team has highlighted that not only metabolic syndrome but also metabolic disorders (increased fasting blood

glucose) and anthropometric disorders (increased body mass index [BMI] and abdominal obesity) may be risk factors for idiopathic infertility.⁸ We also observed that male partners of infertile couples were less physically active than fertile men.⁹ They also had lower plasma levels of antioxidant vitamins.¹⁰

There are many complex mechanisms involved in these phenomena. Overweight, obesity, and metabolic disorders can lead to impaired functioning of the hypothalamic–pituitary–gonadal (HPG) axis. The process of converting steroids into estrogens in peripheral adipose tissue leads to an elevation in estradiol levels, which exerts a negative feedback loop at the hypothalamic–pituitary level. In these men, secondary hypogonadism is therefore observed, characterized by a decrease in serum concentrations of follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone.^{11,12} Hyperinsulinemia leads to a decrease in hepatic production of sex

¹Sorbonne University School of Medicine, Saint-Antoine Research Center, INSERM UMR 938, 27 rue Chaligny, Paris 75012, France; ²Reproductive Biology Department-CECOS, Tenon Hospital, AP-HP, Sorbonne University, Paris 75020, France; ³Clinical Metabolomics Department, Saint-Antoine Hospital, AP-HP, 27 rue Chaligny, Paris 75012, France; ⁴Nutrition Department, Obesity Specialist Centre, Georges Pompidou European Hospital, AP-HP, Paris 75015, France.

*A list of authors and their affiliations appears at the end of the article.

Correspondence: Dr. G Bachelot (guillaume.bachelot@aphp.fr)

Received: 12 December 2023; Accepted: 29 January 2024

hormone-binding globulin (SHBG).¹³ Consequences are increasing levels of free testosterone available for conversion to estradiol. Chronic systemic inflammation, which plays a critical role in the development of metabolic syndrome, is associated with obesity.¹⁴ Levels of inflammatory cytokines (interleukin 1-alpha [IL-1A], interleukin-6 [IL-6], tumor necrosis factor- α [TNF- α], activin A, *etc.*), which are essential for normal spermatogenesis, are disturbed.¹⁵ Systemic oxidative stress is also associated with increased body weight and obesity.¹⁶ Oxidative stress and lipid peroxidation are increased in the testicular microenvironment following the accumulation of adipose tissue.¹⁷ Furthermore, high blood glucose levels have been shown to increase oxidative stress through mitochondrial oxidation of glucose, which releases a substantial amount of free radicals into the cytosol.¹⁸ In addition, oxidative stress leads to the production of nitric oxide (NO) which oxidizes sperm membrane lipids. Oxidative sperm impairs spermatogenesis and sperm motility.¹⁹ Germ cells are thought to be more sensitive to free radical oxidation than somatic cells because their plasma membranes contain a greater quantity of polyunsaturated fatty acids leading to lipid peroxidation.²⁰ Oxidative stress increases sperm damage and results in increased sperm DNA fragmentation, which affects early embryo development and future child health, and increases the risk of miscarriage.^{21,22}

Thus, the mechanisms involved in the impact of obesity and metabolic disorders on male reproductive functions are numerous but strongly interlinked. Until recently, most of the studies have independently investigated the relationship between infertility and anthropometric, metabolic, and dietary elements. Therefore, identifying the parameter(s) involved in male or female infertility is challenging. Artificial intelligence, especially machine learning, has been adopted in scientific and medical research in the last few years, providing access to innovative and more powerful tools. We previously built and evaluated a machine learning (ML) model to identify and stratify fertile couples as well as those with idiopathic infertility. This score was based on a panel of lifestyle-influenced parameters, including the anthropometric, metabolic, and antioxidant status of both partners. This tool combines several easily measurable parameters and generates a score to assess the impact of anthropometric, metabolic, and antioxidant factors on the couple's risk of infertility. Although less efficient, we have also shown that this score, in its version built on male parameters alone, could identify almost 70% of men who were at risk of infertility.¹⁰ However, a score combining different elements influenced by lifestyle has never been used to identify the risk of impaired sperm parameters.

The objective of this study was to evaluate the relationship between semen parameters and anthropometric, metabolic, and nutritional parameters. First, we analyzed each parameter individually, and then, we applied a previously constructed and validated machine learning score to combine them.

MATERIALS AND METHODS

Subjects

"ALIMENTATION et FERTILITE" (ALIFERT) is a cross-sectional case-control study enrolling infertile men between September 2009 and December 2013. This study is registered and approved by National Biomedical Research (Paris, France; Approval No. P071224), Ethics Committee of 'Protection of Persons Committee' (Paris, France; Approval No. AOM 2009-A00256-51), NEudra CT (Paris, France; Approval No. 08180), and ClinicalTrials.gov (No. NCT01093378). The design of ALIFERT was multicenter. The study cohort included 75 infertile men recruited from four infertility centers in France (Jean-Verdier ART Center Hospital, Bondy [JV]; North Hospital ART Center,

Saint-Étienne [SE]; Navarre Polyclinic ART Center, Pau [PAU]; and Cochin Hospital ART Center, Paris [CCH]).

Men were between 18 years and 45 years old and exhibiting idiopathic infertility for over 12 months. Eligibility criteria have been described previously and include the provision of written informed consent.¹⁰

Eligibility criteria for patients were as follows: (1) primary idiopathic infertility >12 months; (2) age between 18 years and 45 years; (3) absence of severe oligozoospermia ($<5 \times 10^6$ ml⁻¹); (4) absence of male reproductive tract abnormalities such as undescended testis, varicocele, or infection; and (5) written informed consent. Patients with any current known or previous metabolic or digestive disease and smokers were excluded.

Assessments

Anthropometric, metabolic, antioxidant, and micronutrient parameters were assessed as described below.

Height, weight, visceral fat (Tanita BC-420MA analyzer; TANITA USA, Issaquah, WA, USA), and waist circumference measured at the narrowest point between the lower edge of the ribs and the iliac crest were measured by the investigator. The measurements were carried out by a trained investigator in the morning under fasting conditions.

Blood samples were obtained after 12 h of fasting for measurement of total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides, and glucose in fresh plasma. Both serum and plasma were frozen at -80°C until further analysis.

After 5 min of rest, systolic and diastolic blood pressures were assessed with a sphygmomanometer cuff around the patient's forearm while lying down. The values for systolic and diastolic pressures were calculated as the mean of the right and left measurements.

Finally, concentrations of serum vitamin D (ng ml⁻¹), vitamin B9 (folic acid-erythro; nmol l⁻¹), vitamin B9 (nmol l⁻¹), vitamin B12 (pmol l⁻¹), alpha-tocopherol (vitamin E; $\mu\text{mol l}^{-1}$), zinc (mmol l⁻¹), selenium (mmol l⁻¹), vitamin C (mg ml⁻¹), alpha-carotene ($\mu\text{mol l}^{-1}$), beta-carotene ($\mu\text{mol l}^{-1}$), lycopene ($\mu\text{mol l}^{-1}$), lutein ($\mu\text{mol l}^{-1}$), beta-cryptoxanthin ($\mu\text{mol l}^{-1}$), and retinol (vitamin A; $\mu\text{mol l}^{-1}$) levels were determined at the Department of Integrated Biology – Nutritional Biology and Oxidative Stress (Grenoble Hospital, Grenoble, France). Ascorbic acid (vitamin C) in serum was assessed by an automated continuous-flow method. The concentrations of retinol, tocopherol, and carotenoids (lutein, beta-cryptoxanthin, lycopene, alpha-carotene, and beta-carotene) in serum were obtained by high-performance liquid chromatographic (HPLC; Biotek-Kontron, Montigny-le-Bretonneux, France). Concentrations of serum zinc were measured by flame atomic absorption spectrometry (model 3110; PerkinElmer, Norwalk, CT, USA) and selenium by atomic absorption spectrometry (4100 ZL; PerkinElmer).

Semen quality parameters

The semen specimens were collected in the laboratory by masturbation into a sterile plastic cup after an abstinence period of 2 days to 5 days. The samples were liquefied at room temperature for 30 min, and conventional semen quality (semen volume, sperm concentration, vitality, and motility) was assessed according to the 2010 World Health Organization (WHO) guidelines.²³ David's criteria were used to evaluate sperm morphology.^{24,25}

Sperm DNA fragmentation

The terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) technique with an *In Situ* Cell Death Detection Kit (Fluorescein; Roche Applied Science, Meylan, France) was used to determine sperm DNA fragmentation.

Machine learning model

In previous work, our team developed and evaluated three ML models, capable of separating infertile/fertile couples, men and women through their bioclinical signature, to facilitate the management of couples with unexplained infertility.¹⁰ This signature was the result of a precise selection of biomarkers among hundreds of parameters investigated. It was based on the findings of a large-scale clinical study evaluating the impact of lifestyle and nutrition on couples' infertility. A major contribution of lifestyle parameters to the discriminatory power of the model had been demonstrated. The result generated and calculated by the model produces a score ranging from 0 (the highest risk of infertility with involvement of nutritional and metabolic factors) to 1 (the lowest risk, therefore rather healthy nutritional and metabolic profile). We calculated the model score for each man in the study cohort, using the formula obtained previously.¹⁰ A score between 0 and 1 was calculated for each man in the cohort, with the aim of assessing more globally the nutritional and metabolic status on fertility (**Figure 1**) and comparing its results with sperm parameters.

Statistical analyses

Python 3.9.12 (Python Software Foundation) was used to carry out the statistical analysis, and the SciPy package version 1.9.3 was used. Data were shown as mean and standard deviation (s.d.). Correlation analysis of metabolic and sperm parameters was investigated by the Pearson's parametric test. $P < 0.05$ was considered statistically significant.

RESULTS

Data were summarized as mean and s.d. and shown in **Table 1**. The association between anthropometric, metabolic, and dietary parameters and semen parameters is presented in **Table 2**.

In this population of idiopathic infertile men, we observed a correlation between some anthropometric characteristics and semen parameters. A decrease in sperm motility was observed with high BMI ($r = -0.233$, $P < 0.05$). In addition, high visceral fat was associated with a decrease in sperm vitality and motility ($r = -0.285$, $P < 0.01$ and $r = -0.270$, $P < 0.05$, respectively).

Regarding the correlation between metabolic characteristics and sperm parameters, high levels of total cholesterol and serum LDL were associated with lower sperm count ($r = -0.272$ and $r = -0.249$, both $P < 0.05$). In addition, lower sperm vitality was observed in patients with high total cholesterol and serum LDL ($r = -0.311$ and $r = -0.326$, both $P < 0.01$). Finally, sperm DNA fragmentation was negatively correlated with serum HDL concentration ($r = -0.364$, $P < 0.01$).

A high concentration of certain antioxidants in the serum is associated with better sperm parameters. Correlation between sperm count and plasma glutathione peroxidase (GPX) activity ($r = 0.230$, $P < 0.05$), as well as between sperm motility and vitamin B12 (cobalamin) concentration, was observed ($r = 0.275$, $P < 0.05$).

An unfavorable machine learning score, based on a combination of anthropometric, metabolic, and oxidative status, was associated with a high level of sperm DNA fragmentation ($r = -0.263$, $P < 0.05$).

DISCUSSION

The aim of this study was to explore the impact of factors such as anthropometric and metabolic status along with the nutritional environment on semen parameters in a population of patients with unexplained infertility. This study explored the relationship between a circulating metabolic and nutritional signature and sperm DNA fragmentation, a qualitative marker of spermatogenesis.

We observed a significant correlation of some anthropometric, metabolic, and nutritional disorders with some sperm characteristics.

Table 1: Baseline characteristics from clinical, biological, and sperm analysis of patients (n=75)

Parameter	Values, mean±s.d.
Baseline, metabolic, and biological parameters	
Age (year)	33.6±5.3
BMI (kg m ⁻²)	26.1±4.3
Waist measurement (cm)	92.6±11.3
Visceral fat (%)	7.0±4.2
Systolic blood pressure (mmHg) ^a	126.8±12.2
Diastolic blood pressure (mmHg) ^a	80.8±10.0
Glycemia (mmol l ⁻¹)	5.0±0.7
Total cholesterol (mmol l ⁻¹)	5.2±1.0
LDL (mmol l ⁻¹)	3.3±1.0
HDL (mmol l ⁻¹)	1.3±0.4
Triglyceride (mmol l ⁻¹)	1.4±0.9
Micronutrients and vitamins	
Vitamin D (ng ml ⁻¹)	23.2±11.3
Folate (nmol l ⁻¹)	12.7±7.2
Cobalamin (pmol l ⁻¹)	300.0±111.0
Retinol (μmol l ⁻¹)	2.1±0.5
Alpha-tocopherol (μmol l ⁻¹)	25.0±6.1
Zinc (μmol l ⁻¹)	12.8±1.9
Selenium (mmol l ⁻¹)	1.2±0.2
GPX (mUI ml ⁻¹)	391.6±61.0
Ascorbic acid (mg ml ⁻¹) ^a	42.3±18.8
Alpha-carotene (μmol l ⁻¹)	0.1±0.1
Lycopene (μmol l ⁻¹)	0.6±0.3
Lutein (μmol l ⁻¹)	0.3±0.2
Beta-cryptoxanthin (μmol l ⁻¹)	0.2±0.2
Glutathione (μmol l ⁻¹)	829.6±200.0
Beta-carotene (μmol l ⁻¹)	0.4±0.3
ML model score	
Semen volume (ml)	3.4±1.6
Sperm concentration (×10 ⁶ ml ⁻¹)	44.2±29.8
Total sperm count (×10 ⁶)	143.7±109.4
Progressive motility (a + b), %	37.5±11.6
Vitality (%)	63.4±15.4
Normal morphology (%)	20.7±9.9
DNA fragmentation rate (%)	28.6±18.4

^aFour patients with unknown systolic blood pressure (mmHg) and diastolic blood pressure (mmHg), and 6 patients with unknown ascorbic acid. s.d.: standard deviation; BMI: body mass index; LDL: low-density lipoprotein; HDL: high-density lipoprotein; ML: machine learning; GPX: glutathione peroxidase

These associations were not as pronounced as those described in other publications.^{2,4-6} This difference may be related to the fact that men in our study did not have major alterations in conventional semen parameters. Moreover, traditional sperm parameters are known to be highly variable between ejaculates, meaning intraindividual variability. DNA fragmentation, on the other hand, appears to be more stable across ejaculates. Traditional univariate analyses, particularly simple pair-wise correlations, may have limitations in this work, given this particular population, and may lack the power to capture more complex associations. The contribution and integration of new methods from artificial intelligence may provide an interesting way to better process available clinical and biological data.

The fertility score previously developed and published by our team was built on anthropometric, metabolic, and antioxidative factors, which were parameters close to those studied in this work. As shown in **Figure 1**, the result of the score was put into perspective with

Table 2: Correlations (Pearson's parametric test) between the studied parameters and sperm quality

Parameter	Semen volume (ml), r (P)	Sperm concentration ($\times 10^6$ ml ⁻¹), r (P)	Total sperm count ($\times 10^6$), r (P)	Progressive motility (a + b, %), r (P)	Vitality (%), r (P)	Normal morphology (%), r (P)	DNA fragmentation rate (%), r (P)
Baseline, metabolic, and biological parameters							
Age (year)	-0.142 (NS)	0.243 (NS)	0.199 (NS)	-0.066 (NS)	-0.119 (NS)	0.045 (NS)	0.107 (NS)
BMI (kg m ⁻²)	-0.307*	-0.080 (NS)	-0.136 (NS)	-0.233*	-0.217 (NS)	-0.068 (NS)	0.078 (NS)
Waist measurement (cm)	-0.286*	-0.102 (NS)	-0.157 (NS)	-0.191 (NS)	-0.169 (NS)	0.015 (NS)	0.152 (NS)
Visceral fat (%)	-0.221 (NS)	-0.060 (NS)	-0.062 (NS)	-0.270*	-0.285**	-0.067 (NS)	0.152 (NS)
Blood pressure systolic (mmHg)	-0.194 (NS)	0.129 (NS)	0.053 (NS)	0.087 (NS)	-0.045 (NS)	0.073 (NS)	-0.120 (NS)
Blood pressure diastolic (mmHg)	-0.23 (NS)	0.187 (NS)	0.078 (NS)	-0.011 (NS)	-0.107 (NS)	0 (NS)	-0.088 (NS)
Glycemia (mmol l ⁻¹)	0.059 (NS)	0.127 (NS)	0.132 (NS)	-0.147 (NS)	-0.166 (NS)	0.018 (NS)	-0.064 (NS)
Total cholesterol (mmol l ⁻¹)	-0.189 (NS)	-0.155 (NS)	-0.272*	-0.211 (NS)	-0.311**	-0.117 (NS)	-0.034 (NS)
LDL (mmol l ⁻¹)	-0.130 (NS)	-0.201 (NS)	-0.249*	-0.208 (NS)	-0.326**	-0.022 (NS)	0.053 (NS)
HDL (mmol l ⁻¹)	0.031 (NS)	-0.045 (NS)	-0.054 (NS)	-0.003 (NS)	0.011 (NS)	-0.186 (NS)	-0.364**
Triglycerides (mmol l ⁻¹)	-0.152 (NS)	0.088 (NS)	-0.040 (NS)	-0.016 (NS)	-0.041 (NS)	-0.066 (NS)	0.100 (NS)
Micronutrients and vitamins							
Vitamin D (ng ml ⁻¹)	0.208 (NS)	-0.028 (NS)	0.095 (NS)	-0.077 (NS)	0.002 (NS)	0.120 (NS)	0.149 (NS)
Folate (nmol l ⁻¹)	0.090 (NS)	-0.121 (NS)	0 (NS)	0.197 (NS)	0.027 (NS)	0.155 (NS)	0.161 (NS)
Cobalamin (pmol l ⁻¹)	-0.173 (NS)	0.057 (NS)	-0.080 (NS)	0.275*	-0.063 (NS)	0.106 (NS)	0.074 (NS)
Retinol (μ mol l ⁻¹)	-0.049 (NS)	-0.046 (NS)	-0.076 (NS)	0.025 (NS)	0.159 (NS)	-0.037 (NS)	0.126 (NS)
Alpha-tocopherol (μ mol l ⁻¹)	-0.042 (NS)	-0.194 (NS)	-0.226 (NS)	0.050 (NS)	-0.035 (NS)	-0.105 (NS)	-0.043 (NS)
Zinc (μ mol l ⁻¹)	-0.039 (NS)	-0.001 (NS)	0.008 (NS)	-0.122 (NS)	0.124 (NS)	0.043 (NS)	0.027 (NS)
Selenium (mmol l ⁻¹)	-0.031 (NS)	0.022 (NS)	0.015 (NS)	0.051 (NS)	-0.105 (NS)	0.030 (NS)	0.070 (NS)
GPX (mUI ml ⁻¹)	0.055 (NS)	0.204 (NS)	0.230*	0.017 (NS)	-0.182 (NS)	-0.053 (NS)	0.002 (NS)
Ascorbic acid (mg ml ⁻¹)	-0.114 (NS)	-0.077 (NS)	-0.145 (NS)	0.141 (NS)	0.067 (NS)	0.091 (NS)	0.129 (NS)
Alpha-carotene (μ mol l ⁻¹)	-0.082 (NS)	-0.003 (NS)	-0.050 (NS)	-0.042 (NS)	0.041 (NS)	-0.05 (NS)	-0.041 (NS)
Lycopene (μ mol l ⁻¹)	0.130 (NS)	-0.111 (NS)	-0.074 (NS)	-0.032 (NS)	-0.221 (NS)	-0.117 (NS)	-0.077 (NS)
Lutein (μ mol l ⁻¹)	-0.086 (NS)	-0.148 (NS)	-0.872 (NS)	0.149 (NS)	-0.050 (NS)	0.009 (NS)	0.073 (NS)
Beta-cryptoxanthin	-0.146 (NS)	-0.028 (NS)	-0.112 (NS)	0.069 (NS)	-0.018 (NS)	0.045 (NS)	-0.086 (NS)
Glutathione (μ mol l ⁻¹)	0.137 (NS)	-0.041 (NS)	0.067 (NS)	-0.003 (NS)	-0.078 (NS)	-0.035 (NS)	0.135 (NS)
Beta-carotene (μ mol l ⁻¹)	0.022 (NS)	-0.078 (NS)	-0.111 (NS)	-0.104 (NS)	-0.140 (NS)	0.044 (NS)	-0.059 (NS)
ML model score	0.091 (NS)	0.008 (NS)	-0.008 (NS)	0.132 (NS)	0.167 (NS)	-0.060 (NS)	-0.263*

*P<0.05; **P<0.01. NS: not significant; BMI: body mass index; LDL: low-density lipoprotein; HDL: high-density lipoprotein; ML: machine learning; GPX: glutathione peroxidase

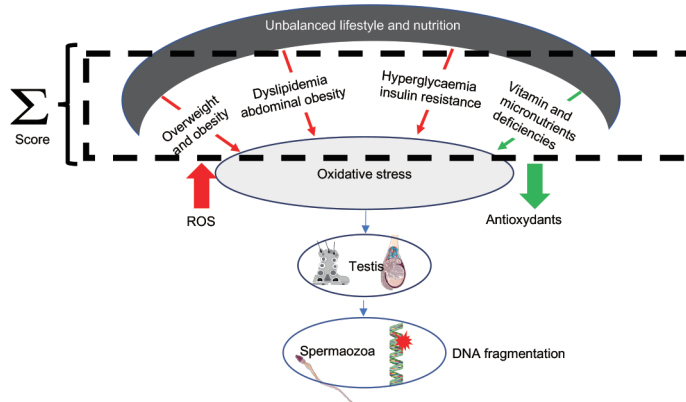


Figure 1: A complex relationship between lifestyle, diet, antioxidants, and sperm DNA fragmentation. An unbalanced lifestyle and diet lead to anthropometric, metabolic, and nutritional changes, which can lead to an imbalance in oxidative stress. These consequences are multiple and interrelated, and the use of artificial intelligence represents a unique opportunity to provide a more holistic evaluation of the influence of the considered marker. This oxidative stress could be one of the main drivers leading to a decline in male reproductive functions and consequently in sperm parameters, in particular sperm DNA fragmentation. ROS: reactive oxygen species.

sperm parameters. Interestingly, this score is the only feature, along with HDL level, that correlates significantly with DNA fragmentation level. Notwithstanding controversial findings in the literature, DNA fragmentation level is widely recognized as a marker of intrinsic

semen quality. Its levels seem to affect embryo quality, implantation, and miscarriage risk.²⁶ Many factors can increase the level of sperm DNA fragmentation, including obesity, high local temperature, drug treatment, exposure to environmental pollutants, and smoking.

Using the strength of multivariate statistical modeling, the score may reveal underlying phenomena that are difficult to identify using conventional univariate methods alone. By correlating significantly with DNA fragmentation, the score may provide a more comprehensive way of assessing male semen quality, particularly in patients with normal or subnormal semen. This assumption must be demonstrated by further studies. Furthermore, from an analytical and technical point of view, DNA fragmentation is a more challenging and less standardized biomarker than other semen analysis parameters.

Although only men with no severe sperm alteration were recruited, a negative correlation between BMI, visceral fat, and sperm motility was observed. Obesity in males has been found to be a risk factor for infertility and increased time to conception.³ In men, many studies, supported by a meta-analysis,²⁷ have reported an alteration of sperm parameters in case of overweight and obesity, more specifically a decrease in sperm concentration and sperm count.^{28–30} Sperm motility also appears to be affected by overweight or obesity.³¹

Dyslipidemia has been reported to negatively affect testicular and epididymal function, sperm maturation and quality, and ejaculatory function.³² Consistent with these findings, in this population of patients with unexplained infertility, we observed an association between lipid levels and some semen parameters. Indeed, despite the absence of patients with dyslipidemia, unfavorable lipid composition (high total and LDL cholesterol and/or low HDL cholesterol) may be negatively associated with semen parameters. However, while blood glucose was a strongly discriminating factor between fertile and infertile couples (as studied previously¹⁰), it does not appear to be correlated with sperm parameters in this population of infertile men. The molecular mechanisms involved in this phenomenon are unclear, but hypotheses have been proposed. Impairment of steroidogenesis has been reported. Cholesterol may modulate the renin–angiotensin system in the testes, leading to the inhibition of steroidogenesis. This, in turn, leads to a decrease in testosterone production.³³ Posttesticular sperm maturation may also be affected. Hypercholesterolemia would induce changes in the epididymal epithelium like an accumulation of cholesterol ester lipid droplets in the smooth muscle of the epididymal tract. Epididymal peristalsis contractions would be impaired, compromising sperm progression inside the epididymal lumen, and thus, the maturation process.³⁴ Eventually, an increase of oxidative stress and an excess of free radical production could also be the consequence of dyslipidemia.³⁵ This last hypothesis is consistent with our findings. Indeed, in our cohort, the HDL level (the “good” cholesterol level) was correlated with the DNA fragmentation level, directly related to oxidative stress.

Besides overweight and lipid and carbohydrate profile, some microelements and vitamins were positively correlated with sperm count and motility. Vitamin B12 levels and GPX activity were correlated with sperm motility and total sperm count, respectively, in this cohort. A meta-analysis by Banihani³⁶ highlighted the importance of vitamin B12 in sperm physiology and quality, particularly on total sperm count.³⁷ Indeed, lower vitamin B12 levels have been reported in infertile men.³⁸ Several beneficial mechanisms of vitamin B12 have been described, including its link to homocysteine metabolism and its antioxidant properties.³⁹

Numerous studies have demonstrated the importance of the GPX during spermatogenesis and its relationship with male fertility.^{40,41} Similarly, previous studies have reported higher GPX activity in seminal plasma in patients with better sperm quality.^{42,43} GPX plays a key function as reactive oxygen species (ROS) scavengers for spermatozoa, to maintain the balance between ROS production and recycling.⁴⁴

A limitation of this study was the lack of measurement of antioxidant status in seminal plasma for the entire cohort of patients.

In addition, in accordance with international recommendations,^{45–47} exploratory andrological assessments (hormone tests, ultrasound, *etc.*) were not carried out in our patients.

This study focused on a particular population of men from couples with unexplained infertility, and a limited population size of 75 patients. However, measurement of sperm parameters alone, which were normal or subnormal in this population, does not appear to be sufficient to evaluate and stratify these patients. These findings suggest that despite normal or subnormal semen parameters, the metabolic and oxidative status factors considered in this study, which are probably in part linked to the lifestyle and behavior of these patients, should be integrated into a holistic and comprehensive diagnostic and therapeutic strategy. Unlike other factors or causes of infertility (genetic or hormonal abnormalities, chemotherapy, *etc.*), in this case, patients can be involved in the management of their health, especially given the reversible character of these indicators. Indeed, an improvement in metabolic and oxidative status could be associated with a reduced risk of infertility, and even a better chance of success when *in vitro* fertilization (IVF) treatment is still required.

Improving the lifestyle of these patients therefore appears to be essential, and even beyond reproductive health alone, for overall physical and mental well-being. Such improvement should be advised by health-care professionals, using dietary and physical activity coaching or guidance properly designed and validated by international guidelines. Interventional studies are needed to confirm the positive impact of lifestyle improvement.⁴⁸

CONCLUSION

Favorable anthropometric, metabolic, and oxidative patterns, which may reflect an appropriate lifestyle, appear to positively impact overall health, in particular reproductive function. Despite normal or subnormal semen parameters, findings reported in this work using patients from a cohort of idiopathic infertile couples show that altered body composition, as well as unfavorable lipid balance and vitamin and microelement deficiency, was associated with worse semen parameters.

This study, consistent with previous publications in this area, suggests that beyond semen quality parameters, in an essential assessment of male fertility, other key factors should be taken into account. A more comprehensive evaluation of the patients' health, affecting their nutritional, metabolic, and oxidative status, seems to be an indispensable element to consider for more holistic care. In this regard, the application of emerging artificial intelligence techniques may provide a unique opportunity to integrate all these parameters and deliver personalized care.

AUTHOR CONTRIBUTIONS

GB participated in the study's conception and design, performed statistical analysis, and drafted the manuscript. RL supervised the study, participated in patient recruitment, interpreted the data, and critically revised the manuscript for intellectual content. CF was involved in study conception and design, patient recruitment, interpretation of data, and critical revision of the manuscript for intellectual content. SC and AL both worked on the study conception and design and critically revised the manuscript for intellectual content. CD contributed to the study conception and design, patient recruitment, interpretation of data, and drafting of the manuscript. The ALIFERT collaborative group members contributed to the study design and were involved in patient recruitment or analysis. All authors read and approved the final manuscript.



COMPETING INTERESTS

All authors declare no competing interests.

ACKNOWLEDGMENTS

The authors acknowledge all the couples involved in the study and the ALIFERT Collaborative Group.

The ALIFERT Collaborative Group is composed of: Isabelle Aknin¹, Isabelle Cedrin-Durnerin², Steven Cens³, Pascale Chavatte-Palmer⁴, Serge Hercberg⁵, Khaled Pocate⁶, Nathalie Sermondade⁷, Claude Uthuriague³, and Jean-Philippe Wolf⁶

¹Reproductive Biology Functional Unit, Histology-Embryology-Cytogenetics, North Hospital, Saint-Étienne 42000, France; ²Reproductive Medicine Department, Jean Verdier Hospital, AP-HP, Bondy 93140, France; ³Assisted Reproductive Technology Centre of PAU, Navarre Polyclinic, Pau 64445, France; ⁴INRA, UMR1198 Biology of Development and Reproduction, Jouy en Josas 78350, France; ⁵University of Paris 13, CRNH IdF, Bobigny 93017, France; ⁶Histology-Embryology-Reproductive Biology Department, Cochin Hospital AP-HP, Paris 75014, France; ⁷Reproductive Biology Department-CECOS, Tenon Hospital, AP-HP, Paris 75020, France.

REFERENCES

- Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, *et al*. Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. *Hum Reprod Update* 2023; 29: 157–76.
- Sermondade N, Faure C, Fezeu L, Shayeg AG, Bonde JP, *et al*. BMI in relation to sperm count: an updated systematic review and collaborative meta-analysis. *Hum Reprod Update* 2013; 19: 221–31.
- Ramlau-Hansen CH, Thulstrup AM, Nohr EA, Bonde JP, Sørensen TI, *et al*. Subfecundity in overweight and obese couples. *Hum Reprod* 2007; 22: 1634–7.
- Esmaili V, Zendehelel M, Shahverdi A, Alizadeh A. Relationship between obesity-related markers, biochemical metabolic parameters, hormonal profiles and sperm parameters among men attending an infertility clinic. *Andrologia* 2022; 54: e14524.
- Dupont C, Faure C, Sermondade N, Boubaya M, Eustache F, *et al*. Obesity leads to higher risk of sperm DNA damage in infertile patients. *Asian J Androl* 2013; 15: 622–5.
- Leisegang K, Sengupta P, Agarwal A, Henkel R. Obesity and male infertility: mechanisms and management. *Andrologia* 2021; 53: e13617.
- Sun B, Messerlian C, Sun ZH, Duan P, Chen HG, *et al*. Physical activity and sedentary time in relation to semen quality in healthy men screened as potential sperm donors. *Hum Reprod* 2019; 34: 2330–9.
- Dupont C, Faure C, Daoud F, Gautier B, Czernichow S, *et al*. Metabolic syndrome and smoking are independent risk factors of male idiopathic infertility. *Basic Clin Androl* 2019; 29: 9.
- Foucaut AM, Faure C, Julia C, Czernichow S, Levy R, *et al*. Sedentary behavior, physical inactivity and body composition in relation to idiopathic infertility among men and women. *PLoS One* 2019; 14: e0210770.
- Bachelot G, Lévy R, Bachelot A, Faure C, Czernichow S, *et al*. Proof of concept and development of a couple-based machine learning model to stratify infertile patients with idiopathic infertility. *Sci Rep* 2021; 11: 24003.
- Hammiche F, Laven JS, Twigg JM, Boellaard WP, Steegers EA, *et al*. Body mass index and central adiposity are associated with sperm quality in men of subfertile couples. *Hum Reprod* 2012; 27: 2365–72.
- Carrageta DF, Oliveira PF, Alves MG, Monteiro MP. Obesity and male hypogonadism: tales of a vicious cycle. *Obes Rev* 2019; 20: 1148–58.
- Hautanen A. Synthesis and regulation of sex hormone-binding globulin in obesity. *Int J Obes* 2000; 24 Suppl 2: S64–70.
- Monteiro R, Azevedo I. Chronic inflammation in obesity and the metabolic syndrome. *Mediators Inflamm* 2010; 2010: 289645.
- Fève B, Bastard JP, Vidal H. [Relationship between obesity, inflammation and insulin resistance: new concepts]. *C R Biol* 2006; 329: 587–97. [Article in French]
- Furukawa S, Fujita T, Shimabukuro M, Iwaki M, Yamada Y, *et al*. Increased oxidative stress in obesity and its impact on metabolic syndrome. *J Clin Invest* 2004; 114: 1752–61.
- Vincent HK, Innes KE, Vincent KR. Oxidative stress and potential interventions to reduce oxidative stress in overweight and obesity. *Diabetes Obes Metab* 2007; 9: 813–39.
- Baccetti B. Insulin-dependent diabetes in men is associated with hypothalamopituitary derangement and with impairment in semen quality. *Hum Reprod* 2002; 17: 2673–7.
- Yu T, Botham JL, Yoon Y. Increased production of reactive oxygen species in hyperglycemic conditions requires dynamic change of mitochondrial morphology. *Proc Natl Acad Sci U S A* 2006; 103: 2653–8.

- Shi GJ, Li Y, Cao QH, Wu HX, Tang XY, *et al*. *In vitro* and *in vivo* evidence that quercetin protects against diabetes and its complications: a systematic review of the literature. *Biomed Pharmacother* 2019; 109: 1085–99.
- Esteves SC. Interventions to prevent sperm DNA damage effects on reproduction. *Adv Exp Med Biol* 2019; 1166: 119–48.
- McQueen DB, Zhang J, Robins JC. Sperm DNA fragmentation and recurrent pregnancy loss: a systematic review and meta-analysis. *Fertil Steril* 2019; 112: 54–60.e3.
- World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Semen. 5th ed. Geneva: World Health Organization; 2010.
- Auger J, Eustache F, Andersen AG, Irvine DS, Jørgensen N, *et al*. Sperm morphological defects related to environment, lifestyle and medical history of 1001 male partners of pregnant women from four European cities. *Hum Reprod* 2001; 16: 2710–7.
- Jouannet P, Ducot B, Feneux D, Spira A. Male factors and the likelihood of pregnancy in infertile couples. I. Study of sperm characteristics. *Int J Androl* 1988; 11: 379–4.
- Haddock L, Gordon S, Lewis SE, Larsen P, Shehata A, *et al*. Sperm DNA fragmentation is a novel biomarker for early pregnancy loss. *Reprod Biomed Online* 2021; 42: 175–84.
- Sermondade N. Obesity and increased risk for oligozoospermia and azoospermia. *Arch Intern Med* 2012; 172: 440.
- Kort HI. Impact of body mass index values on sperm quantity and quality. *J Androl* 2006; 27: 450–2.
- Chavarro JE, Toth TL, Wright DL, Meeker JD, Hauser R. Body mass index in relation to semen quality, sperm DNA integrity, and serum reproductive hormone levels among men attending an infertility clinic. *Fertil Steril* 2010; 93: 2222–31.
- Belloc S, Cohen-Bacrie M, Amar E, Izard V, Benkhalifa M, *et al*. High body mass index has a deleterious effect on semen parameters except morphology: results from a large cohort study. *Fertil Steril* 2014; 102: 1268–73.
- Sekhvat L, Moein MR. The effect of male body mass index on sperm parameters. *Aging Male* 2010; 13: 155–8.
- Pushendra A, Gc J. Hyper-lipidemia and male fertility: a critical review of literature. *Andrology* 2015; 4: 141.
- Martínez-Martos JM, Arrazola M, Mayas MD, Carrera-González MP, García MJ, *et al*. Diet-induced hypercholesterolemia impaired testicular steroidogenesis in mice through the renin-angiotensin system. *Gen Comp Endocrinol* 2011; 173: 15–9.
- Ouvrier A, Alves G, Damon-Soubeyrand C, Marceau G, Cadet R, *et al*. Dietary cholesterol-induced post-testicular infertility. *PLoS One* 2011; 6: e26966.
- Selley ML, Lacey MJ, Bartlett MR, Copeland CM, Ardlie NG. Content of significant amounts of a cytotoxic end-product of lipid peroxidation in human semen. *Reproduction* 1991; 92: 291–8.
- Banihani SA. Vitamin B12 and semen quality. *Biomolecules* 2017; 7: 42.
- Boxmeer JC, Smit M, Weber RF, Lindemans J, Romijn JC, *et al*. Seminal plasma cobalamin significantly correlates with sperm concentration in men undergoing IVF or ICSI procedures. *J Androl* 2007; 28: 521–7.
- Dhillon VS, Shahid M, Husain SA. Associations of *MTHFR DNMT3b* 4977 bp deletion in mtDNA and *GSTM1* deletion, and aberrant CpG island hypermethylation of *GSTM1* in non-obstructive infertility in Indian men. *MHR Basic Sci Reprod Med* 2007; 13: 213–22.
- Hu JH, Tian WQ, Zhao XL, Zan LS, Xin YP, *et al*. The cryoprotective effects of Vitamin B12 supplementation on bovine semen quality: effects of Vitamin B12 on bovine semen quality. *Reprod Domest Anim* 2011; 46: 66–73.
- Imai H, Suzuki K, Ishizaka K, Ichinose S, Oshima H, *et al*. Failure of the expression of phospholipid hydroperoxide glutathione peroxidase in the spermatozoa of human infertile males. *Biol Reprod* 2001; 64: 674–83.
- Kralikova M, Crha I, Huser M, Melounova J, Zakova J, *et al*. The intracellular concentration of homocysteine and related thiols is negatively correlated to sperm quality after highly effective method of sperm lysis. *Andrologia* 2017; 49: e12702.
- Giannattasio A, De Rosa M, Smeraglia R, Zarrilli S, Cimmino A, *et al*. Glutathione peroxidase (GPX) activity in seminal plasma of healthy and infertile males. *J Endocrinol Invest* 2002; 25: 983–6.
- Crisol L, Matorras R, Aspichueta F, Expósito A, Hernández ML, *et al*. Glutathione peroxidase activity in seminal plasma and its relationship to classical sperm parameters and *in vitro* fertilization-intracytoplasmic sperm injection outcome. *Fertil Steril* 2012; 97: 852–7.e1.
- Drevet JR. The antioxidant glutathione peroxidase family and spermatozoa: a complex story. *Mol Cell Endocrinol* 2006; 250: 70–9.
- Minhas S, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, *et al*. European association of urology guidelines on male sexual and reproductive health: 2021 Update on Male Infertility. *Eur Urol* 2021; 80: 603–20.
- Schlegel PN, Sigman M, Collura B, De Jonge CJ, Eisenberg ML, *et al*. Diagnosis and treatment of infertility in men: AUA/ASRM Guideline Part I. *J Urol* 2021; 205: 36–43.

- 47 Schlegel PN, Sigman M, Collura B, De Jonge CJ, Eisenberg ML, *et al*. Diagnosis and treatment of infertility in men: AUA/ASRM guideline part II. *Fertil Steril* 2021; 115: 62–9.
- 48 Dupont C, Aegerter P, Foucaut AM, Reyre A, Lhuissier FJ, *et al*. Effectiveness of a therapeutic multiple-lifestyle intervention taking into account the periconceptional environment in the management of infertile couples: study design of a randomized controlled trial – The PEPCI study. *BMC Pregnancy Childbirth* 2020; 20: 322.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

©The Author(s)(2024)

Résultats II : Combinaison de métabolomique et d'intelligence artificielle pour simplifier le diagnostic différentiel SOPK vs hyperplasie congénitale des surrénales

Contexte scientifique et objectifs

Comme décrit dans l'introduction, les mécanismes physiopathologiques responsables du SOPK et de l'HCS sont très différents, mais ces deux maladies sont à l'origine de patientes ayant des présentations cliniques similaires voire identiques. L'exclusion du diagnostic d'HCS apparaît donc comme une étape essentielle dans l'exploration de l'hyperandrogénie, mais aussi dans l'établissement formel du diagnostic de SOPK.

Importance de distinguer correctement SOPK vs NC21OHD : deux maladies de physiopathologie différentes, mais à présentation clinique identique

L'exclusion du NC21OHD dans un contexte d'hyperandrogénie ou en cas de suspicion de SOPK revêt donc d'une double importance : thérapeutique et génétique.

Thérapeutique : En effet, la prise en charge thérapeutique en cas de désir de grossesse chez une patiente atteinte d'une HCS est le traitement par hydrocortisone, qui aura également pour effet une baisse de la progestérone (effet sur l'implantation) et une réduction du surrisque de fausse couche précoce spontanée ⁷⁰. Ce traitement par glucocorticoïdes, par un effet de rétrocontrôle négatif sur l'ACTH et la CRH, réduit la sécrétion d'androgènes surrénaliens et leurs précurseurs, et permet généralement de normaliser les taux d'androgènes dans 90% des cas entraînant une amélioration voire une normalisation des cycles menstruels ⁷¹. La prise en charge du SOPK est quant à elle beaucoup plus complexe, avec en outre l'utilisation de traitement anti-androgéniques pour lutter contre l'hyperandrogénie, d'inducteur de l'ovulation ou encore la FIV en cas d'infertilité. La prise en charge du risque métabolique associé au SOPK et à l'insulinorésistance permet de limiter l'effet amplificateur de l'hyperinsulinisme sur la sécrétion des androgènes ovariens (et bien entendu limiter leurs conséquences).

Génétique : Les déficits en 21-hydroxylases sont dus à des mutations du gène CYP21A2 codant pour le cytochrome P450c21, situé sur le bras court du chromosome 6. Les mutations les plus fréquentes responsables des formes non classiques sont la V281L (55%), P30L (4%) et P453S (4%)³⁵. Il existe une bonne corrélation génotype/phénotype : certaines mutations sont responsables de formes classiques, d'autres de formes non classiques. Dans le cadre des formes non classiques, on retrouve une activité enzymatique résiduelle qui varie entre 20 et 50%, alors que dans les formes classiques elle est inférieure à 20%. L'HCS est une maladie autosomique récessive, deux allèles doivent donc être mutés pour observer l'expression de la maladie. Les patientes atteintes d'HCS peuvent être homozygotes ou le plus souvent hétérozygotes composites, et l'expression de la maladie est déterminée par la mutation la moins sévère. La forme non classique est ainsi due, soit à la présence de deux mutations entraînant une forme non classique, soit à la présence d'une mutation responsable d'une forme classique et une autre d'une forme non classique (jusqu'à 60% des cas !)⁷². Ces mutations sévères peuvent être transmises à la descendance avec le risque que, si le conjoint est lui aussi porteur d'une mutation sévère du gène CYP21A2 à l'état hétérozygote, l'enfant puisse être atteint par une forme classique de la maladie (= grave). En tout état de cause, le dépistage systématique du conjoint est donc nécessaire pour toute patiente atteinte de NC21OHD, dans la mesure où la fréquence des hétérozygotes pour une mutation sévère dans la population générale est estimée à 1/60⁷².

Objectif

La problématique réside donc dans la stratégie actuelle pour distinguer ces deux entités, laquelle comporte plusieurs inconvénients et limites. Ainsi, la seule mesure de la 17OHP semble insuffisante et la détermination d'un seuil de 17OHP sur taux de base reste débattue. Cette situation impose de stimuler la glande surrénale pour exacerber le bloc enzymatique et ainsi caractériser une augmentation nette de la 17OHP (> 10 ng/mL pour le seuil post-stimulation). Comme décrit dans l'introduction, la spectrométrie de masse constitue un véritable tournant technologique et analytique, en permettant d'offrir une quantification absolue et surtout simultanée des profils moléculaires de plusieurs stéroïdes. L'une des conséquences des approches multiplexes par MS est le profond changement de paradigme qu'elles engendrent: les stéroïdes ne sont plus considérés de manière indépendante les uns des autres (comme c'est le cas avec les approches d'immunoanalyse), mais selon des empreintes métaboliques. Cette nouvelle approche a nécessité le développement de modèles mathématiques multivariés

capables d'appréhender des systèmes biologiques complexes en condensant l'information par le biais de projections et/ou de combinaisons de variables moléculaires.

Notre hypothèse est donc qu'en appliquant des méthodes d'analyse de données à cette cartographie métabolique (voie de biosynthèse des stéroïdes) accessibles grâce aux techniques analytiques telles que la LS-MS/MS, nous pourrions simplifier le diagnostic différentiel entre SOPK et bloc de 21-hydroxylase non classique, et ainsi nous affranchir du test dynamique à l'ACTH. Pour ce faire, nous avons commencé par mener une première preuve de concept sur une population pédiatrique, atteinte de formes plus franches (annexe) ⁷³. Selon le degré d'atteinte, l'HCS-NC peut se révéler plus tôt dans l'enfance. Les enfants atteints d'HCS-NC présentent une pilosité pubienne avant 8 ans chez la fille et 9 ans chez le garçon, une accélération de la vitesse de croissance et de la maturation osseuse, secondaire à un bloc enzymatique plus sévère. Cependant, la survenue d'une pilosité pubienne précoce est, dans la majorité des cas, une variante de la normale ou prémature adrénarche (PP). La prémature adrénarche est un diagnostic d'élimination chez un enfant qui ne présente ni signe d'hyperandrogénie, ni accélération de la vitesse de croissance, ni avance de maturation osseuse. Cela correspond à une activation précoce mais non pathologique de la glande surrénale. Une épreuve dynamique à l'ACTH est nécessaire en pratique pour distinguer ces deux entités. L'objectif de cette étude était de développer un outil capable de court-circuiter cette épreuve dynamique.

Nous avons montré qu'en appliquant une modélisation multivariée à une empreinte métabolique de 16 stéroïdes nous pouvions distinguer des enfants avec une puberté prématurée d'enfants atteints d'HCS-NC sans épreuve dynamique et de manière fiable. Ce travail a porté sur une cohorte de 97 patients. Sur le plan méthodologique, cette étude a été menée en monocentrique (hôpital Trousseau, AP-HP), et nous avons utilisé une validation temporelle : une cohorte d'entraînement, composée de patients recrutés sur une première période et une cohorte de test, avec des patients recrutés sur une seconde période. Chaque empreinte métabolique (16 stéroïdes) a été dosée par LC-MS/MS, comme décrit dans la méthode publiée par l'équipe du département de métabolisme (28). Du point de vue de la modélisation mathématique, l'algorithme de l'analyse discriminante par la méthode des moindres carrés partiels ou « PLS-DA » a été appliquée. Le modèle OPLS-DA ainsi évalué sur la cohorte de test présentait une performance de 100% (sensibilité et spécificité).

L'objectif du travail présenté dans cette partie était d'étendre cette méthodologie (empreinte métabolique + modélisation) à une population adulte plus hétérogène, avec des formes moins marquées, dans le but d'identifier des blocages de la 21-hydroxylase au sein d'une population hyperandrogénique, dont certaines étaient atteintes de SOPK. Une première cohorte, utilisée pour l'entraînement de l'algorithme, était composée de 19 patientes HCS-NC et 19 contrôles provenant du service d'endocrinologie de la reproduction de l'hôpital Bicêtre (Pr Jacques Young). Une seconde cohorte externe de validation provenant du département d'endocrinologie et médecine de la reproduction de l'hôpital la Pitié-Salpêtrière (Pr Anne Bachelot) a été utilisée pour mesurer les performances du modèle ainsi construit. Elle se composait de 17 HCS-NC, de 72 contrôles et d'une population confondante de 266 patientes SOPK. Les 283 patients (266 + 17) ont subi un test dynamique de l'ACTH, et les 17 HCS-NC ont été confirmés par la génétique. Ceci constituerait une première étape dans la simplification de l'arbre décisionnel pour l'exploration de l'hyperandrogénie, permettant d'éviter la réalisation d'un test dynamique de l'ACTH (gain de temps, confort du patient, etc.).

Résumé des résultats

Construction et évaluation du modèle

Fort de cette première preuve de concept dans une population pédiatrique, nous avons voulu étendre cette méthode à une population adulte plus hétérogène, qui consultait pour des troubles du cycle ou de l'infertilité, toujours dans le but de se passer d'un test dynamique à l'ACTH. Conformément aux résultats publiés précédemment ⁷⁴, l'utilisation de la spectrométrie de masse pour déterminer la concentration en 17OHP n'était pas suffisante dans notre cohorte de réplique pour exclure correctement les HCS-NC du groupe contrôle/SOPK. En se basant sur la seule mesure de la [17OHP], nous avons comptabilisé 25 faux positifs et 1 faux négatif. En revanche, l'utilisation d'un panel de 15 stéroïdes combiné à une modélisation multivariée a permis d'exclure sans erreur les 17 patients NC21OHD (100%) de la cohorte test, malgré la présence d'une population confondante de 266 SOPK. Il convient toutefois de noter que les patientes atteintes de SOPK se situent toutes dans la zone " contrôle " : le modèle actuel n'est donc pas en mesure de distinguer les patientes atteintes de SOPK des témoins (il n'a d'ailleurs pas été conçu et formé à cette fin).

Stéroïdes importants dans le modèle

Les stéroïdes les plus contributifs dans la construction du modèle étaient les suivants : 21DF, 21DB, 11 β -OHA, 17OHP, A, B et T. Ces résultats, indissociables de la seule exactitude de classification, permettent de corroborer les hypothèses et connaissances des altérations biochimiques impliquées dans ces maladies, et renforcent la pertinence du modèle ainsi construit. En cas de HCS-NC par bloc en 21-hydroxylase (CYP21A2), une accumulation de Progestérone et de 17OHP en amont du bloc, qui sont tous deux également substrat de la 11 β -hydroxylase (CYP11B1), conduit à une augmentation des concentrations sériques en 21DB (21-deoxycorticoesterone ou 11 β -Hydroxyprogesterone) et 21DF (21-Deoxycortisol), respectivement. L'activation renforcée de la 11 β -hydroxylase a donc également pour conséquence une augmentation de 11 β -OH- Δ 4-androstènedione (qui se retrouve être le 3ème stéroïde le plus discriminant dans notre modèle). Ces trois stéroïdes, 21DF, 21DB et 11BOHA) apparaissent comme étant les plus importants dans le modèle OPLS-DA, en plus de la 17OHP.

Les résultats de cette étude sont encourageants et démontrent qu'un test dynamique à l'ACTH pourrait être évité pour exclure le diagnostic d'HCS-NC dans un contexte d'hyperandrogénie, malgré la présence d'une population confondante hétérogène de 266 patientes atteintes de SOPK. Par ailleurs, plusieurs patientes (n=31) de la cohorte test avaient été prélevées en phase lutéale (considérée comme telle lorsque [P] > 10 ng/mL) : le moment du prélèvement n'a pas eu d'impact sur le résultat du modèle. Ces premiers résultats, qui doivent être confirmés dans une étude plus large, laissent penser que cette contrainte de prélèvement sanguin pourrait être levée.

Éléments de discussion

Ce travail constituait une première étape dans la simplification de l'exploration des hyperandrogénies. Une validation prospective à plus large échelle est maintenant nécessaire pour confirmer ces résultats.

Modèle

Le degré élevé de multicolinéarité dû aux différentes espèces moléculaires composant la même voie de biosynthèse (dans notre cas, la biosynthèse des stéroïdes) a une fois de plus orienté notre choix vers l'utilisation de l'algorithme OPLS-DA. Nous avons choisi d'inclure uniquement

notre choix vers l'utilisation de l'algorithme OPLS-DA. Nous avons choisi d'inclure uniquement des HCS-NC et des contrôles dans la cohorte d'entraînement, afin que le modèle puisse apprendre de manière plus précise les spécificités et les caractéristiques de l'empreinte métabolique du bloc en 21-hydroxylase. Compte tenu de la diversité de profil des patientes atteintes de SOPK, nous ne les avons intégrés que dans la cohorte de test. Dans cette cohorte test (n=355), l'algorithme a été en mesure d'identifier un bloc en 21-hydroxylase sans confondre avec une empreinte de SOPK, malgré des cliniques identiques.

Stéroïdes les plus discriminants

En ce qui concerne les stéroïdes les plus pertinents pour la discrimination, deux stéroïdes d'origine strictement surrénalienne arrivent en tête du classement : le 21-désoxycortisol (21DF) et le 21-désoxycorticostérone (21DB). Ces deux stéroïdes, 21DF et 21DB, sont les produits de l'action directe de l'enzyme surrénalienne cytochrome P450 11 β -hydroxylase (CYP11B1) sur la 17OHP et la progestérone respectivement, qui s'accumulent en cas de bloc en 21-hydroxylase. Cette enzyme est abondamment exprimée dans les zones fasciculées et réticulées de la glande surrénale, mais son expression dans les gonades est négligeable. Le troisième stéroïde le plus discriminant, la 11 β -OHA, qui est le deuxième stéroïde le plus abondant produit par les glandes surrénales, est également le produit de l'action de la CYP11B1 sur l'androstènedione. Ensuite, dans les tissus périphériques, en particulier au niveau rénal, les quantités abondantes de 11 β -OHA produites par les glandes surrénales peuvent être oxydées en 11KA4 par l'enzyme 11 β -hydroxystéroïde déshydrogénase de type 2 (11 β HSD2), puis convertie en 11KT par l'AKR1C3 (Aldo-keto reductase family 1 member C3) dans d'autres tissus périphériques dont le tissu adipeux. Curieusement, la 17OHP, le marqueur actuellement utilisé pour le diagnostic, n'arrive qu'en quatrième position. L'une des explications les plus probables est sa double origine, ovarienne et surrénalienne, qui le rend moins spécifique que d'autres marqueurs tels que le 21DF ⁷⁵.

Forces et faiblesses

Les cohortes utilisées dans ce travail étaient la principale force de cette étude. En plus d'un phénotypage rigoureux, nous disposions des résultats génétiques pour l'étude des mutations du gène CYP21A2 pour toutes les HCS-NC. Toutes les patientes atteintes de SOPK ont effectué un test dynamique à l'ACTH, dont les valeurs, toutes inférieures à 7ng/mL, ont permis de réduire au minimum le risque de bloc en 21-hydroxylase dans cette population

confondante. La nature multicentrique de l'étude a également permis une validation du modèle dans les règles de l'art. De plus, le choix des variables d'entrée était pertinent : l'objectif du modèle était d'identifier une maladie métabolique affectant une enzyme dans la voie de biosynthèse des stéroïdes. La modélisation précise des stéroïdes de cette voie de biosynthèse, directement liés à la physiopathologie de la maladie, garantissant une proximité étroite entre la physiopathologie et les données modélisées, a entraîné une performance remarquable du modèle.

Une des limites de cette étude publiée est l'absence de patientes présentant d'autres pathologies responsables d'hyperandrogénie, tel que le syndrome de Cushing ou les tumeurs surrenaliennes et ovariennes. Bien que représentant un défi clinique moins important que celui du diagnostic différentiel HCS-NC/SOPK, de telles patientes ont été par la suite testées dans le modèle construit. De même, les contrôles et les PCOS étaient positionnés dans la même zone du modèle. Cette étape fera l'objet de développements ultérieurs. Il convient également de noter que les analyses LC-MS/MS ont toutes été effectuées dans le même laboratoire (métabolomique clinique à Saint-Antoine). Bien que des contrôles de qualité externes réguliers soient effectués, censés garantir la cohérence entre les résultats obtenus dans différents laboratoires en France, il aurait été intéressant de tester des profils obtenus dans un autre laboratoire dans le modèle. Nous pouvons également supposer que la cartographie de la voie de biosynthèse des stéroïdes, plutôt qu'un seul dosage, nettement plus informatif, réduirait, voire amortirait, cette variabilité analytique potentielle.


Conclusion et perspectives

En définitive, les résultats obtenus et présentés dans ce travail sont encourageants et montrent que la combinaison d'une empreinte métabolique et d'outils issus de l'intelligence artificielle permet de simplifier le diagnostic différentiel NC21OHD/SOPK, indépendamment de la phase du cycle menstruel. La cartographie de la voie de biosynthèse des stéroïdes est beaucoup plus informative que la seule concentration en 17OHP, et permet d'éviter l'épreuve dynamique à l'ACTH. Ce premier modèle ainsi développé constitue la première étape d'un projet plus ambitieux d'exploration des hyperandrogénies et troubles du cycle chez la femme en âge de procréer. Ce modèle pourrait améliorer la prise en charge de ces patientes, en simplifiant le diagnostic, améliorant le confort tout en faisant gagner du temps aux praticiens. Un travail reste à faire auprès des acteurs et des utilisateurs pour intégrer ces modèles dans

l'intelligence artificielle pour renforcer l'arsenal diagnostique à travers le développement d'outils plus efficaces, plus performants et plus confortables.

De ce travail, on peut noter deux grandes perspectives : 1) confirmer et étendre la capacité du modèle à identifier les NC21OHD depuis une population plus large, plus hétérogène et représentative de la pratique courante et en levant d'autres contraintes du prélèvement et 2) s'intéresser de plus près au groupe SOPK, lui-même hétérogène, en intégrant des marqueurs additionnels explorant d'autres voies métaboliques et systèmes biologiques. Cette deuxième perspective fera l'objet d'une thèse d'université, co-dirigée par les Prs Lamazière et Bachelot, dans la continuité de mon travail. Pour ce qui est de la première perspective, la validation du modèle sur une cohorte plus large et d'un nouveau centre est en cours de rédaction. Il s'agit de patientes du service d'endocrinologie de l'hôpital Saint-Antoine (Pr Christin-Maître et Dr Bruno Donadille). Après réentraînement avec les femmes de PSL et ajustement des coefficients du modèle PLS-DA, 80 SOPK et 29 NC21OHD ont été testés dans le nouveau modèle. Celui-là a permis d'identifier la présence de patiente sous traitement (2 patientes traitées par hydrocortisone, abaissant les concentrations de certains stéroïdes et classant à tort la patiente en zone contrôle/SOPK). En dehors de ces 2 patientes, sous traitement, l'ensemble des 107 femmes a été bien classé. Ces études ont inclus des patientes SOPK et NC21OHD, mais en pratique courante (= terrain de jeu réel du modèle à termes), d'autres étiologies peuvent consulter avec des motifs similaires. En ce sens, ont été également testés dans le modèle 113 patientes reçues en consultation pour troubles de l'ovulation et/ou hyperandrogénie, dont les diagnostics finaux retenus étaient les suivantes : 72 hyperandrogénies idiopathiques, 28 aménorrhée fonctionnelle, 3 syndromes de Cushing, et 10 hyperprolactinémies. L'ensemble de ces patientes ont bien été classées par le modèle (travail en cours de rédaction).

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^{1,2,3}, Anne Bachelot⁴, Marion Bonnier⁴,
Joe-Elie Salem⁵, Dominique Farabos², Severine Trabado^{6,7},
Charlotte Dupont^{1,2}, Peter Kamenicky^{7,8,9}, Muriel Houang^{2,10},
Jean Fiet^{2,3}, Yves Le Bouc^{2,10}, Jacques Young^{7,8,9}, and
Antonin Lamazière ^{2,3,*}

¹Service de Biologie de La Reproduction-CECOS, Hôpital Tenon, AP-HP.Sorbonne Université, Paris, France ²Sorbonne Université, Saint Antoine Research Center, INSERM UMR 938, Paris, France ³Département de Métabolomique Clinique, Hôpital Saint Antoine, AP-HP.Sorbonne Université, Paris, France ⁴Service d'Endocrinologie et Médecine de La Reproduction, Centre de Référence Des Maladies Endocriniennes Rares de La Croissance et du Développement, Centre Des Pathologies Gynécologiques Rares, Hôpital Pitié Salpêtrière (APHP), Sorbonne Université, Paris, France ⁵CIC-1901 and Department of Pharmacology, Pitié-Salpêtrière Hospital, AP-HP, Sorbonne Université, Inserm, Paris, France ⁶Service de Génétique Moléculaire, Pharmacogénétique et Hormonologie, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, Le Kremlin-Bicêtre, France ⁷INSERM UMR-UI185, Paris-Saclay University, Le Kremlin-Bicêtre, France ⁸University Paris-Saclay, Orsay, France ⁹Department of Reproductive Endocrinology, Assistance Publique-Hôpitaux de Paris, Bicêtre Hospital, Le Kremlin-Bicêtre, France ¹⁰Explorations Fonctionnelles Endocriniennes, Hôpital Armand Trousseau, AP-HP, Paris, France

*Correspondence address. Département de Métabolomique Clinique, Hôpital Saint Antoine, AP-HP.Sorbonne Université, 27 Rue Chaligny, 75012 Paris, France. E-mail: antonin.lamaziere@aphp.fr  <https://orcid.org/0000-0003-0813-7370>

Submitted on June 18, 2022; resubmitted on November 3, 2022; editorial decision on November 10, 2022

STUDY QUESTION: Can a combination of metabolomic signature and machine learning (ML) models distinguish nonclassic 21-hydroxylase deficiency (NC21OHD) from polycystic ovary syndrome (PCOS) without adrenocorticotrophic hormone (ACTH) testing?

SUMMARY ANSWER: A single sampling methodology may be an alternative to the dynamic ACTH test in order to exclude the diagnosis of NC21OHD in the presence of a clinical hyperandrogenic presentation at any time of the menstrual cycle.

WHAT IS KNOWN ALREADY: The clinical presentation of patients with NC21OHD is similar with that for other disorders of androgen excess. Currently, cosyntropin stimulation remains the gold standard diagnosis of NC21OHD.

STUDY DESIGN, SIZE, DURATION: The study was designed using a bicentric recruitment: an internal training set included 19 women with NC21OHD and 19 controls used for developing the model; a test set included 17 NC21OHD, 72 controls and 266 PCOS patients used to evaluate the performance of the diagnostic strategy thanks to an ML approach.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Fifteen steroid species were measured in serum by liquid chromatography–mass spectrometry (LC-MS/MS). This set of 15 steroids (defined as ‘steroidome’) used to map the steroid biosynthesis pathway was the input for our models.

MAIN RESULTS AND THE ROLE OF CHANCE: From a single sample, modeling involving metabolic pathway mapping by profiling 15 circulating steroids allowed us to identify perfectly NC21OHD from a confounding PCOS population. The constructed model using baseline LC-MS/MS-acquired steroid fingerprinting successfully excluded all 17 NC21OHDs (sensitivity and specificity of 100%) from

266 PCOS from an external testing cohort of originally 549 women, without the use of ACTH testing. Blood sampling timing during the menstrual cycle phase did not impact the efficiency of our model.

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

WIDER IMPLICATIONS OF THE FINDINGS: This tool makes it possible to propose a new semiology for the management of hyperandrogenism. The model presents better diagnostic performances compared to the current reference strategy. The management of patients may be facilitated by limiting the use of ACTH tests. Finally, the modeling process allows a classification of steroid contributions to rationalize the biomarker approach and highlight some underlying pathophysiological mechanisms.

STUDY FUNDING/COMPETING INTEREST(S): This study was supported by 'Agence Française de Lutte contre le dopage' and DIM Région Ile de France. This study was supported by the French institutional PHRC 2010-AOR10032 funding source and APHP. All authors declare no competing financial interests.

TRIAL REGISTRATION NUMBER: N/A.

Key words: 21-hydroxylase deficiency / late-onset congenital adrenal hyperplasia / hirsutism / machine learning / algorithm / polycystic ovary syndrome / hyperandrogenism

Introduction

Twenty-one-hydroxylase deficiency (21OHD) is a rare genetic disorder with autosomal recessive inheritance caused by deleterious mutations in *CYP21A2* (El-Maouche et al., 2017). Twenty-one OHD is the most common cause (95%) of congenital adrenal hyperplasia (CAH) (Speiser et al., 1992). The different clinical forms of 21OHD depend on the severity of the mutations. Besides the classical 21OHD form (C21OHD), which is due to complete or severe 21OH deficiency revealed at birth, there is the nonclassical 21OHD form (NC21OHD), which is due to partial enzymatic defect, with late-onset symptoms; diagnosis usually being made in late childhood or after puberty. The partial preservation of cortisol and aldosterone secretion in patients with NC21OHD (Kamenický et al., 2020) is explained by the existence of at least one of the alleles of a *CYP21A2* mutation encoding a 21-hydroxylase that retains a minimum residual enzyme activity of 50% (New et al., 2013). NC21OHD is most often revealed in adolescent girls and young women by a clinical pattern of hyperandrogenism accompanied by oligomenorrhea associated with increased circulating testosterone (Admoni et al., 2006). Infertility and episodes of miscarriage may also occur in some women (New, 2006). The clinical presentation of patients with NC21OHD is similar to that seen in a common female condition, polycystic ovarian syndrome (PCOS), wherein signs of hyperandrogenism associated with oligomenorrhea and infertility are also very prevalent (Matheson and Bain, 2019). As a result, many national and international consensuses recommend that the diagnosis of NC21OHD be systematically ruled out before the diagnosis of PCOS is definitively confirmed (Teede et al., 2018). To exclude NC21OHD, many consensuses recommend that an assay of 17-hydroxyprogesterone (17OHP) be performed in the morning during the follicular phase and that the diagnosis of NC21OHD be considered if the basal serum concentration of 17OHP is increased (Azziz et al., 1999; Speiser et al., 2000). The thresholds proposed in the literature are variable and depend on the 17OHP assay method used—that is, either immunoassay or mass spectrometry (Makela and Ellis, 1988). Thus, depending on the centers and methods, the diagnosis of NC21OHD will be considered if baseline 17OHP levels are ranging between 2 and 4 ng/ml (Azziz et al., 1999; Young et al., 2010). In such

cases, it is recommended that adrenocorticotrophic hormone (ACTH) stimulation be performed and the diagnosis established if the 17OHP peak value is greater than 10 ng/ml (Dewailly et al., 1986; Bello et al., 2017). However, a significant percentage of women with NC21OHD may have circulating levels of 17OHP close to or even below 2 ng/ml (Escobar-Morreale et al., 2008; Bidet et al., 2009, 2010). In addition, patients with increased serum 17OHP levels, but without NC21OHD, have also been reported (Rosenfield and Ehrmann, 2016).

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

The concept of profiling or multiplexing the circulating steroidome is an emerging paradigm shift, as this method appears to be more accurate than single-steroid immunoassay quantitation to stratify or classify patients from different conditions. Specifically, in a recent preliminary work, investigators have shown that, compared to an isolated measure of serum 17OHP, simultaneous measurements of a panel of several steroids in the basal state more accurately distinguish subjects with NC21OHD from a control population (Turcu et al., 2020a). In parallel to these advances in hormonal evaluation, machine learning (ML) approaches are emerging as interesting tools in diagnostic procedures due to their classification strength compared to classical univariate analysis traditionally used in clinical practice. As NC21OHD is an adrenal disorder and PCOS has an ovarian origin, we hypothesized that an

evaluation of a steroid panel including both predominantly adrenal steroids and others with a dual adrenal and ovarian origin would help to better distinguish between these two conditions.

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

Materials and methods

Study overview, population and data source

Study overview

As shown in Fig. 1, the study was designed using a bicentric recruitment: an internal training set included 38 female subjects (19 with NC21OHD and 19 healthy controls) from the Department of Reproductive Endocrinology, Bicêtre Hôpital, Le Kremlin-Bicêtre,

France (KB); an external test set included 355 women (17 NC21OHD, 72 controls and 266 PCOS) from the Department of Reproductive Endocrinology, Pitié-Salpêtrière Hospital, Paris, France (PSL), evaluated from 2019 to 2020.

The training set was used for developing the ML model. The test set was used to evaluate the performance of the model. PCOS patients (with a clinical presentation similar to those with NC21OHD) from PSL hospital were included in the test set to confirm that the model succeeded in classifying PCOS patients out of the NC21OHD group. The structure of this test cohort was designed to be in accordance with the prevalence of those conditions in a context of hyperandrogenism.

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

Subjects—population

Training set (KB). Nineteen women with NC21OHD with biallelic *CYP21A2* mutations were included as the training set in this study. These patients were previously evaluated in a clinical study (Kamenický *et al.*, 2020) prospective case–control open-label blinded-endpoint non-inferiority (ClinicalTrials.gov Id: NCT01862380) at a

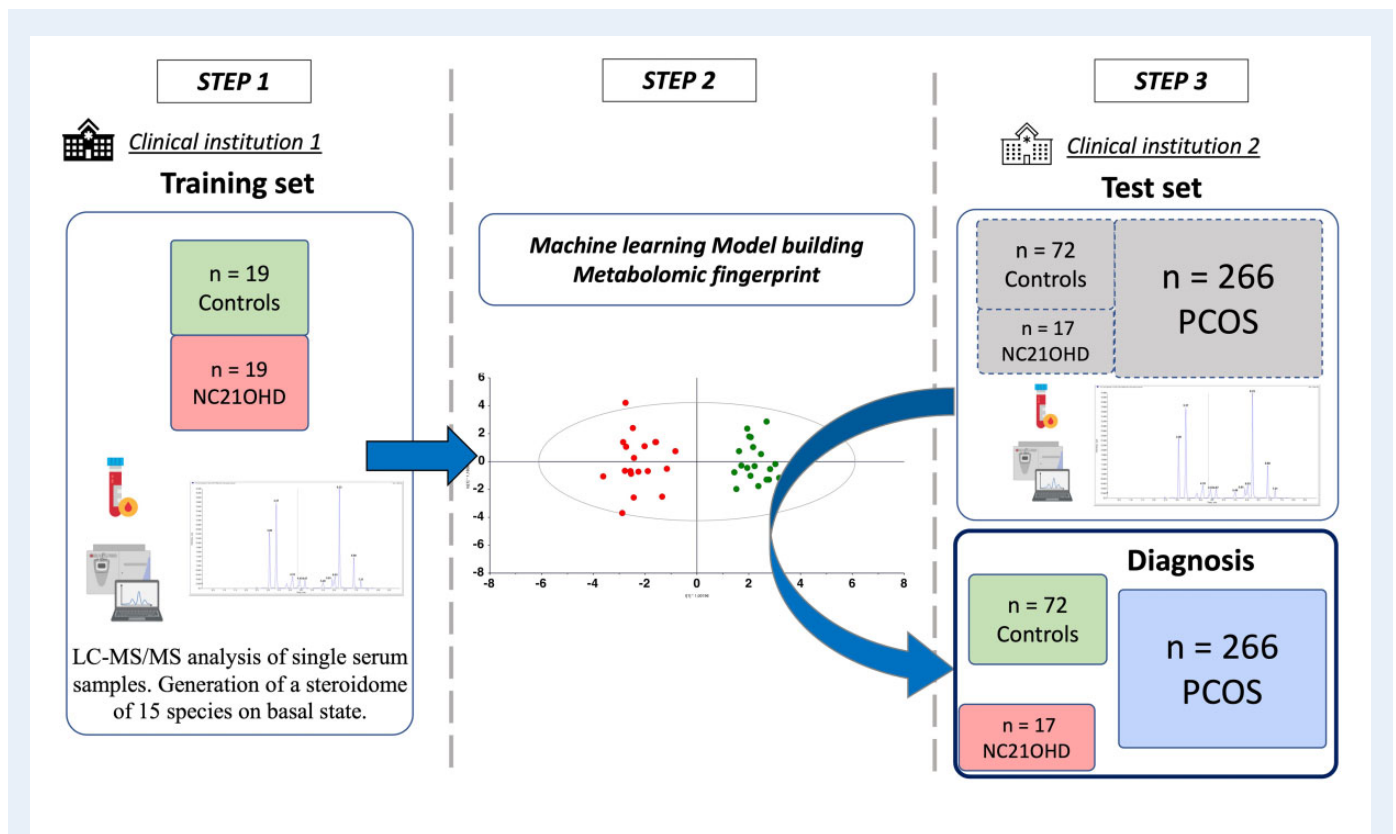


Figure 1. Study overview, step-by-step from steroidome quantitation to explainable clinical decision. Step 1: Fifteen steroid species were measured in serum by liquid chromatography–mass spectrometry (LC-MS/MS). Step 2: Orthogonal partial least squares discriminant analysis (OPLS-DA) was performed on the training set (Kremlin-Bictère (KB)). The Y axis represents the decision boundary and allowed to define two 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 Pitié-Salpêtrière (PSL). After measuring their steroidome (similarly to Step 1), the individuals of the test cohort were put into the model. The goal was to distinguish women with NC21OHD from both normal and polycystic ovary syndrome (PCOS) women without the use of an ACTH stimulation test.

tertiary referral medical center and clinical investigation center (2013–2016). This study was approved by the local ethics committee (Comité de Protection des Personnes Ile de France VII, approval no. 11-034) and the competent authority (ANSM). All subjects gave written informed consent to participate to the study. In brief, these eligible NC21OHD patients were premenopausal women aged 18–45 years with biallelic *CYP21A2* mutations (homozygotes or compound heterozygotes) and serum 17OHP concentration >10 ng/ml after stimulation with ACTH (250 µg of Synacthen®) performed at inclusion visit. Main clinical and genetics characteristics of these patients have been described elsewhere (Kamenický et al., 2020). In brief, seven of these patients carried the mild p.V281L *CYP21A2* mutation in the homozygous state and 12 were compound heterozygotes with at least one minor mutation (for details see Supplementary Table S1). Twelve patients received, at some point in their lives, hydrocortisone (10–20 mg per day) and/or dexamethasone (0.25–0.5 mg per day), prescribed transiently for important hyperandrogenism or impaired fertility, but all had received these treatments more than 1 year before inclusion. Nineteen age- and BMI-matched healthy women, whose characteristics have been detailed elsewhere (Kamenický et al., 2020), were included as controls. These eligible controls were premenopausal healthy women with serum 17OHP <2.0 ng/ml after stimulation with 250 µg of ACTH.

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

Test set (PSL). Two groups of patients with hyperandrogenism and/or hyperandrogenemia were recruited from this institution between 2019 and 2020 (Supplementary Fig. S1). We included 266 PCOS patients and 17 women with NC21OHD patients that were selected from routine clinical practice.

Thirteen of these 17 NC21OHD patients carried the mild p.V281L *CYP21A2* mutation in the homozygous state, and 4 were compound heterozygotes (V281L/P30L, IVS2-13C-1>G/P30L, IVS2-13C-1>G/V281L, I26C>T, I13G>A, I10T>C/I26C>T, I13G>A) with at least one minor mutation (for details see Supplementary Table S1). The diagnosis of PCOS was based on 2018 International Guidelines which resume the 2003 Rotterdam criteria (Teede et al., 2018). PCOS diagnosis was based on the presence of two out of the following three criteria: irregular menstrual cycles, clinical hyperandrogenism and/or hyperandrogenemia and the polycystic ovary on ultrasound after excluding hyperprolactinemia and Cushing syndrome (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004). CAH diagnosis was excluded according to the guidelines, i.e. a 17OHP value below 10 ng/ml after ACTH testing.

Data from 72 healthy female controls were also included from a multicenter prospective observational case–control study (2011–2014) exploring early cardiovascular damage in adult men and women with CAH due to late-onset 21 α -hydroxylase deficiency (CARDIOHCS (NCT01807364)) (Rosenbaum et al., 2019).

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

All patients gave written informed consent to participate. This study was performed in accordance with the Declaration of Helsinki. Healthy volunteers of this study were recruited by the Centre d'Investigation Clinique Paris Est (CIC-1421, Pitié-Salpêtrière Hospital, Paris). The ethics committee ('Comité de Protection des Personnes') approved the study.

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

Fifteen steroid species (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)) were measured in serum by LC-MS/MS as described elsewhere (Fiet et al., 2017). This set of 15 steroids (defined as the 'steroidome') used to map the steroid biosynthesis pathway were the input for our models. All the steroid analyses were performed in the same clinical laboratory (Saint-Antoine Hospital, Paris, France). There were no missing values.

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

After performing regular statistical analysis (mean comparison, Mann–Whitney–Wilcoxon using Python 3.7.6 environment and the SciPy Package version 1.4.1), we trained an ML model using the steroidome (set of 15 steroids) as input to detect (or predict) NC21OHD from both normal and PCOS women without the use of an ACTH stimulation test. We used orthogonal partial least squares (OPLS) as the standard metabolomic ML algorithm. As previously shown in Bachelot et al. (2021) and Agnani et al. (2022), we applied the OPLS algorithm for its performance on metabolomic data and the possibility to obtain easily interpretable score plot. The OPLS model was performed using SIMCA 16.0.1 software (Umetrics, Sweden).

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

As a classification task, the model's performance was evaluated using confusion matrix and accuracy score (the fraction of correct predictions).

Eventually, an additional evaluation was performed. We compared the performance of the model versus the standard cutoff based on the basal 17OHP concentration of 2 and 10 ng/ml.

Results

Study cohorts—baseline characteristics

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

In the test set (i.e. PSL), PCOS and NC21OHD patients exhibited similar baseline characteristics (Table 1), except for oligomenorrhea,

Table 1 Clinical and biological characteristics of polycystic ovary syndrome (PCOS) and non-classic 21OH deficiency (NC21OHD) patients from test set.

	PCOS (n = 266) [§]	NC21OHD (n = 17) ^{§§}	P-value
Baseline characteristics			
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)	**
LC-MS/MS steroids			
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)	**

Data are expressed as median (interquartile range). Unpaired two-tailed Mann–Whitney *U* test.

P* < 0.05; *P* < 0.001; n.s., not significant.

[§]One 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 1 h after ACTH stimulation, 58 patients with unknown follicular number right and 60 patients with unknown follicular number left.

^{§§}Two patients with unknown FSH, LH and estradiol, one patient with unknown 17OHP 1 h after ACTH stimulation and six patients with unknown follicular number left and right.

follicle number (right and left) and serum LH. As expected, these parameters were increased in PCOS compared with NC21OHD.

As a bicentric study, basal steroidomic profiles were compared in the different control (KB versus PSL) and NC21OHD (KB versus PSL) groups. As shown in [Supplementary Table SII](#), levels of 17OHP, progesterone, DOC, 21-DF, corticosterone, aldosterone and cortisone were homogeneous in patients and controls from both institutions. In the control groups, some steroid species—such as testosterone, 21 deoxycortisol and 17OH-pregnenolone—exhibited differences between the two institutions, thereby reflecting natural inter-individual variabilities, while remaining within the same orders of magnitude and within the normal value ranges.

[Figure 2A and B](#) respectively show the circulating basal concentrations of 17OHP and circulating concentrations of 17OHP after ACTH stimulation in the two groups of NC21OHD patients (*n* = 36) and PCOS patients (*n* = 266), which were significantly different (*P* < 0.0001). Some overlapping among NCAH and PCOS patients was observed for basal 17OHP concentrations. However, all PCOS (*n* = 266) exhibited a

concentration well below the 10 ng/ml cutoff values of 17OHP after injection. Likewise, none of the NC21OHD patients (*n* = 36) exhibited post-stimulation 17OHP values below 10 ng/ml. Thus, there was strictly no overlap between NC21OHDs (with a minimum stimulated 17OHP concentration = 16.40 ng/ml) and PCOSs (with a maximum elevated 17OHP concentration = 4.55 ng/ml). These data are in accordance with the literature on 21OHD, which proposed that when a patient was genotyped with a biallelic mutation, the peak after ACTH was consistently above 10 ng/ml ([Bidet et al., 2009, 2010](#)). These results excluded the possibility that any of the PCOS patients in the test set would be affected by 21OHD.

From model to diagnosis

The main goal of our study was to avoid the ACTH stimulation test through the use of mathematical algorithms and metabolomics. For that purpose, we built the model described in [Fig. 1](#), which was implemented with the circulating basal steroids signature (15 steroids) and

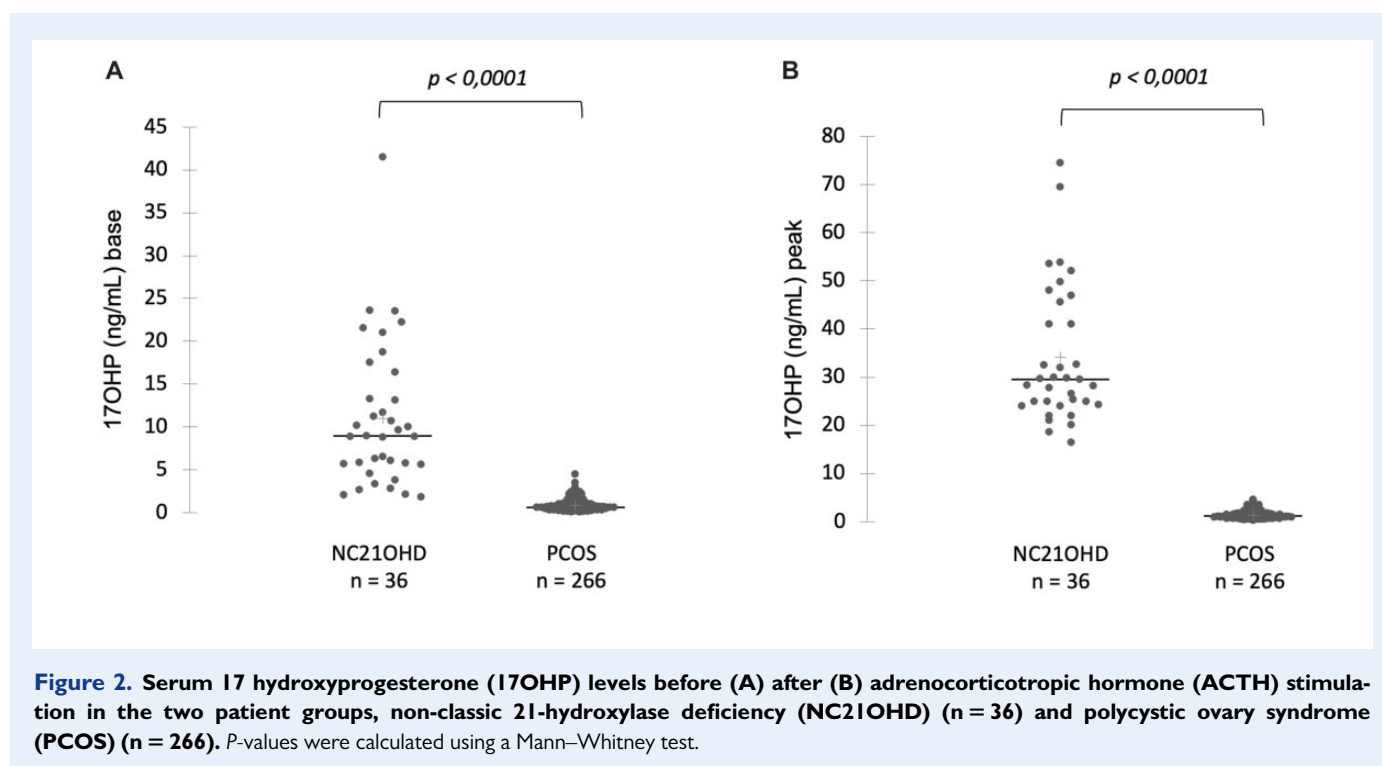


Figure 2. Serum 17 hydroxyprogesterone (17OHP) levels before (A) after (B) adrenocorticotrophic hormone (ACTH) stimulation in the two patient groups, non-classic 21-hydroxylase deficiency (NC21OHD) (n = 36) and polycystic ovary syndrome (PCOS) (n = 266). P-values were calculated using a Mann–Whitney test.

trained to efficiently differentiate two groups of individuals: female controls and NC21OHD patients.

Machine learning modeling

Step 1: Circulating basal steroids signatures were acquired from LC-MS/MS and preprocessing (normalization) (Fig. 1).

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

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

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

As mentioned above, one important clinical issue is the differentiation between NC21OHD and polycystic ovary syndrome (PCOS) in

female patients with hyperandrogenism (Table I). Therefore, PCOS patients from PSL were tested in the model. The objective was to ensure that the algorithm was able to accurately identify and segregate the NC21OHD patients, not only from controls but also from PCOS patients. After being tested on the NC21OHD/controls model, PCOS patients were plotted on OPLS-DA (Fig. 3D); all were located in the controls area (266 out 266).

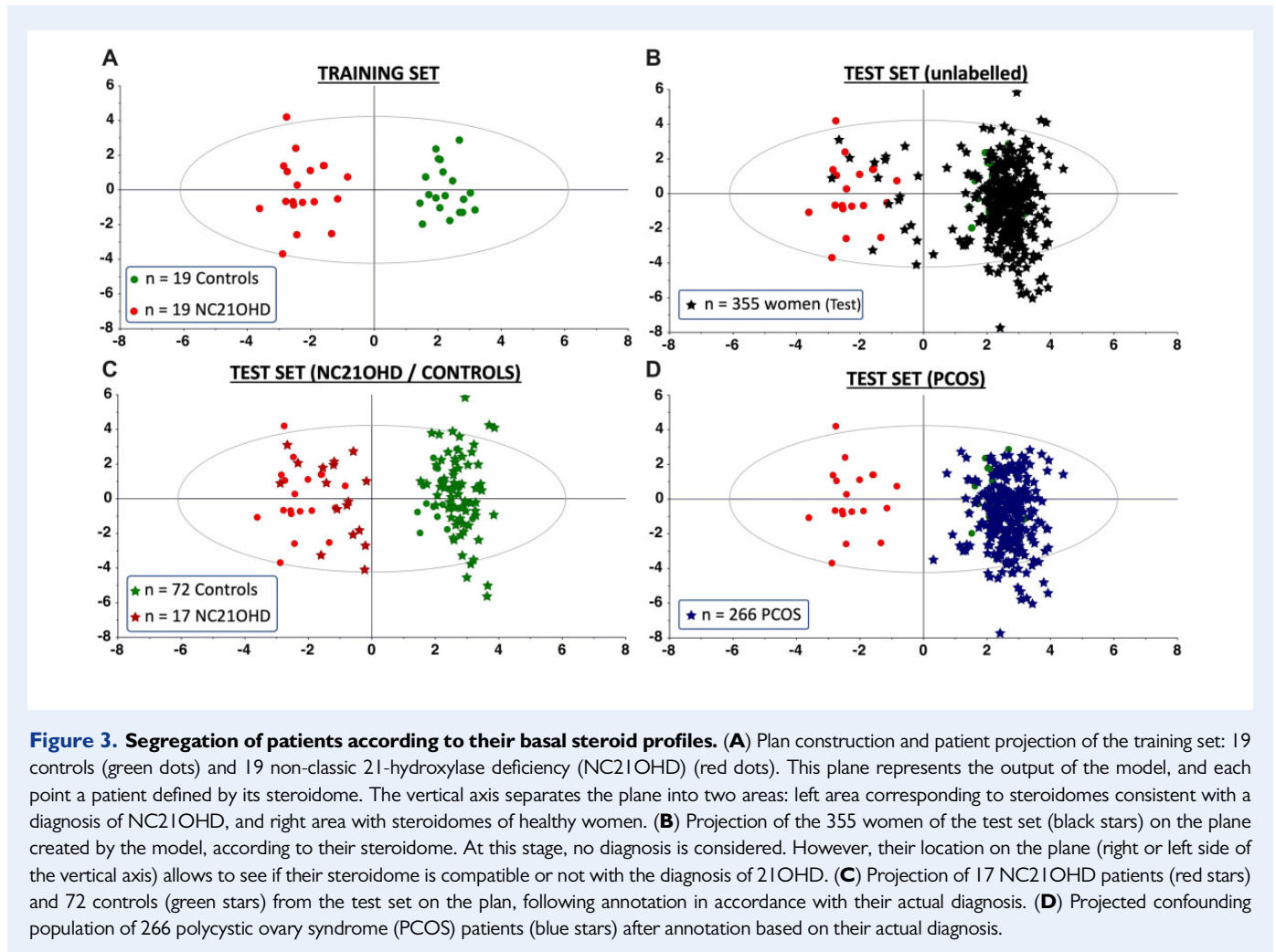
The Variable Importance for the Projection (VIP) plot summarized the importance (contribution) of the steroids in the model (Fig. 4). The VIP plot displays the VIP values as a column plot with CIs (95%) derived from jack-knifing. The most segregative basal steroid species were those with VIP values larger than 1, which included 21-DF, 21-DB, 11 β -OHA, 17OHP, androstenedione (A), corticosterone (B) and testosterone (T). Interestingly, except for corticosterone, all these steroids precursors are upstream of the enzymatic blockade observed in subjects affected by NC21OHD.

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

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

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

As shown in Table II, despite the use of LC-MS/MS (more sensitive and more specific than immunoassay), using a 17OHP cutoff value of 2 ng/ml (Table II), the test has a sensitivity of 94.1% and a specificity of 92.6% for the identification of NC21OHD cases. It corresponded to



I false negative and 25 false positives on the test set (PSL). Using a 17OHP cutoff value of nonstimulated 10 ng/ml (Table II), the test showed a sensitivity of 29.4% and a specificity of 100%.

Based on our data, the ROC analysis (Youden's index) suggested a basal 17OHP measured by LC-MS/MS cutoff level of 1.76 ng/ml as the optimal cutoff. This cutoff provided 94.1% sensitivity, but yielded 36 false positives results, leading to a more widespread use of genetic testing.

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

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

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

therefore evaluate the impact of the luteal phase of the cycle on the performance of our model.

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

Discussion

To the best of our knowledge, this proof-of-concept study is the first attempt to use a relevant steroid profile combined with an ML approach in order to build a model that reliably distinguishes women with NC21OHD from controls as well as PCOS women without the use of an ACTH stimulation test and regardless of the phase of the menstrual cycle.

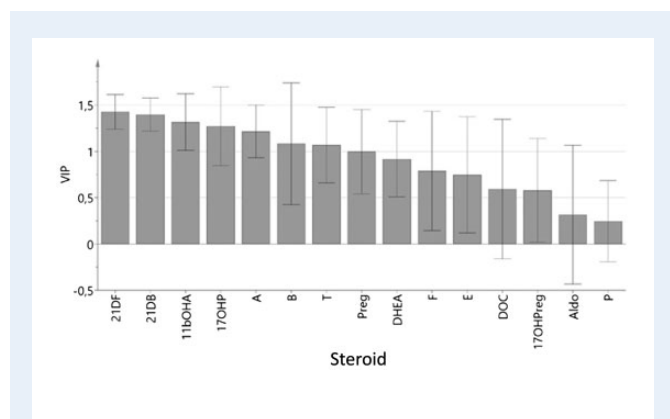


Figure 4. Major contributing steroids to distinguish non-classic 21-hydroxylase deficiency (NC21OHD) from controls and polycystic ovary syndrome (PCOS) within the model. The Variable Importance for the Projection (VIP) plot displays the VIP values as a column plot with CIs (95%) derived from jack-knifing (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 21-deoxycortisol (21-DF), 21-deoxycorticosterone (21-DB), 11 β -hydroxyandrostenedione (11 β -OHA), 17-OH-progesterone (17OHP), androstenedione (A), corticosterone (B) and testosterone (T).

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

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

We have shown here that isolated basal levels of serum 17OHP cut-offs, even quantified using LC-MS/MS, do not allow sufficient sensibility and specificity to effectively distinguish NC21OHD from PCOS. Indeed, despite using the LC-MS/MS method, which is more sensitive and specific than the conventional immunological assays (Wong et al., 1992; Minutti et al., 2004; Fiet et al., 2017), the mere application of the classic thresholds of 2 or 10 ng/ml for basal 17OHP on the patients of the replication cohort yields many misclassifications (1 false negative and 25 false positives for 2 ng/ml basal 17OHP cutoff).

Table II Comparison of the discriminating power of the basal 17-hydroxyprogesterone (17OHP) threshold used for screening (2 ng/ml); the threshold used for diagnosis (10 ng/ml) as well as the machine learning model in identifying non-classic 21OH deficiency (N21HOD) cases.

Screening value: 2 ng/ml (17OHP)				
Actual 17 NC21OHD				
		No	Yes	Total
Predicted NC21OHD	No	True negative = 313	False negative = 1	314
	Yes	False positive = 25	True positive = 16	41
Total		338	17	355
Diagnosis value: 10 ng/ml (17OHP)				
Actual 17 NC21OHD				
		No	Yes	Total
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				
		No	Yes	Total
Predicted NC21OHD	No	True negative = 338	0	338
	Yes	0	True positive = 17	17
Total		338	17	355

Contingency table using basal (17OHP) cutoff value of 2 ng/ml has a sensitivity of 94.1% and a specificity of 92.6%; contingency table using basal (17OHP) cutoff value of 10 ng/ml showed a sensitivity of 29.4% and a specificity of 100%. Contingency table using the model developed from the 15 steroids in prime rate. The model exhibiting sensitivity and specificity of 100%.

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

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

Turcu et al. (2020a) recently proposed a logistic regression restricted to a small set of three circulating steroids (17OHP, 21DF and

corticosterone) measured by LC-MS to discriminate between patients with NC21OHD and unaffected individuals. This preliminary monocentric study exhibited an interesting performance to distinguish healthy women from NC21OHD. However, it did not evaluate the ability of this approach to differentiate NC21OHD from a larger representative cohort of women with PCOS. In addition, the absence of an external test set did not fully allow an unbiased evaluation of the performance of the logistic regression. Considering the vast heterogeneity of PCOS clinical expression and its epidemiology, the inclusion by Turcu *et al.* (2020a) of only 27 patients with PCOS in the group of 'unaffected individuals' may have been biased and not representative of the well-recognized wide hormonal spectrum observed in women with PCOS. However, this strategy (multivariate versus 17OHP only) is mathematically relevant because it reduces the risk of spurious results (type I error/false positive and type II error/false negative) (Wold *et al.*, 2001; Eriksson *et al.*, 2013); however, one may argue that in Turcu *et al.* (2020a), the authors focused only on few pre-selected steroids. In our model, we used a less restricted strategy including 15 steroids to more comprehensively map the steroidal pathway, incorporating steroids that are predominantly—if not exclusively—of adrenal origin (21DB, 21DF and 11 β OH androstenedione). In our experience, these steroids are particularly adapted to compare the level hormone synthesis by each gland, distinguishing adrenal from ovarian disease (Fiet *et al.*, 2017; Pretorius *et al.*, 2017; Turcu *et al.*, 2020b; Agnani *et al.*, 2022).

Mathematically, a set of variables is considered to be collinear if there are one or more relations or associations between those variables. These 15 steroids are interrelated in their biosynthetic pathway, resulting in a high degree of collinearity. Consequently, and based on our previous work (Bachelot *et al.*, 2021; Agnani *et al.*, 2022), we used the OPLS approach: the latter is less affected by multicollinearity than logistic regression because its components are uncorrelated (i.e. orthogonal) (Wold *et al.*, 2001; Pérez-Enciso and Tenenhaus, 2003; Kjeldahl and Bro, 2010). On a larger set of patients ($n=366$), our model allowed us to successfully exclude all (100%) patients with NC21OHD from the independent test cohort (including PCOS, controls and NC21OHD). Practically, we could reduce the dimension of our system by considering fewer variables (e.g. the four most discriminating steroids: 21DF, 21DB, 11BOHA, 17OHP) for an equally efficient NC21OHD/PCOS differential diagnosis. From a practical point of view, it must be stressed that a quantitative profiling by LC-MS/MS of the 15 steroids in a 10-min chromatographic run is not a limiting factor and indeed facilitates its real-life clinical application (Fiet *et al.*, 2017).

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

For years, 17OHP has been used—and is still used by many practitioners—to diagnose NC21OHD (Carmina *et al.*, 2017). However, as stated before we have shown here the interest of 21-deoxysteroid metabolites in the diagnosis of NC21OHD. Indeed, 21DF and

21DB—being strictly adrenal metabolites—may represent more specific biomarkers than 17OHP, which has a double ovarian and adrenal origin (Turcu *et al.*, 2014). The exclusively adrenal origin of these steroids is likely an explanation for the significant improvement of the diagnostic performance of our model compared to the isolated use of a simple 17OHP. In women with PCOS, the main source of androstenedione is in the ovary, and this androgen is not submitted to the direct action of adrenal 11-hydroxylase; consequently, the 11 β hydroxyandrostenedione serum concentration found was similar to that in normal subjects, and the ratio of androstenedione to 11-OHA was higher than in normal women (Polson *et al.*, 1988).

Because of the ovarian and adrenal origin, basal 17OHP should be measured during the follicular phase to detect NC21OHD. Indeed, an ovarian surge of 17OHP occurs until 3 ng/ml during the luteal phase of the menstrual cycle. Moreover, slightly elevated 17OHP has been reported in women with PCOS (Pall *et al.*, 2010). The inclusion of several patients ($n=31$, 9.2%) in our series who were in the luteal phase (i.e. with serum progesterone level above 10 ng/ml) suggests that our model is accurate independently of the timing of blood sampling.

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

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

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

Conclusion

The aim of this work was to develop a unique sampling methodology to serve as an alternative to the dynamic ACTH test to exclude the diagnosis of NC21OHD in the presence of a clinical hyperandrogenic presentation at any time of the menstrual cycle, facilitating easier management and greater patient comfort. This methodology was based on a combination of ML with metabolic profiling of 15 molecular species (steroids) in LC-MS/MS at the base level. This first proof of concept is

promising; however, it will require further evaluation in real clinical situations in the future.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

The data underlying this article will be shared upon reasonable request to the corresponding author.

Acknowledgements

We thank Nicolas Helin and Bettina Ribault for their technical support.

Authors' roles

G.B. performed lipidomics analysis, statistical modeling analysis, drafted the article, participated in the study conception and design, interpretation of data. M.B., J.-E.S., D.F., S.T., C.D., P.K., M.H., J.F. and Y.L.B. participated in acquisition of data, patients care and in critical revision of the article for intellectual content. J.Y. and A.B. supervised the study, participated in study conception and design, drafting the article and in critical revision of the article for intellectual content. A.L. supervised the study, lipidomics analysis, statistical modeling analysis, participated in study conception and design, interpretation of data and drafting the article.

Funding

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

Conflict of interest

All authors declare no competing financial interests.

References

- Admoni O, Israel S, Lavi I, Gur M, Tenenbaum-Rakover Y. Hyperandrogenism in carriers of CYP21 mutations: the role of genotype. *Clin Endocrinol (Oxf)* 2006;**64**:645–651.
- Agnani H, Bachelot G, Eguether T, Ribault B, Fiet J, Le Bouc Y, Netchine I, Houang M, Lamazière A. A proof of concept of a machine learning algorithm to predict late-onset 21-hydroxylase deficiency in children with premature pubic hair. *J Steroid Biochem Mol Biol* 2022;**220**:106085.
- Azziz R, Hincapie LA, Knochenhauer ES, Dewailly D, Fox L, Boots LR. Screening for 21-hydroxylase-deficient nonclassic adrenal hyperplasia among hyperandrogenic women: a prospective study. *Fertil Steril* 1999;**72**:915–925.
- Bachelot G, Lévy R, Bachelot A, Faure C, Czernichow S, Dupont C, Lamazière A; Alifert Collaborative Group. Proof of concept and

- development of a couple-based machine learning model to stratify infertile patients with idiopathic infertility. *Sci Rep* 2021;**11**:24003.
- Bello R, Lebenthal Y, Lazar L, Shalitin S, Tenenbaum A, Phillip M, de Vries L. Basal 17-hydroxyprogesterone cannot accurately predict nonclassical congenital adrenal hyperplasia in children and adolescents. *Acta Paediatr* 2017;**106**:155–160.
- Bidet M, Bellanné-Chantelot C, Galand-Portier M-B, Golmard J-L, Tardy V, Morel Y, Clauin S, Coussieu C, Boudou P, Mowzowicz I et al. Fertility in women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *J Clin Endocrinol Metab* 2010;**95**:1182–1190.
- Bidet M, Bellanné-Chantelot C, Galand-Portier M-B, Tardy V, Billaud L, Laborde K, Coussieu C, Morel Y, Vaury C, Golmard J-L et al. Clinical and molecular characterization of a cohort of 161 unrelated women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency and 330 family members. *J Clin Endocrinol Metab* 2009;**94**:1570–1578.
- Carmina E, Dewailly D, Escobar-Morreale HF, Kelestimur F, Moran C, Oberfield S, Witchel SF, Azziz R. Non-classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency revisited: an update with a special focus on adolescent and adult women. *Hum Reprod Update* 2017;**23**:580–599.
- Dewailly D, Vantyghem-Haudiquet MC, Sainsard C, Buvat J, Cappoen JP, Ardaens K, Racadot A, Lefebvre J, Fossati P. Clinical and biological phenotypes in late-onset 21-hydroxylase deficiency. *J Clin Endocrinol Metab* 1986;**63**:418–423.
- El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. *Lancet* 2017;**390**:2194–2210.
- Eriksson L, Byrne T, Johansson E, Trygg J, Vikström C. *Multi- and Megavariable Data Analysis Basic Principles and Applications*. Sweden: Umetrics Academy, 2013.
- Escobar-Morreale HF, Sanchón R, San Millán JL. A prospective study of the prevalence of nonclassical congenital adrenal hyperplasia among women presenting with hyperandrogenic symptoms and signs. *J Clin Endocrinol Metab* 2008;**93**:527–533.
- Fiet J, Le Bouc Y, Guéchet J, Hélin N, Maubert MA, Farabos D, Lamazière A. A liquid chromatography/tandem mass spectrometry profile of 16 serum steroids, including 21-deoxycortisol and 21-deoxycorticosterone, for management of congenital adrenal hyperplasia. *J Endocr Soc* 2017;**1**:186–201.
- Kamenický P, Blanchard A, Lamazière A, Piedvache C, Donadille B, Duranteau L, Bry H, Gautier J-F, Salenave S, Raffin-Sanson M-L. Cortisol and aldosterone responses to hypoglycemia and Na depletion in women with non-classic 21-hydroxylase deficiency. *J Clin Endocrinol Metab* 2020;**105**:dgz005.
- Kjeldahl K, Bro R. Some common misunderstandings in chemometrics. *J Chemometrics* 2010;**24**:558–564.
- Kuttann F, Coullin P, Girard F, Billaud L, Vincens M, Boucekkinne C, Thalabard JC, Maudelonde T, Spritzer P, Mowszowicz I. Late-onset adrenal hyperplasia in hirsutism. *N Engl J Med* 1985;**313**:224–231.
- Livadas S, Dracopoulou M, Dastamani A, Sertedaki A, Maniati-Christidi M, Magiakou A-M, Kanaka-Gantenbein C, Chrousos GP, Dacou-Voutetakis C. The spectrum of clinical, hormonal and molecular findings in 280 individuals with nonclassical congenital adrenal hyperplasia caused by mutations of the CYP21A2 gene. *Clin Endocrinol (Oxf)* 2015;**82**:543–549.

- Makela SK, Ellis G. Nonspecificity of a direct 17 alpha-hydroxyprogesterone radioimmunoassay kit when used with samples from neonates. *Clin Chem* 1988;**34**:2070–2075.
- Matheson E, Bain J. Hirsutism in women. *Am Fam Physician* 2019;**100**:168–175.
- Minutti CZ, Lacey JM, Magera MJ, Hahn SH, McCann M, Schulze A, Cheillan D, Dorche C, Chace DH, Lymp JF et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab* 2004;**89**:3687–3693.
- New MI. Extensive clinical experience: nonclassical 21-hydroxylase deficiency. *J Clin Endocrinol Metab* 2006;**91**:4205–4214.
- New MI, Abraham M, Gonzalez B, Dumic M, Razzaghy-Azar M, Chitayat D, Sun L, Zaidi M, Wilson RC, Yuen T. Genotype-phenotype correlation in 1,507 families with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. *Proc Natl Acad Sci USA* 2013;**110**:2611–2616.
- Pall M, Azziz R, Beires J, Pignatelli D. The phenotype of hirsute women: a comparison of polycystic ovary syndrome and 21-hydroxylase-deficient nonclassic adrenal hyperplasia. *Fertil Steril* 2010;**94**:684–689.
- Papadakis G, Kandaraki EA, Tseniklidi E, Papalou O, Diamanti-Kandaraki E. Polycystic ovary syndrome and NC-CAH: distinct characteristics and common findings. *Front Endocrinol (Lausanne)* 2019;**10**:388.
- Pérez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Hum Genet* 2003;**112**:581–592.
- Polson DW, Reed MJ, Franks S, Scanlon MJ, James VHT. Serum 11 β -hydroxyandrostenedione as an indicator of the source of excess androgen production in women with polycystic ovaries. *J Clin Endocrinol Metab* 1988;**66**:946–950.
- Pretorius E, Artl W, Storbeck K-H. A new dawn for androgens: novel lessons from 11-oxygenated C19 steroids. *Mol Cell Endocrinol* 2017;**441**:76–85.
- Pugeat M, Déchaud H, Raverot V, Denuzière A, Cohen R, Boudou P. Recommandations pour l'exploration des hyperandrogénies. *Annales D'endocrinologie* 2010;**71**:e3–e8.
- Rosenbaum D, Gallo A, Lethielleux G, Bruckert E, Levy BI, Tanguy ML, Dulon J, Dahmoune N, Salem JE, Bittar R et al.; CARDIOHCS study group. Early central blood pressure elevation in adult patients with 21-hydroxylase deficiency. *J Hypertens* 2019;**37**:175–181.
- Rosenfield RL, Ehrmann DA. The pathogenesis of polycystic ovary syndrome (PCOS): the hypothesis of PCOS as functional ovarian hyperandrogenism revisited. *Endocr Rev* 2016;**37**:467–520.
- Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004;**19**:41–47.
- Speiser PW, Dupont J, Zhu D, Serrat J, Buegeleisen M, Tusie-Luna MT, Lesser M, New MI, White PC. Disease expression and molecular genotype in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *J Clin Invest* 1992;**90**:584–595.
- Speiser PW, Knochenhauer ES, Dewailly D, Fruzzetti F, Marcondes JA, Azziz R. A multicenter study of women with nonclassical congenital adrenal hyperplasia: relationship between genotype and phenotype. *Mol Genet Metab* 2000;**71**:527–534.
- Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ; International PCOS Network. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Hum Reprod* 2018;**33**:1602–1618.
- Turcu A, Smith JM, Auchus R, Rainey WE. Adrenal androgens and androgen precursors—definition, synthesis, regulation and physiologic actions. *Compr Physiol* 2014;**4**:1369–1381.
- Turcu AF, El-Maouche D, Zhao L, Nanba AT, Gaynor A, Veeraraghavan P, Auchus RJ, Merke DP. Androgen excess and diagnostic steroid biomarkers for nonclassic 21-hydroxylase deficiency without cosyntropin stimulation. *Eur J Endocrinol* 2020a;**183**:63–71.
- Turcu AF, Wannachalee T, Tsodikov A, Nanba AT, Ren J, Shields JJ, O'Day PJ, Giachero D, Rainey WE, Auchus RJ. Comprehensive analysis of steroid biomarkers for guiding primary aldosteronism subtyping. *Hypertension* 2020b;**75**:183–192.
- Ueland GÅ, Dahl SR, Methlie P, Hessen S, Husebye ES, Thorsby PM. Adrenal steroid profiling as a diagnostic tool to differentiate polycystic ovary syndrome from nonclassic congenital adrenal hyperplasia: pinpointing easy screening possibilities and normal cutoff levels using liquid chromatography tandem mass spectrometry. *Fertil Steril* 2022;**118**:384–391.
- Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 2001;**58**:109–130.
- Wong T, Shackleton CH, Covey TR, Ellis G. Identification of the steroids in neonatal plasma that interfere with 17 alpha-hydroxyprogesterone radioimmunoassays. *Clin Chem* 1992;**38**:1830–1837.
- Young J, Tardy V, Brac de la Perrière A, Bachelot A, Morel Y. Dépistage et prise en charge du déficit en 21-hydroxylase à révélation tardive chez les femmes avec hyperandrogénie. *Ann Endocrinol* 2010;**71**:e15–e19.

Discussion et Perspectives

Discussion et perspectives générales

Cette thèse aborde des problématiques et des pathologies courantes liées à l'infertilité, pour lesquelles il existe encore de nombreuses incertitudes, et un recours fréquent à des traitements empiriques. En effet, la complexité de ces pathologies rend l'établissement d'un consensus difficile, notamment en raison de l'hétérogénéité des populations étudiées et des difficultés de comparaison entre les études. Cette complexité est illustrée par les révisions régulières des critères diagnostiques et thérapeutiques.

Sur le plan évolutif, la reproduction et la fertilité occupent une place centrale, et l'une des théories suggère l'existence d'un décalage entre l'évolution culturelle (rapide) et l'évolution biologique (lente). Cette divergence ou conflit évolutif entre le mode de vie contemporain, l'environnement moderne et la biologie, façonnée au fil de longues périodes, pourrait en partie expliquer ou contribuer à certaines conditions pathologiques.

Le mode de vie moderne semble jouer un rôle essentiel dans l'apparition et l'évolution de ces troubles métaboliques complexes, nécessitant ainsi une prise en charge globale, à la fois individualisée et à l'échelle de la santé publique. Les répercussions de ce mode de vie inadéquat sont difficiles à comprendre, à définir et à évaluer, d'autant plus que ces conséquences se manifestent à des degrés différents d'un individu à l'autre.

Ces répercussions peuvent être directes, mais se manifestent surtout par le biais du métabolisme, ou plutôt par les altérations qu'elles induisent sur les métabolismes, qui sont finement régulés. Ces différents métabolismes ou axes métaboliques (stéroïdes, acides gras, statuts glucidique et oxydatif, axe gonadotrope...) interagissent de diverses manières, ce qui ajoute une complexité supplémentaire. Ces problématiques abordées dans ce travail nécessitent donc d'être étudiées sous plusieurs angles afin d'être appréhendées de manière la plus complète possible. C'est ainsi que ces questions se révèlent être un terrain propice à l'application des méthodes de l'intelligence artificielle.

Comme nous l'avons vu dans cette thèse, et compte tenu des multiples similarités retrouvées et observées, un dénominateur commun métabolique pourrait exister entre certains cas d'infertilité dits inexplicables et les troubles ovulatoire type SOPK. Ce dénominateur

commun métabolique pourrait permettre d'homogénéiser et rationaliser les prises en charge, notamment sur les leviers du mode de vie. Bien que les recommandations hygiéno-diététiques occupent une place importante dans la prise en charge de ces pathologies, il existe encore un manque de preuves solides quant à leur efficacité, seules ou en association avec des techniques d'assistance médicale à la procréation telles que la FIV ou l'IIU.

Sur le plan mathématique, on peut schématiquement envisager que l'IA puisse être utilisée pour 2 grands types d'intervention en biologie et en médecine : 1) exécuter des tâches qu'un médecin est en mesure de faire (comme identifier la présence d'un bloc en 21-hydroxylase), d'une manière plus simple et/ou plus directe et/ou plus efficace, et 2) réaliser des tâches inaccessibles au médecin (prédire l'apparition d'une maladie ou d'une grossesse, identifier la présence d'une infertilité inexplicée). Les deux applications présentées dans ce travail illustrent parfaitement ces deux types d'interventions et le potentiel de l'IA.

Les résultats obtenus à partir de ces deux applications révèlent l'importance de la ou des hypothèses en amont dans le choix des variables pertinentes. En ML, on parle souvent de la connaissance métier (dans le triptyque mathématiques/informatique/domaine d'application), élément central dans la réussite d'une tâche de modélisation. En biologie et en médecine, ces hypothèses reposent généralement sur la connaissance de la physiologie et/ou physiopathologie de la maladie étudiée. Concrètement, plus notre compréhension des mécanismes de la maladie que nous étudions est approfondie et complète, plus le choix des variables à modéliser sera pertinent, et plus le modèle sera performant. Autrement dit, plus les données intégrées dans le modèle sont proches du phénomène étudié, plus le modèle sera efficace (et moins il faudra d'exemples, et donc de données, pour capturer l'information pertinente).

A ce titre, l'exemple de l'identification de la NC21OHD est illustratif : la voie de biosynthèse des stéroïdes est mesurée, cartographiée et modélisée, dans une pathologie causée par un bloc enzymatique dans cette même voie. La proximité entre les données collectées pour le modèle et le phénomène à expliquer est presque totale, de sorte que le modèle peut saisir l'information contenue dans les données en utilisant peu de patients ou exemples. Inversement, les mécanismes moléculaires impliqués dans l'infertilité inexplicée étant, par définition, beaucoup moins bien connus, la modélisation des données de la cohorte ALIFERT a conduit à un modèle moins performant. Notons que l'utilisation des VIP dans un modèle PLS permet également de faire des allers-retours entre la modélisation et la physiopathologie, et éventuellement de générer de nouvelles hypothèses et d'intégrer de nouvelles variables. Cette

compétence est particulièrement pertinente pour approfondir la compréhension des mécanismes moléculaires.

D'un point de vue pratique, travailler sur des données rétrospectives qui n'ont pas été prévues à cet effet soulève plusieurs difficultés : données manquantes, prétraitement fastidieux, etc... Ces limites soulignent l'importance du design de l'étude et de la qualité du phénotypage plutôt que des techniques IA appliquées. Les professionnels de santé sont ainsi au cœur de la construction des modèles d'IA. Les différentes étapes du processus, de la conception de l'étude au déploiement, impliquent donc une diversité de profils et de compétences. Une collaboration étroite entre ces différents experts complémentaires est donc indispensable pour développer un modèle prédictif utilisable dans la pratique clinique quotidienne.

Enfin, et dans un objectif d'utilisation en routine clinique, l'étape de validation des modèles d'IA est particulièrement importante dans le domaine médical. Le cadre juridique des dispositifs médicaux numériques utilisant l'intelligence artificielle est en constante évolution pour répondre aux défis posés par ces technologies émergentes. Les autorités réglementaires, telles que la Food and Drug Administration (FDA) aux États-Unis et l'Agence européenne des médicaments (EMA) en Europe, élaborent des lignes directrices et des normes pour encadrer l'utilisation sûre et efficace de ces dispositifs. Ces directives visent à assurer la sécurité des patients, la fiabilité des résultats et la protection des données personnelles. Elles définissent les exigences en matière de validation clinique, de surveillance continue, de traçabilité des données, de transparence des algorithmes et de responsabilité des fabricants. Les dispositifs médicaux utilisant l'intelligence artificielle doivent généralement obtenir une autorisation réglementaire avant leur commercialisation, qui peut varier en fonction du niveau de risque associé au dispositif.

En ce sens, le modèle couple développé au cours de cette thèse, qui a fait par ailleurs l'objet d'un brevet AP-HP/Sorbonne Université, a été certifié dispositif médical de classe I au cours de l'été 2023 (<https://www.alifert.com/>).

Miniaturisation des prélèvements

Malgré la puissance de la combinaison omique/IA, l'une des principales limites réside dans la dépendance à des techniques analytiques de pointe telles que la LC-MS/MS. Dans le but de diffuser cette approche et ces scores à une plus grande échelle et de manière plus accessible, une perspective consiste à développer des techniques de micro-prélèvement. En effet, conscients que la LC-MS/MS est une technique sophistiquée non disponible dans tous les laboratoires, le recours à des méthodes de micro-prélèvements à domicile pourrait permettre une diffusion plus large de ces technologies et de leur combinaison avec les systèmes d'IA associés. De plus, face à la nécessité pour valider les modèles et à la difficulté pour recruter des patients dans les études, l'utilisation de ces méthodes pourrait faciliter et élargir géographiquement les recrutements.

Fort de l'expérience du département de métabolomique clinique de Saint-Antoine dans le dépistage néonatal à l'aide de taches de sang sur buvard, nous avons commencé la mise en place du dosage des biomarqueurs (en particulier les vitamines et les antioxydants types caroténoïdes) impliqués dans le score d'infertilité inexplicée. Une première étape a consisté à développer la méthode en LC-MS/MS, de manière à permettre un dosage de l'ensemble de ces biomarqueurs sur une seule et même analyse (ou un « run » unique). Une seconde étape consiste en la transposition de ces dosages à partir d'une goutte de sang séchée.

Explorations métaboliques du SOPK : mécanismes moléculaires

Dans cette thèse, nous nous sommes focalisés sur un point très précis de l'exclusion des NC21OHD en contexte d'hyperandrogénie, et notamment du SOPK. Comme nous avons pu le constater dans ce travail, les mécanismes moléculaires impliqués dans la physiopathologie du SOPK sont nombreux, complexes et interconnectés ⁷⁶⁻⁷⁸. L'utilisation d'autres méthodes d'IA pourrait également constituer un atout dans une perspective plus fondamentale et mécanistique, telle que l'exploration métabolique du SOPK.

En ce sens, le développement de l'analyse de données à multi-blocs est une approche utilisée pour explorer et analyser des ensembles de données complexes et hétérogènes qui sont structurés en plusieurs blocs. L'analyse de ces différents blocs, particulièrement adaptée à l'interprétation des interactions biologiques, proviennent de différentes sources ou modalités ^{79,80}. Un même panel de marqueurs peut être mesuré dans différents organes (analyse dite multi-

organes), pour une analyse à l'échelle de l'organisme, permettant d'identifier des schémas d'expression métabolique organe spécifique. Plusieurs types de marqueurs, provenant de la même matrice, peuvent également être mesurés : on parle alors d'analyse multi-omiques ⁸¹. L'objectif principal de ce type d'analyse est de découvrir des relations, des structures ou des profils sous-jacents qui peuvent exister entre les différentes sources de données. Pour ce faire, des techniques de réduction de dimensionnalité et dérivées, telles que celles appliquées dans cette thèse, sont souvent utilisées.

L'intégration des différentes voies et systèmes biochimiques impliqués dans le SOPK, à l'aide de méthodes multi-blocs, pourrait favoriser une meilleure prise en compte de ces différents éléments. Cette approche intégrative contribuerait à comprendre leurs implications et leurs interactions dans la physiopathologie de la maladie, à identifier des sous-groupes de malades et, en fin de compte, à mieux prévenir et traiter ces patientes.

Conclusion

Notre objectif a été d'explorer et de développer des outils diagnostiques dans des conditions complexes de l'infertilité, comme l'infertilité inexplicée ou le SOPK. Nous avons en effet montré, au travers de deux applications, que l'application d'outils innovants issus de l'intelligence artificielle sur des données hétérogènes permettait de mieux identifier, ou de manière simplifiée, les patientes ou les couples infertiles.

Ces outils pourraient potentiellement redéfinir l'infertilité inexplicée à long terme en remettant en question la notion d'absence de diagnostic et en démontrant l'existence d'une composante métabolique importante, réduisant ainsi progressivement la part des situations inexplicées. Par ailleurs, la nature hétérogène du SOPK imposera l'application de méthodes de modélisation multi-blocs, sophistiquées, en vue d'une meilleure identification des sous-groupes de patientes et de la prise en compte des comorbidités.

Enfin, l'intégration de ces nouvelles méthodes d'intelligence artificielle devra passer par un changement de sémiologie, mais permettra, à terme, d'améliorer la prise en charge des patients à travers une médecine plus prédictive, préventive et personnalisée. L'intégration du numérique et plus largement des nouvelles technologies en santé ne pourra se faire que par l'établissement d'un consensus et d'une approche collaborative transdisciplinaire, pour garantir la maîtrise totale de toutes les étapes du processus de développement. Sur le plan réglementaire, les pays s'organisent pour proposer et mettre en place un cadre législatif adaptée aux contraintes et à la rapidité d'évolution caractéristique de ces technologies. Pour finir, une réflexion éthique s'impose : les professionnels de santé devront réfléchir aux implications éthiques de l'utilisation de l'IA dans leur pratique, notamment en ce qui concerne la confidentialité des données, la responsabilité et l'équité dans l'accès aux soins de santé.

Bibliographie

1. World Health Organization; 2023. Infertility Prevalence Estimates, 1990–2021.
2. Samir HAMAMAH, Salomé BERLIOUX, Comité de pilotage. Rapport sur les causes d'infertilité - Vers une stratégie nationale de lutte contre l'infertilité. Published online February 21, 2022.
3. Levine H, Jørgensen N, Martino-Andrade A, et al. Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. *Hum Reprod Update*. 2023;29(2):157-176. doi:10.1093/humupd/dmac035
4. Hoek J, Schoenmakers S, Baart EB, et al. Preconceptional Maternal Vegetable Intake and Paternal Smoking Are Associated with Pre-implantation Embryo Quality. *Reprod Sci*. 2020;27(11):2018-2028. doi:10.1007/s43032-020-00220-8
5. Hojeij B, Rousian M, Sinclair KD, Dinnyes A, Steegers-Theunissen RPM, Schoenmakers S. Periconceptional biomarkers for maternal obesity: a systematic review. *Rev Endocr Metab Disord*. 2023;24(2):139-175. doi:10.1007/s11154-022-09762-5
6. Husen SC, Kemper NDHE, Go ATJI, Willemsen SP, Rousian M, Steegers-Theunissen RPM. Periconceptional maternal folate status and the impact on embryonic head and brain structures: the Rotterdam Periconceptional Cohort. *Reprod Biomed Online*. 2022;44(3):515-523. doi:10.1016/j.rbmo.2021.11.002
7. Van Zundert SKM, Van Rossem L, Willemsen SP, Van Der Meer L, Ernst-Smelt HE, Steegers-Theunissen RPM. Periconceptional maternal social, lifestyle and medical risk factors impair embryonic growth: The Rotterdam Periconceptional Cohort. *Reprod Biomed Online*. 2022;44(6):1123-1133. doi:10.1016/j.rbmo.2022.02.011
8. Thonneau P, Marchand S, Tallec A, et al. Incidence and main causes of infertility in a resident population (1 850 000) of three French regions (1988–1989)*. *Hum Reprod*. 1991;6(6):811-816. doi:10.1093/oxfordjournals.humrep.a137433
9. Carson SA, Kallen AN. Diagnosis and Management of Infertility: A Review. *JAMA*. 2021;326(1):65. doi:10.1001/jama.2021.4788
10. Buckett W, Sierra S. The management of unexplained infertility: an evidence-based guideline from the Canadian Fertility and Andrology Society. *Reprod Biomed Online*. 2019;39(4):633-640. doi:10.1016/j.rbmo.2019.05.023
11. Penzias A, Bendikson K, Falcone T, et al. Evidence-based treatments for couples with unexplained infertility: a guideline. *Fertil Steril*. 2020;113(2):305-322. doi:10.1016/j.fertnstert.2019.10.014
12. The Guideline Group on Unexplained Infertility, Romualdi D, Ata B, et al. Evidence-based guideline: unexplained infertility. *Hum Reprod*. 2023;38(10):1881-1890. doi:10.1093/humrep/dead150

13. Zegers-Hochschild F, Adamson GD, Dyer S, et al. The International Glossary on Infertility and Fertility Care, 2017. *Fertil Steril*. 2017;108(3):393-406. doi:10.1016/j.fertnstert.2017.06.005
14. Raperport C, Desai J, Qureshi D, et al. The definition of unexplained infertility: A systematic review. *BJOG Int J Obstet Gynaecol*. Published online November 13, 2023;1471-0528.17697. doi:10.1111/1471-0528.17697
15. Teede HJ, Tay CT, Laven JJE, et al. Recommendations From the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. *J Clin Endocrinol Metab*. 2023;108(10):2447-2469. doi:10.1210/clinem/dgad463
16. Rolland L, Amar-Hoffet A, Lubin V, Préaubert L, Miquel L, Courbiere B. Unexplained infertility: A French national survey of clinical practices. *J Gynecol Obstet Hum Reprod*. 2022;51(3):102301. doi:10.1016/j.jogoh.2021.102301
17. Munro MG, Critchley HOD, Fraser IS, the FIGO Menstrual Disorders Committee. The two FIGO systems for normal and abnormal uterine bleeding symptoms and classification of causes of abnormal uterine bleeding in the reproductive years: 2018 revisions. *Int J Gynecol Obstet*. 2018;143(3):393-408. doi:10.1002/ijgo.12666
18. Broeze KA, Opmeer BC, Van Geloven N, et al. Are patient characteristics associated with the accuracy of hysterosalpingography in diagnosing tubal pathology? An individual patient data meta-analysis. *Hum Reprod Update*. 2011;17(3):293-300. doi:10.1093/humupd/dmq056
19. Wang Y, Qian L. Three- or four-dimensional hysterosalpingo contrast sonography for diagnosing tubal patency in infertile females: a systematic review with meta-analysis. *Br J Radiol*. 2016;89(1063):20151013. doi:10.1259/bjr.20151013
20. Alcázar JL, Martinez A, Duarte M, et al. Two-dimensional hysterosalpingo-contrast-sonography compared to three/four-dimensional hysterosalpingo-contrast-sonography for the assessment of tubal occlusion in women with infertility/subfertility: a systematic review with meta-analysis. *Hum Fertil*. 2022;25(1):43-55. doi:10.1080/14647273.2020.1769204
21. Jurkovic D, Geipel A, Gruboeck K, Jauniaux E, Natucci M, Campbell S. Three-dimensional ultrasound for the assessment of uterine anatomy and detection of congenital anomalies: a comparison with hysterosalpingography and two-dimensional sonography. *Ultrasound Obstet Gynecol*. 1995;5(4):233-237. doi:10.1046/j.1469-0705.1995.05040233.x
22. Caliskan E, Ozkan S, Cakiroglu Y, Sarisoy HT, Corakci A, Ozeren S. Diagnostic accuracy of real-time 3D sonography in the diagnosis of congenital Mullerian anomalies in high-risk patients with respect to the phase of the menstrual cycle. *J Clin Ultrasound*. 2010;38(3):123-127. doi:10.1002/jcu.20662
23. Ludwin A, Pityński K, Ludwin I, Banas T, Knafel A. Two- and Three-Dimensional Ultrasonography and Sonohysterography versus Hysteroscopy With Laparoscopy in the Differential Diagnosis of Septate, Bicornuate, and Arcuate Uteri. *J Minim Invasive Gynecol*. 2013;20(1):90-99. doi:10.1016/j.jmig.2012.09.011

24. World Health Organization. WHO laboratory manual for the examination and processing of human semen, sixth edition. Published online 2021.
25. O'Neill CL, Parrella A, Keating D, Cheung S, Rosenwaks Z, Palermo GD. A treatment algorithm for couples with unexplained infertility based on sperm chromatin assessment. *J Assist Reprod Genet.* 2018;35(10):1911-1917. doi:10.1007/s10815-018-1270-x
26. Borges E, Zanetti BF, Setti AS, Braga DPDAF, Provenza RR, Iaconelli A. Sperm DNA fragmentation is correlated with poor embryo development, lower implantation rate, and higher miscarriage rate in reproductive cycles of non-male factor infertility. *Fertil Steril.* 2019;112(3):483-490. doi:10.1016/j.fertnstert.2019.04.029
27. Repalle D, Saritha KV, Bhandari S. Sperm DNA fragmentation negatively influences the cumulative live birth rate in the intracytoplasmic sperm injection cycles of couples with unexplained infertility. *Clin Exp Reprod Med.* 2022;49(3):185-195. doi:10.5653/cerm.2021.05169
28. Veena BS, Upadhya S, Adiga SK, Pratap KN. Evaluation of oxidative stress, antioxidants and prolactin in infertile women. *Indian J Clin Biochem.* 2008;23(2):186-190. doi:10.1007/s12291-008-0041-3
29. Pekel A, Gönenç A, Turhan NÖ, Kafalı H. Changes of sFas and sFasL, oxidative stress markers in serum and follicular fluid of patients undergoing IVF. *J Assist Reprod Genet.* 2015;32(2):233-241. doi:10.1007/s10815-014-0396-8
30. Lazzarino G, Pallisco R, Bilotta G, et al. Altered Follicular Fluid Metabolic Pattern Correlates with Female Infertility and Outcome Measures of In Vitro Fertilization. *Int J Mol Sci.* 2021;22(16):8735. doi:10.3390/ijms22168735
31. Şentürk R, Tola EN, Bozkurt M, Doğuç DK. The role of oxidant status on the etiopathogenesis of unexplained infertility and intracytoplasmic sperm injection – embryo transfer success: a case-control study. *J Obstet Gynaecol.* 2022;42(5):1312-1318. doi:10.1080/01443615.2021.1960294
32. Wang R, Danhof NA, Tjon-Kon-Fat RI, et al. Interventions for unexplained infertility: a systematic review and network meta-analysis. Cochrane Gynaecology and Fertility Group, ed. *Cochrane Database Syst Rev.* 2019;2019(9). doi:10.1002/14651858.CD012692.pub2
33. Duffy JMN, Adamson GD, Benson E, et al. Top 10 priorities for future infertility research: an international consensus development study. *Hum Reprod.* 2020;35(12):2715-2724. doi:10.1093/humrep/deaa242
34. Lai S, Wang R, Van Wely M, et al. IVF versus IUI with ovarian stimulation for unexplained infertility: a collaborative individual participant data meta-analysis. *Hum Reprod Update.* 2024;30(2):174-185. doi:10.1093/humupd/dmad033
35. El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. *The Lancet.* 2017;390(10108):2194-2210. doi:10.1016/S0140-6736(17)31431-9
36. Pugeat M, Déchaud H, Raverot V, Denuzière A, Cohen R, Boudou P. Recommandations pour l'exploration des hyperandrogénies. Published online 2010:30.

37. Young J, Tardy V, Brac de la Perrière A, Bachelot A, Morel Y. Dépistage et prise en charge du déficit en 21-hydroxylase à révélation tardive chez les femmes avec hyperandrogénie. *Ann Endocrinol.* 2010;71(1):e15-e19. doi:10.1016/j.ando.2009.12.010
38. Azziz R, Carmina E, Chen Z, et al. Polycystic ovary syndrome. *Nat Rev Dis Primer.* 2016;2(1):16057. doi:10.1038/nrdp.2016.57
39. Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med.* 2010;8(1):41. doi:10.1186/1741-7015-8-41
40. Practice Committee of the American Society for Reproductive Medicine. Diagnostic evaluation of the infertile female: a committee opinion. *Fertil Steril.* 2015;103(6):e44-e50. doi:10.1016/j.fertnstert.2015.03.019
41. Turcu AF, Rege J, Auchus RJ, Rainey WE. 11-Oxygenated androgens in health and disease. *Nat Rev Endocrinol.* 2020;16(5):284-296. doi:10.1038/s41574-020-0336-x
42. Storbeck KH, O'Reilly MW. The clinical and biochemical significance of 11-oxygenated androgens in human health and disease. *Eur J Endocrinol.* 2023;188(4):R98-R109. doi:10.1093/ejendo/lvad047
43. Foucaut AM, Faure C, Julia C, Czernichow S, Levy R, Dupont C. Sedentary behavior, physical inactivity and body composition in relation to idiopathic infertility among men and women. *PLoS One.* 2019;14(4):e0210770. doi:10.1371/journal.pone.0210770
44. Dupont C, Faure C, Daoud F, Gautier B, Czernichow S, Lévy R. Metabolic syndrome and smoking are independent risk factors of male idiopathic infertility. *Basic Clin Androl.* 2019;29:9. doi:10.1186/s12610-019-0090-x
45. Schiffer L, Barnard L, Baranowski ES, et al. Human steroid biosynthesis, metabolism and excretion are differentially reflected by serum and urine steroid metabolomes: A comprehensive review. *J Steroid Biochem Mol Biol.* 2019;194:105439. doi:10.1016/j.jsbmb.2019.105439
46. Fiet J, Le Bouc Y, Guéchet J, et al. A Liquid Chromatography/Tandem Mass Spectrometry Profile of 16 Serum Steroids, Including 21-Deoxycortisol and 21-Deoxycorticosterone, for Management of Congenital Adrenal Hyperplasia. *J Endocr Soc.* 2017;1(3):186-201. doi:10.1210/js.2016-1048
47. Babu M, Snyder M. Multi-Omics Profiling for Health. *Mol Cell Proteomics.* 2023;22(6):100561. doi:10.1016/j.mcpro.2023.100561
48. Karczewski KJ, Snyder MP. Integrative omics for health and disease. *Nat Rev Genet.* 2018;19(5):299-310. doi:10.1038/nrg.2018.4
49. Rutledge J, Oh H, Wyss-Coray T. Measuring biological age using omics data. *Nat Rev Genet.* 2022;23(12):715-727. doi:10.1038/s41576-022-00511-7
50. Tessier AJ, Wang F, Liang L, et al. Plasma metabolites of a healthy lifestyle in relation to mortality and longevity: Four prospective US cohort studies. *Med.* 2024;5(3):224-238.e5. doi:10.1016/j.medj.2024.01.010

51. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*. 1989;246(4926):64-71. doi:10.1126/science.2675315
52. Mellon FA. MASS SPECTROMETRY | Principles and Instrumentation. In: *Encyclopedia of Food Sciences and Nutrition*. Elsevier; 2003:3739-3749. doi:10.1016/B0-12-227055-X/00746-X
53. Yost RA, Enke CG. Selected ion fragmentation with a tandem quadrupole mass spectrometer. *J Am Chem Soc*. 1978;100(7):2274-2275. doi:10.1021/ja00475a072
54. Bachelot G, Rivet-Danon D, Sermondade N, et al. Artificial intelligence: toward a better strategy for predicting testicular biopsy results in the context of azoospermia. *Ann Biol Clin (Paris)*. Published online December 2023. doi:doi:10.1684/abc.2024.1882
55. Tenenhaus M. *La Régression PLS: Théorie et Pratique*. Editions Technip; 1998.
56. WOLD H. Estimation of principal components and related models by iterative least squares. *Multivar Anal*. Published online 1966:391-420.
57. Wold H. Nonlinear iterative partial least squares (NIPALS) modelling: some current developments. In: *Multivariate Analysis—III*. Elsevier; 1973:383-407.
58. Pérez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Hum Genet*. 2003;112(5-6):581-592. doi:10.1007/s00439-003-0921-9
59. Farkouh A, Salvio G, Kuroda S, Saleh R, Vogiatzi P, Agarwal A. Sperm DNA integrity and male infertility: a narrative review and guide for the reproductive physicians. *Transl Androl Urol*. 2022;11(7):1023-1044. doi:10.21037/tau-22-149
60. Agarwal A, Majzoub A, Baskaran S, et al. Sperm DNA Fragmentation: A New Guideline for Clinicians. *World J Mens Health*. 2020;38(4):412. doi:10.5534/wjmh.200128
61. Santi D, Spaggiari G, Simoni M. Sperm DNA fragmentation index as a promising predictive tool for male infertility diagnosis and treatment management – meta-analyses. *Reprod Biomed Online*. 2018;37(3):315-326. doi:10.1016/j.rbmo.2018.06.023
62. Aitken RJ. DNA damage in human spermatozoa; important contributor to mutagenesis in the offspring. *Transl Androl Urol*. 2017;6(S4):S761-S764. doi:10.21037/tau.2017.09.13
63. Aitken RJ, Gibb Z, Baker MA, Drevet J, Gharagozloo P. Causes and consequences of oxidative stress in spermatozoa. *Reprod Fertil Dev*. 2016;28(2):1. doi:10.1071/RD15325
64. Aitken RJ, Curry BJ. Redox Regulation of Human Sperm Function: From the Physiological Control of Sperm Capacitation to the Etiology of Infertility and DNA Damage in the Germ Line. *Antioxid Redox Signal*. 2011;14(3):367-381. doi:10.1089/ars.2010.3186
65. Henkel R, Kierspel E, Stalf T, et al. Effect of reactive oxygen species produced by spermatozoa and leukocytes on sperm functions in non-leukocytospermic patients. *Fertil Steril*. 2005;83(3):635-642. doi:10.1016/j.fertnstert.2004.11.022

66. Mahfouz RZ, Du Plessis SS, Aziz N, Sharma R, Sabanegh E, Agarwal A. Sperm viability, apoptosis, and intracellular reactive oxygen species levels in human spermatozoa before and after induction of oxidative stress. *Fertil Steril*. 2010;93(3):814-821. doi:10.1016/j.fertnstert.2008.10.068
67. Homa S, Vassiliou A, Stone J, et al. A Comparison Between Two Assays for Measuring Seminal Oxidative Stress and their Relationship with Sperm DNA Fragmentation and Semen Parameters. *Genes*. 2019;10(3):236. doi:10.3390/genes10030236
68. Pearce KL, Hill A, Tremellen KP. Obesity related metabolic endotoxemia is associated with oxidative stress and impaired sperm DNA integrity. *Basic Clin Androl*. 2019;29(1):6. doi:10.1186/s12610-019-0087-5
69. Oleszczuk K, Augustinsson L, Bayat N, Giwercman A, Bungum M. Prevalence of high DNA fragmentation index in male partners of unexplained infertile couples. *Andrology*. 2013;1(3):357-360. doi:10.1111/j.2047-2927.2012.00041.x
70. Speiser PW, Azziz R, Baskin LS, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2010;95(9):4133-4160. doi:10.1210/jc.2009-2631
71. Bidet M, Bellanné-Chantelot C, Galand-Portier MB, et al. Fertility in women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *J Clin Endocrinol Metab*. 2010;95(3):1182-1190. doi:10.1210/jc.2009-1383
72. Bidet M, Bellanné-Chantelot C, Galand-Portier MB, et al. Clinical and molecular characterization of a cohort of 161 unrelated women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency and 330 family members. *J Clin Endocrinol Metab*. 2009;94(5):1570-1578. doi:10.1210/jc.2008-1582
73. Agnani H, Bachelot G, Eguether T, et al. A proof of concept of a machine learning algorithm to predict late-onset 21-hydroxylase deficiency in children with premature pubic hair. *J Steroid Biochem Mol Biol*. 2022;220:106085. doi:10.1016/j.jsbmb.2022.106085
74. Ueland GÅ, Dahl SR, Methlie P, Hessen S, Husebye ES, Thorsby PM. Adrenal steroid profiling as a diagnostic tool to differentiate polycystic ovary syndrome from nonclassic congenital adrenal hyperplasia: pinpointing easy screening possibilities and normal cutoff levels using liquid chromatography tandem mass spectrometry. *Fertil Steril*. 2022;118(2):384-391. doi:10.1016/j.fertnstert.2022.05.012
75. Fiet J, Villette JM, Galons H, et al. The application of a new highly-sensitive radioimmunoassay for plasma 21-deoxycortisol to the detection of steroid-21-hydroxylase deficiency. *Ann Clin Biochem*. 1994;31 (Pt 1):56-64. doi:10.1177/000456329403100110
76. Rodriguez Paris V, Bertoldo MJ. The Mechanism of Androgen Actions in PCOS Etiology. *Med Sci*. 2019;7(9):89. doi:10.3390/medsci7090089
77. Kempegowda P, Melson E, Manolopoulos KN, Arlt W, O'Reilly MW. Implicating androgen excess in propagating metabolic disease in polycystic ovary syndrome. *Ther Adv Endocrinol Metab*. 2020;11:204201882093431. doi:10.1177/2042018820934319

78. Michalakis K, Mintziori G, Kaprara A, Tarlatzis BC, Goulis DG. The complex interaction between obesity, metabolic syndrome and reproductive axis: a narrative review. *Metabolism*. 2013;62(4):457-478. doi:10.1016/j.metabol.2012.08.012
79. Lejeune FX, Ichou F, Camenen E, et al. A Multimodal Omics Exploration of the Motor and Non-Motor Symptoms of Parkinson's Disease. *Int J Transl Med*. 2022;2(1):97-112. doi:10.3390/ijtm2010009
80. Xicota L, Ichou F, Lejeune FX, et al. Multi-omics signature of brain amyloid deposition in asymptomatic individuals at-risk for Alzheimer's disease: The INSIGHT-preAD study. *EBioMedicine*. 2019;47:518-528. doi:10.1016/j.ebiom.2019.08.051
81. Reinke SN, Galindo-Prieto B, Skotare T, et al. OnPLS-Based Multi-Block Data Integration: A Multivariate Approach to Interrogating Biological Interactions in Asthma. *Anal Chem*. 2018;90(22):13400-13408. doi:10.1021/acs.analchem.8b03205

Annexe

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Steroid Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/jsbmb

A proof of concept of a machine learning algorithm to predict late-onset 21-hydroxylase deficiency in children with premature pubic hair

Hélène Agnani^{a,c}, Guillaume Bachelot^b, Thibaut Eguether^b, Bettina Ribault^b, Jean Fiet^b, Yves Le Bouc^{a,c}, Irène Netchine^{a,c}, Muriel Houang^{a,c,1}, Antonin Lamazière^{b,*,1}

^a Sorbonne Université, Saint Antoine Research center, CRSA, INSERM, Sorbonne Université, Paris, France

^b Sorbonne Université, Saint Antoine Research center, CRSA, INSERM, Département de Métabolisme Clinique, Hôpital Saint Antoine, AP-HP/Sorbonne Université, Paris, France

^c Explorations Fonctionnelles Endocriniennes, Hôpital Armand Trousseau, AP-HP, 26 Av Dr Netter, Paris 75012, France

ARTICLE INFO

Keywords:

Late-onset 21-hydroxylase deficiency
Premature pubarche
Steroid profile
Machine learning
Mass spectrometry

ABSTRACT

In children with premature pubarche (PP), late onset 21-hydroxylase deficiency (21-OHD), also known as non-classical congenital adrenal hyperplasia (NCCAH), can be routinely ruled out by an adrenocorticotrophic hormone (ACTH) test. Using liquid chromatography–tandem mass spectrometry (LC-MS/MS), a quantitative assay of the circulating steroidome can be obtained from a single blood sample. We hypothesized that, by applying multivariate machine learning (ML) models to basal steroid profiles and clinical parameters of 97 patients, we could distinguish children with PP from those with NCCAH, without the need for ACTH testing. Every child presenting with PP at the Trousseau Pediatric Endocrinology Unit between 2016 and 2018 had a basal and stimulated steroidome. Patients with central precocious puberty were excluded. The first set of patients (year 1, training set, $n = 58$), including 8 children with NCCAH verified by ACTH test and genetic analysis, was used to train the model. Subsequently, a validation set of an additional set of patients (year 2, $n = 39$ with 5 NCCAH) was obtained to validate our model. We designed a score based on an ML approach (orthogonal partial least squares discriminant analysis). A metabolic footprint was assigned for each patient using clinical data, bone age, and adrenal steroid levels recorded by LC-MS/MS. Supervised multivariate analysis of the training set (year 1) and validation set (year 2) was used to validate our score. Based on selected variables, the prediction score was accurate (100%) at differentiating premature pubarche from late onset 21-OHD patients. The most significant variables were 21-deoxycorticosterone, 17-hydroxyprogesterone, and 21-deoxycortisol steroids. We proposed a new test that has excellent sensitivity and specificity for the diagnosis of NCCAH, due to an ML approach.

1. Introduction

Premature pubarche (PP), defined as the development of pubic hair before the age of 8 for girls and 9 for boys, is a common feature in children and is usually due to physiological adrenarche [1]. In individuals with PP, it is necessary, however, to rule out an adrenal pathology, especially late onset congenital adrenal hyperplasia (NCCAH), which is due to steroid 21-hydroxylase deficiency (21-OHD).

NCCAH, one of the most common genetic autosomal recessive disorders, results from molecular abnormalities of the *CYP21A2* gene. It occurs in 0.1% of the general population [2], with some variation between ethnic groups [3]. In NCCAH, residual enzyme activity of about 50–70% [4] leads to less severe features than classical neonatal congenital adrenal hyperplasia (CAH). The accumulation of androgens upstream of the enzyme defect [5] is responsible for hirsutism, acne [6], accelerated bone maturation [7], and short final height [8]. The

Abbreviations: 11OHA4, 11 β -hydroxyandrostenedione; 11-DF, 11-deoxycortisol; 17-OHP, 17 α -hydroxyprogesterone; 17-OHPreg, 17 α -hydroxypregnenolone; 21-DB, 21-deoxycorticosterone; 21-DF, 21-deoxycortisol; ACTH, Adreno corticotrophic hormone; AI, Adrenal insufficiency; ALDO, Aldosterone; B, Corticosterone; BA, Bone age; BMI, Body mass index; A4, Androstenedione; DHEA, Dehydroepiandrosterone; DOC, 11-deoxycorticosterone; E, Cortisone; F, Cortisol; GV, Growth velocity; LC-MS/MS, Liquid chromatography tandem mass spectrometry; NCCAH, Late-onset 21-hydroxylase deficiency; P, Progesterone; Preg, Pregnenolone; T, Testosterone; ML, Machine Learning.

* Corresponding author.

E-mail address: antonin.lamaziere@aphp.fr (A. Lamazière).

¹ The two authors equally contribute to this work

<https://doi.org/10.1016/j.jsbmb.2022.106085>

Received 18 June 2021; Received in revised form 27 February 2022; Accepted 1 March 2022

Available online 12 March 2022

0960-0760/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

phenotype of NCCAH is determined by the less severe mutation of *CYP21A2* [9]. However, molecular diagnosis is mandatory for familial planning [10].

Based on the current consensus [11], adrenocorticotrophic hormone (ACTH) testing is performed on a routine basis for all children with PP. NCCAH is currently defined by a post-ACTH test level of 17-hydroxyprogesterone (17-OHP) in the serum greater than 10 ng/ml [12]. After parental consent (according to the Declaration of Helsinki), these results must be confirmed by a genetic test of the *CYP21A2* gene. The use of mass spectrometry for multidimensional steroid profiling in 21-OHD is well established, starting with use of gas chromatography–mass spectrometry in the 1980 s where urinary steroid profiles were useful, especially in newborns [13,14]. Furthermore, with the development of liquid chromatography–tandem mass spectrometry methods (LC-MS/MS), the diagnosis of adrenal diseases [15,16] and especially of CAH [17,18] has improved specificity for steroid analysis when compared with immunoassays [19], because these methods are free from cross reactions of 17-OHP with others steroids [20] and can reduce false positive results. Spectrometry methods can also reduce the number of unnecessary tests and ease the anxiety of the patients [21]. With LC-MS/MS, a circulating steroidome quantitative assay [22], including more than 16 molecular species, is now available in a single analysis.

We hypothesized that combining patient anthropometric data with the analytical performances of mass spectrometry and the use of multivariate statistical approaches, such as machine learning algorithms, could lead to the development of a new practical diagnostic tool for clinical routine analysis. The aim of the present study was to generate a score based on multivariate training to easily and accurately distinguish children with PP from those with NCCAH. Routine ACTH tests would therefore be unnecessary and limited to ruling out adrenal insufficiency (AI).

2. Patients and methods

2.1. Patients

The exploration of leftover samples of serum of all children who had routine ACTH tests to screen for NCCAH at the Trousseau Pediatric Endocrinology Unit from June 2016 to July 2018 was extended to 16

steroid profile analyses. All patients came to the endocrine unit because of premature pubic hair, to rule out either adrenal disease or precocious puberty. Only children with early pubic hair (before 8 years in girls, and before 9 years in boys) were included in the study. Children with central precocious puberty (CPP) (> Tanner stage 1 associated with an increase in luteizing hormone LH > 5 after gonadotropin-releasing hormone (GnRH) testing) were excluded. One younger patient (1 year and 3 months old) without premature pubic hair was included because of high basal 17-OHP levels in an unrelated context of cryptorchidism; one patient was older than the others (12 years and 10 months old) and had been tested because of Prader-Willi syndrome-associated PP [23].

The first set of patients (year 1, n = 58), recruited from June 2016 to July 2017, was used to train the model. Subsequently, a validation set (year 2) of an additional group of patients (n = 39) was obtained to validate our model (Fig. 1). The latter had the same inclusion and exclusion criteria as patients in the training set. The inclusion of patients in the validation set started in November 2017 and ended in July 2018. The study was conducted in accordance with the Declaration of Helsinki.

The model was first trained and tuned (feature selection) on a training set of children (n = 63 with exclusion of 5 boys with CPP) from a pediatric endocrinology unit (year 1) and internally evaluated using cross-validation. A validation set of children (year 2), recruited consecutively to the training set from the same clinical institution, was used for external evaluation (n = 39). The results were matched and labeled according to the expected diagnosis output.

2.2. ACTH test

All patients underwent ACTH testing at 8:30 AM, after a night of fasting. Physical examination was performed by a pediatric endocrinologist. The data collected included height in standard deviation score (SDS), growth velocity in SDS, body mass index (BMI) in Z-score, pubertal status as defined by Tanner, and bone age (BA) increment (BA–chronological age). Growth curves were standardized according to Sempé [24]. BMI was analyzed according to Rolland-Cachera standards [25]. The BA was determined according to the method of Greulich and Pyle [26]. When the patients had breast development or increased testicular volumes, they also had a GnRH test. Steroids were assayed at baseline, then stimulated by ACTH and assayed by LC-MS/MS as

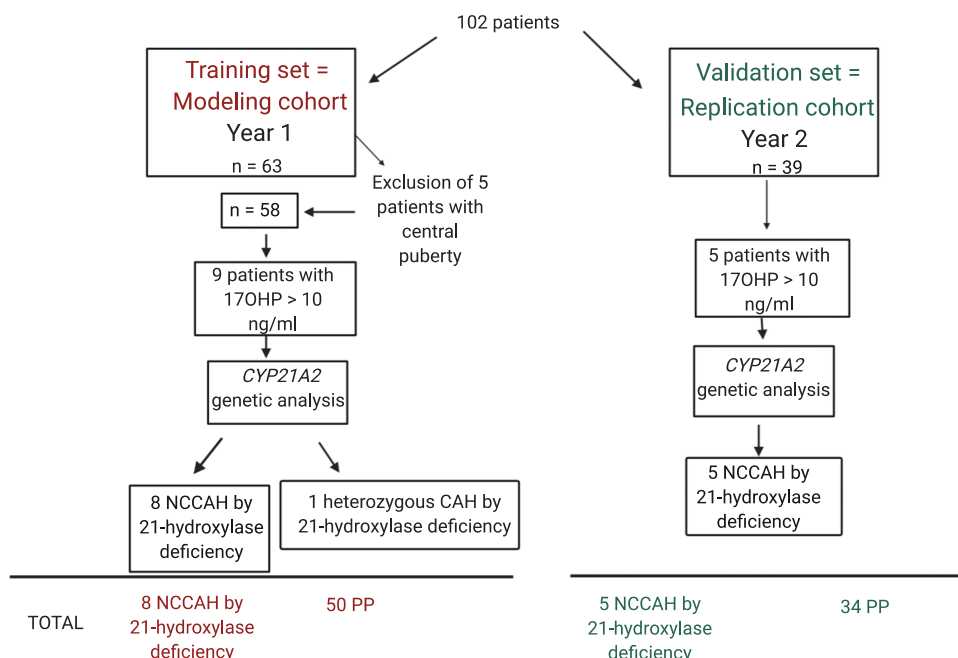


Fig. 1. Data flow diagram for the study (17-OHP > 10 ng/ml corresponds to the 17-OHP stimulated concentration).

previously described [22]. Informed consent for *CYP21A2* gene analysis was obtained from the patients themselves and their parents when 17-OHP levels were above 10 ng/ml, as approved by our local ethics committee. Following the criteria defined by Kuttenn et al. [12], patients with 17-OHP concentrations above 10 ng/ml after ACTH test and bearing abnormalities in *CYP21A2* gene on different alleles were defined as NCCAH. In our cohort, all NCCAH patients presented a 17-OHP concentration above 22 ng/ml after ACTH testing. All PP patients exhibited a 17-OHP concentration below 5 ng/ml after ACTH test (except one heterozygous patient with a 9 ng/ml peak 17-OHP).

2.3. LC-MS/MS steroid profiling

A panel of 16 steroids was analyzed before ACTH test (basal) and 60 min after 250 µg ACTH injection (peak) in the Department of Clinical Metabolomics at the Saint Antoine Hospital [22]. The 16 steroids were 11-deoxycortisol (11-DF), 21-deoxycortisol (21-DF), 11β-hydroxyandrostenedione (11OHA4), androstenedione (A4), testosterone (T), 17α-hydroxyprogesterone (17-OHP), 17α-hydroxypregnenolone (17-OHPreg), dehydroepiandrosterone (DHEA), aldosterone (ALDO), cortisone (E), cortisol (F), corticosterone (B), 11-deoxycorticosterone (DOC), 21-deoxycorticosterone (21-DB), progesterone (P) and pregnenolone (Preg). Steroids were measured in serum by LC-MS/MS as described elsewhere [22].

Briefly, a mixture of the deuterated internal standard (150 µL) was added to 50–100 µL of serum. The solution was mixed and left standing for 5 min, then loaded into an Isolute SLE + 0.4 ml cartridge (Biotage, Uppsala, Sweden). The samples were allowed to adsorb for 5 min before elution of the steroids through the addition of 2 × 0.9 ml methylene chloride. The eluates, which contained the non-conjugated steroids, were evaporated until dry and reconstituted to 150 µL in methanol/water (50/50, volume-to-volume ratio). Steroids were chromatographically separated by high-performance liquid chromatography using a Shimadzu Nexera XR system (Shimadzu France, Marne la Vallee, France) and a Coreshell C18 column (Kinetex, 2.6 µm 100 Å 100 × 2.1 mm; Phenomenex, Le Pecq, France). Detection was performed using a triple quadrupole mass spectrometer (Triple Quad 6500, ABSciex, Foster City, CA). Upon collection, the LC-MS/MS data were analyzed using MultiQuant software (ABSciex, Foster City, CA, version 3.0) with built-in queries and quality control rules that allowed for compound-specific criteria for flagging outlier results. These flagging criteria included accuracies for standards and quality control, quantifier ion/qualifier ion ratios, and lower/upper calculated concentration limits. For each calibration curve, the regression line used for quantification was calculated using least-squares weighting (1/x).

2.4. Data processing and statistical analysis

Biological and clinical data were combined and subjected to multivariate analyses and were compared between the NCCAH children and the other children (referred to as the PP group) from the training set. Univariate analysis was first performed using Mann–Whitney testing according to the number of observations (statistical significance threshold $p < 0.05$).

The raw data were then loaded into SIMCA 15 software (version 15, Umetrics, Västerbotten, Sweden), using principal component analysis (PCA) and orthogonal partial least square discriminant analysis (OPLS-DA) after standardization (removing the mean and scaling to unit variance). A metabolic footprint was assigned for each patient using clinical data and the baseline adrenal steroid levels were recorded by the LC-MS/MS. The goal was to determine whether basal levels can be as accurate as an ACTH test in predicting NCCAH.

2.5. Unsupervised multivariate analysis of the training set

The purpose of this first step was descriptive; the goal was to provide

an overview of the distribution of the two groups of patients and the variability of the system. PCA, an unsupervised method (dimension reduction), was applied to the data to observe the possible presence of trends and groupings, which were previously not obvious by observing the raw data. PCA allowed the detection of potential outliers as well. PCA is used to preprocess and reduce the dimensionality of datasets, while preserving the original structure and relationships inherent to the dataset. The resulting data were displayed as score plots representing the distribution of the samples in multivariate space.

2.6. Supervised multivariate analysis of the training set

The objective of this next step was to determine if the distribution was significantly different between the NCCAH patients and the PP patients of the training set. A discriminant analysis of variables was performed. The assignment of patients in each group (NCCAH and PP) was made a priori for the construction of the model. OPLS-DA, as a supervised and useful multivariate ml algorithm [27], was applied to the training set. Based on recent studies, the OPLS approach was considered particularly appropriate, given the input data type (multicollinearity) and size of the present study [28].

The developed model can be evaluated using several parameters, such as goodness of fit (R^2), a test of permutation, and the capability to predict (Q^2). Eventually, owing to the validation set, we were able to build a predictive classification score. In order to determine the relevance of the variables and to rank their predictive capacity in the model, we used variable importance in projection (VIP) [29].

3. Results

3.1. Patient's characteristics from the training and validation set

All patients visited the endocrine unit because of precocious pubic hair to exclude adrenal disease or precocious puberty. A total of 97 patients (81 girls and 16 boys) were included in our study – 58 patients in the training set and 39 in the validation set (Fig. 1).

All patients were in Tanner stage 1 (prepubertal stage) (Table 1). Thirteen patients were diagnosed with NCCAH by 21-OHD, confirmed by *CYP21A2* genetic analysis: 8 patients from the training set and 5 patients from the validation set (Table 2). All these patients presented with a V281L mutation; 54% had an additional severe mutation or deletion. 12 of the 13 had partial AI, defined as a post-ACTH cortisol peak < 180 ng/ml. Basal levels of 17-OHP ranged from 3.62 ng/ml to 87.94 ng/ml. Peak 17-OHP levels after ACTH testing ranged from 22.73 ng/ml to 92.57 ng/ml (Table 2).

A total of 58 patients (45 girls and 13 boys) with a mean age of 7.14 years (+/– 2.32) were included in our training set (Fig. 1). Five boys had an increase in testicular volume (> G1) with peak LH > 5 after GnRH test and were not included in the study. Nine patients after ACTH testing had a peak 17-OHP level > 9 ng/ml. *CYP21A2* gene analysis was performed on these patients. A total of eight patients were diagnosed with homozygosity for NCCAH and a diagnosis of heterozygosity for 21-hydroxylase was made in one girl, an 8-year-old with pubic hair appearance at the age of 7 years and 3 months. she had an accelerated

Table 1

Clinical characteristics (means ± SD) of patients from both training and validation sets.

	Training set		Validation set	
	NCCAH	PP	NCCAH	PP
Age (years)	8.2 ± 3.2	7 ± 2.2	7.3 ± 1.7	7.5 ± 1.5
% of girls	50	82	100	91
Body Mass Index (SD)	1.3 ± 1.2	1.6 ± 2.2	2 ± 0.8	1.4 ± 2.1
Bone Age (delta)	3 ± 1.5	1.3 ± 1	2 ± 0.9	1.4 ± 0.9
Growth Velocity (SD)	0.8 ± 1.6	2 ± 2.1	4.8 ± 0.4	1.3 ± 1.5

Table 2

Basal and peak (after ACTH test) steroid concentrations (in ng/ml) of 17-OHP, 21-DF, F, and genotypes of NCCAH patients. Patients 1–8 belonged to the training set and patients 9–13 to the validation set.

	NCCAH patients	Gender	Basal 17 α -hydroxy progesterone	Basal 21-deoxycortisol	Basal cortisol	Peak 17 α -hydroxyprogesterone	Peak 21-deoxycortisol	Peak cortisol	Genotypes
Training set (year 1)	1	M	16.9	4.06	170.00	35.14	7.31	206	V281L/V281L
	2	F	5.16	0.54	86.79	27.41	5.65	128.36	V281L/P453S
	3	F	23.74	3.18	149.34	54.37	12.15	169.49	V281L/IVS2-13C>G
	4	F	28.5	2.17	137.31	59.8	9.7	152.28	V281L/IVS2-13C>G
	5	F	5.13	0.81	61.53	38.5	11.57	97.82	V281L/IVS2-13C>G
	6	M	3.62	0.04	47.88	22.73	3.72	122.81	V281L/V281L
	7	M	87.94	15.38	127.16	92.57	15.07	138.57	V281L/deletion
	8	M	18.76	2.02	83.78	52.75	9.59	113.23	V281L/deletion
Validation set(year 2)	9	F	11.73	2.56	112.2	39.8	9.93	131.64	V281L/[Q318X; A265V]
	10	F	3.91	0.09	57.36	45.11	0.04	161.86	V281L/M383V
	11	F	34.16	7.37	152.67	58.20	15.66	169.29	V281L/P30L
	12	F	8.44	0.81	89.11	51.40	13.77	170.63	V281L/V281L
	13	F	12.42	1.66	79.01	73.36	17.59	135.80	V281L/R356W

growth velocity (10 cm in one year, 5.4 SDS) and a two-year BA advance; 17-OHP basal level was 0.57 ng/ml with a peak at 9.06 ng/ml. Sequencing of the *CYP21A2* gene found a moderate mutation p.H38L in a heterozygous state.

Table 3 compares the basal biological characteristics of the children in the training set. The difference between NCCAH and PP children was significant ($p < 0.05$) for A4, T, 17-OHP, 21-DB, 21-DF, PREG, 11OHA4, and P.

Table 3

Steroid concentrations of the training set (mean and range), $n = 8$ for NCCAH and $n = 50$ for PP. All variables are in basal values. All steroid concentrations are expressed in ng/ml except for DOC, 21-DB and ALDO in pg/ml.

	NCCAH n= 8	PP n=50	P values
Steroids	Mean (range)	Mean (range)	
Aldosterone	165 (49-311)	114 (10-712)	NS
Dehydroepiandrosterone	2.85 (0.1-6.36)	2.15 (0.1-13.83)	NS
Androstenedione	0.84 (0.11-1.28)	0.28 (0.04-0.81)	**
Testosterone	0.2 (0.07-0.44)	0.07 (0.01-0.32)	**
Pregnenolone	1.36 (0.77-2.23)	0.82 (0.04-3.19)	***
11 β -hydroxyandrostenedione	2.05 (0.72-3.73)	0.59 (0.1-1.89)	**
17 α -hydroxypregnenolone	3.26 (0.96-5.71)	1.80 (0.01-15.04)	***
Progesterone	0.88 (0.15-3.4)	0.09 (0.001-0.99)	**
21-deoxycorticosterone	20 (12-31)	5 (5-5)	**
11-deoxycorticosterone	31 (9-71)	54 (6-227)	NS
17 α -hydroxyprogesterone	14.19 (3.62-28.5)	0.28 (0.05-1.03)	**
21-deoxycortisol	1.92 (0.04-4.06)	0.03 (0.001-0.51)	**
Corticosterone	0.67 (0.24-1.01)	2.74 (0.37-26.95)	***
11-deoxycortisol	0.55 (0.23-1.04)	0.46 (0.006-3.07)	NS
Cortisone	26.58 (18.01-35.56)	21.99 (9.48-32.22)	NS
Cortisol	106.1 (47.88-170)	79.97 (26.81-186.11)	NS

NS: p -value > 0.05

** p -value < 0.001

*** p -value {0.001-0.05}

There was no overlap for the 17-OHP and 21-DB values between the two groups in both sets, with significantly increased 17-OHP, 21-DF, and 21-DB levels in the NCCAH children (Fig. 2).

Unsupervised multivariate analysis of the training set was first completed. The descriptive analysis showed a good separation between the NCCAH group and the PP group ($R2X = 0.350$). The variables with poor discrimination were removed from the study, which were growth velocity and BMI.

3.2. Model construction

Secondly, a score plot was obtained from a supervised multivariate analysis of the training set using OPLS-DA (Fig. 3A). The observations were projected in the system's maximum variability plan ($R2X = 0.900$ and $Q2X = 0.868$) on two components. NCCAH patients were segregated on the left side of the score plot (red dots) while PP patients were segregated on the right quadrants (green dots). There was a good separation of the groups without any overlap. The heterozygous patient was segregated in the PP group.

To build a prediction score, the contribution of each variable of our system was calculated (Supplementary Figure 1). The higher the absolute value of the coefficient of a variable, the more discriminatory and important the variable was for prediction and group separation. The coefficients were calculated using the algorithm established by the OPLS-DA regression. As shown in Supplementary Figure 1, the variables 21-DB, 17-OHP, 21-DF, 11OHA4, and P were the top 5 highest contribution coefficients for NCCAH group discrimination followed by BA, the only discriminating clinical factor with a VIP score > 1.

3.3. Validation of the model with the replication cohort

The capability to predict the model was good according to internal cross validation $QX2 = 0.868$. A validation set of 39 patients (36 girls and 3 boys) was used to assess a second step validation. The average age was 7.50 years (± 1.46), and no patient was excluded. Five patients had a peak 17-OHP level > 10 ng/ml and were diagnosed as NCCAH after *CYP21A2* gene analysis (see Table 1, patients 9–13).

To note, by definition, the data from the validation set were not used to refit the model; they were only used to test its accuracy. To mimic future prospective routine diagnosis, observations from the validation

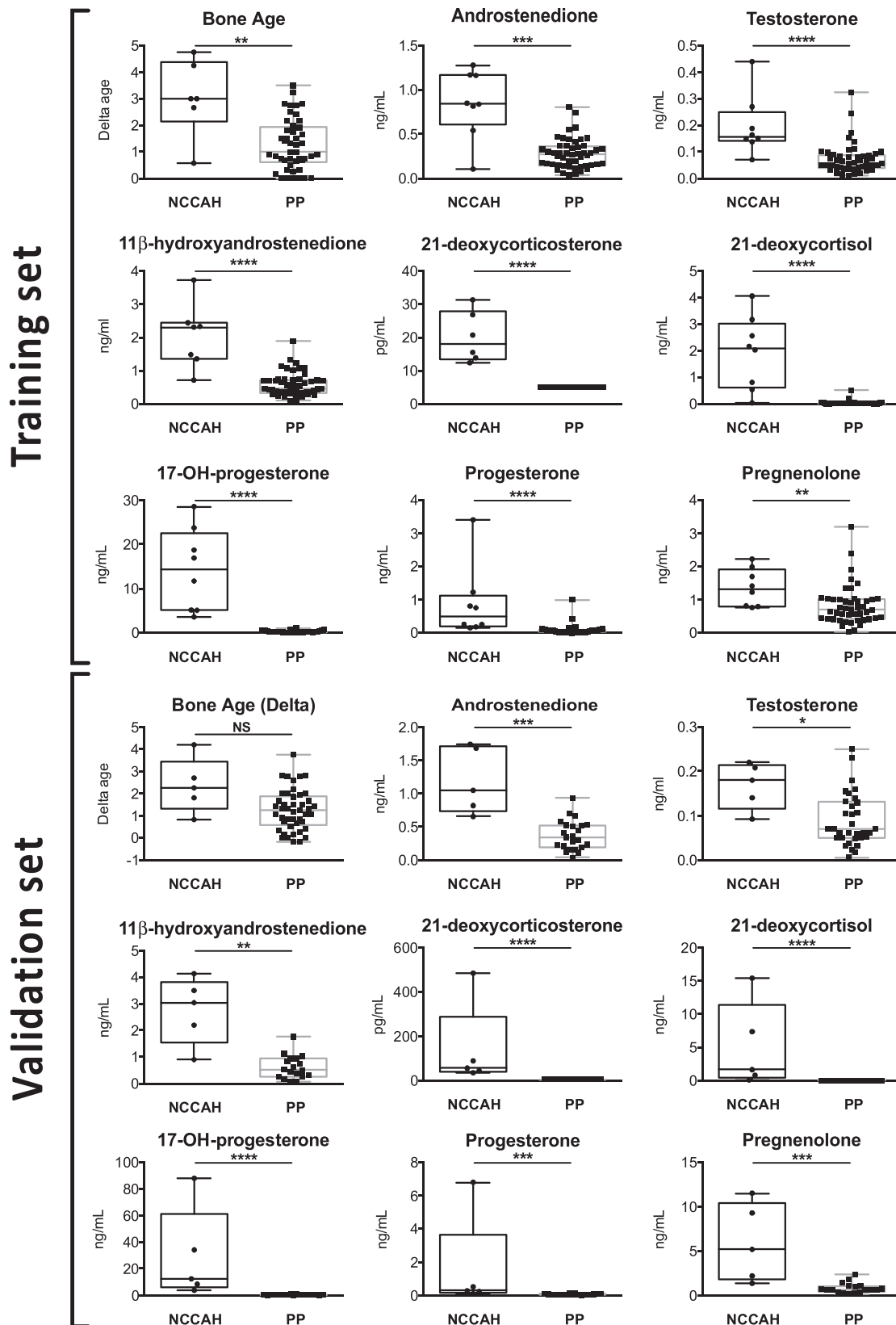


Fig. 2. Box plots of bone age and some circulating steroid concentrations at their basal state from the two groups of patients from the training and the validation set: late onset NCCAH children (n = 8 training set and n = 5 validation set) and the PP children (n = 50 training set and n = 34 validation set). All steroids are expressed in ng/ml except for 21-DB in pg/ml. Bone age is expressed as a difference in years (BA–chronological age).

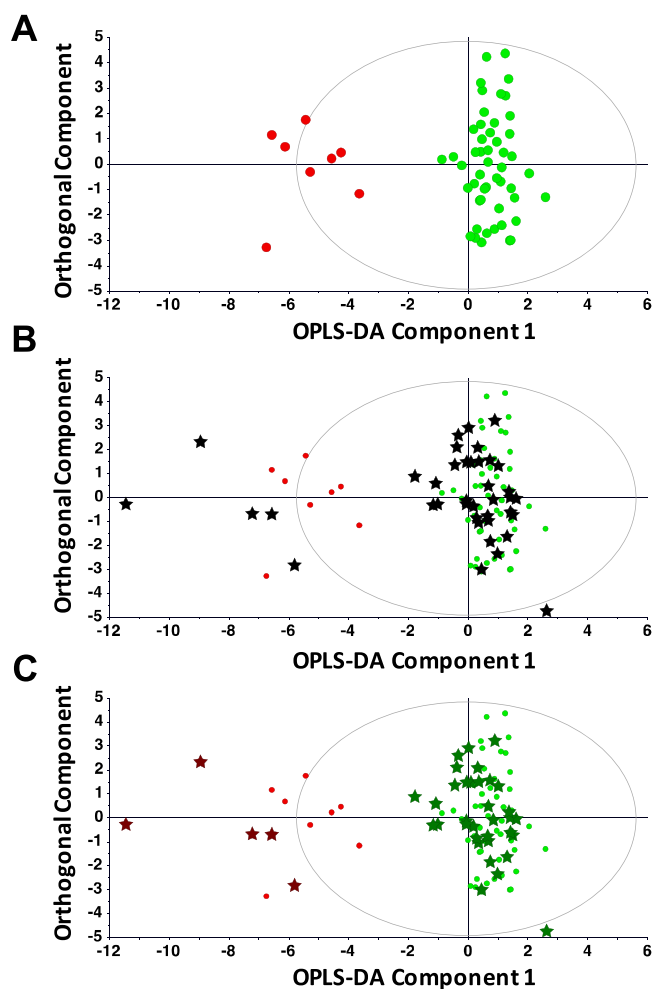


Fig. 3. Score plots from orthogonal partial least squares discriminant analysis (incremental process with model training and external evaluation) of late onset NCCAH and PP children as assessed by their basal circulating signature (16 steroids and BA). Panel A: Score scatter plot plots of the training set generated with a supervised approach from known NCCAH ($n = 8$, red dots) and PP ($n = 50$, green dots) children. R^2 and Q^2 were 0.900 and 0.868 respectively. Panel B: Score plot of the initial training set (NCCAH ($n = 8$, red dots) and PP ($n = 50$, green dots)) superimposed with new children from the validation set ($n = 39$ children) labelled with black dots, indicating no a priori class attribution. Panel C: Score plot of the initial training set (NCCAH ($n = 8$, red dots) and PP ($n = 50$, green dots)) superimposed with the new children from the validation set labeled according to their retrospective final diagnosis: NCCAH ($n = 5$, red stars) and PP ($n = 34$ green stars). Validation accuracy = 1. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

Table 4

Model prediction results. Assignment in each class (NCCAH or PP) of the patients from the training set and the validation set. The predictive results obtained on the training set were obtained using cross-validation.

		Number of Patients	Correct classification rate	NCCAH predicted	PP predicted
Training set	NCCAH	8	100%	8	0
	PP	50	100%	0	50
Validation set	NCCAH ^a	5	100%	5	0
	PP ^a	34	100%	0	34
	Total	97	100%	13	84
	Fisher's prob	5.2e-10			

^a diagnosis made with post-hoc confirmation by CYP21A2 analysis

set were first labelled with black stars, as no a priori class attribution was achieved (Fig. 3B). Eventually, as seen in Fig. 3C, to evaluate the true model performances, the black stars corresponding to children from the validation set were labeled according to their retrospective final diagnosis: NCCAH ($n = 5$, red stars) and PP ($n = 34$ green stars). Five patients of the validation set were segregated in the same area of the score plot as the NCCAH patients based on the training set. These 5 patients were eventually *a posteriori* genotyped as NCCAH. The misclassification table (Table 4) summarizes the prediction results of the model.

The first column displays the total number of observations in the training set and the validation set, the average percentage correctly classified, and the number of observations classified to each class. As mentioned in the No class column, the model never classified any of the patients to other classes than those available (i.e., YPred score value < 0). Fisher's probability is the probability of the table occurring by chance and is satisfied when $p < 0.05$ for 95% confidence. Here, the success rate for the prediction of the membership class was 100% with either no false positive or false negative scores. The five NCCAH patients in the validation set (post-hoc confirmation by CYP21A2 analysis) were all assigned to the NCCAH group in our model. Therefore, thanks to the OPLS model integrating steroid fingerprint and clinical data, the classification of NCCAH patients and PP patients was optimum with a validation accuracy of 100%.

4. Discussion

The objective of the present study was to develop, based on a mathematical ML model, a score including basal circulating steroid profile and clinical data to avoid the use of an ACTH testing in the differential diagnosis of NCCAH in a pediatric population.

Indeed, ACTH tests are useful in detecting AI in children with NCCAH. In our study, similar to other pediatric studies [30–32], a high number of NCCAH patients presented a cortisol peak below 180 ng/ml (12 NCCAH patients out of 13). These data underlined the importance of glucocorticoid stress dose in children undergoing stressful situations like surgery or illness.

Our findings uncovered clinical and biological differences between children with premature pubic hair and NCCAH with significant results for BA and basal levels of 21-DB, 17-OHP, 21-DF, A4, T, Preg, and 11OHA4. There were no overlaps for 21-DB and 21-DF (concentration threshold = 12.42 pg/ml), or for 17-OHP (concentration threshold = 3.62 ng/ml). It was consistent with previous studies, in particular for 21-deoxysteroids [19,20,22]. Indeed, 21-DF and 21-DB are strictly adrenal metabolites and therefore represent potentially more specific biomarkers for NCCAH diagnoses than 17-OHP, which has a dual ovarian and adrenal origin [22,43,44]. Some authors have already proposed a threshold of basal 17-OHP [34], but this remains controversial [32,33]. With the validated mass spectrometry approach (LC-MS/MS), we were able to routinely quantify simultaneously in 10 min 16 circulating steroids via a direct measurement method with excellent sensitivity and specificity (certified by the French national quality control organization). We can thereby attribute a specific steroidome, hormonal signature, or “fingerprint” to each patient [22].

In our patients, basal measurements of 17-OHP were surprisingly

elevated, and could already distinguish NCCAH patients from PP patients in univariate analysis. This was not always the case for our patients, for whom a threshold value of 17-OHP was always discussed, presumably because our pediatric patients had more severe features than late onset adrenal hyperplasia diagnosed in adulthood and because they were bearing a severe mutation in 54% of the cases. However, the use of 17-OHP which characterizes 21-OHD is not without criticism [35]. False negative rates of up to 22% have been reported in infant screening [36], particularly when mothers were exposed to glucocorticoids prenatally [37]. False positives in some premature infants or in other forms of CAH including 11-hydroxylase deficiency were also reported [38]. Moreover a single basal 17-OHP concentration is often insufficient to diagnose non-classic 21-OHD carriers [39].

The use of mass spectrometry gives a more specific view of adrenal steroid levels in 21-OHD compared with immunoassays [40]. Furthermore, in terms of cost, according to our own experience in our clinical laboratory, a panel of about 16 steroids can be more advantageous than relying on specific expensive immunoassay kits. This method is more specific and more sensitive than radioimmunoassay (RIA) because it allows the quantifying of steroid profiles including 21-DF within 150 μ L of serum. It also allows the measurement of other steroids such as 21-DB, which is particularly informative for CAH. To our knowledge, 21-DB has not been extensively studied until now. According to our previous data, plasma basal 21-DB concentrations measured by RIA [41], and more recently in LC-MS/MS [22], could represent an interesting additional biomarker to identify patients with NCCAH. Although 17-OHP is usually quantified on its own, the addition and combination of new strictly adrenal steroids, such as 21-DF and 21-DB, could enhance the specificity for the diagnosis of NCCAH. Miller et al. even suggest replacing 17-OHP with 21-DF in the newborn screening program [35].

With the ability to routinely and simultaneously quantify 16 circulating steroids, we highlighted the use of some steroids not often used in practice (such as 21-DB with no overlap) and other better-known steroids (like 21-DF). 21-DF has already been evaluated in children with 21-OHD, and was deemed an excellent specific marker of this disease [20]. 21-DF is useful in the diagnosis and the follow-up of NCCAH and for the detection of heterozygotes [42].

Determining the status of heterozygotes and NCCAH is very important, especially in those carrying severe alleles, for genetic counseling to anticipate the risk of AI and genital ambiguity at birth [43]. In our study, 54% of NCCAH children had a severe mutation. It is further essential to explain the screening to the family's proband and the future partner in family planning [44,45].

A non-classical form of mutation (M283V of the NCCAH patient no. 10, Table 1) has been described in just one case of a patient with NCCAH [46]. Notably, all of our NCCAH patients had the V281L mutation, the most common mutation in patients with NCCAH [47]. Three had the IVS2-13 C > G allele, the most frequent mutation in classical forms [9].

Our findings demonstrated a difference in clinical features and advanced BA, as is already known in the literature [7]. Even using subjective methods, it is easy to perform routinely in clinical settings [26]. In our study, all children with NCCAH demonstrated advanced BA and six had advanced BA of more than two years. Some reports propose glucocorticoid therapy for these patients to avoid a short final stature [11]. In combination with biological measurements, BA was an important part of our design score for the diagnosis of NCCAH, suggesting that an advanced BA may be the only difference between clinical features of NCCAH patients and PP [34]. We limited our study to prepubertal children from 7 to 9 years old (mean age 7.14 years (+/- 2.32)) to focus on adrenal steroids responsible for adrenarche and to free ourselves from sex variations induced by gonadal puberty. Analysis of steroids only secreted by the adrenal gland could be a further area of study [47].

The model we developed predicted with 100% sensitivity and specificity the diagnoses of NCCAH in a population of children with premature pubic hair (n = 97), and could probably be extended to older populations with mild features. The constitutions of the training set (first

year) and the validation set (second year) led us to evaluate the robustness of our model more accurately than with a simple internal cross-validation. Multivariate data analysis [48] and ML methods have already been evaluated in adrenal diseases and can shorten diagnosis in some cases [49,50]. In this study, OPLS-DA analysis was chosen as the ML algorithm since our dataset included many steroids with potentially high collinearity. We trained and evaluated 5 other ML models (support vector machine, nearest neighbors classifier, decision tree, random forest, and Gaussian Naive Bayes) which confirmed the discriminative power of the data, independently of the chosen method (not shown). However, their performances were either worse or equivalent with OPLS-DA. Moreover, this approach gave us the opportunity to present the data on 2D score plots, creating a bird's eye view to summarize and study the degree of similarity (or dissimilarity) of the patient's bio-clinical signatures.

In future, multivariate mathematical approaches such as these, which can integrate anthropometric data (gender, age, BMI, etc.), biological measurements, and even radiological images, could generate a unique fingerprint for each patient and offer new potentialities in the future diagnosis toolbox. Model training iteration processes could eventually be performed [51, 52]. Indeed, with the selected features, additional patients could be included to build refined models over time. The proposed model could therefore find its place in the care management pathway of children with premature pubarche and late onset 21-OHD. The method may be extended to higher ages; this work is in progress with adult patients from non-pediatric endocrinology units, considering age and sex references for adult patients. Prospective interventional studies will be needed to validate this model for clinical use [50].

5. Conclusion

To conclude, LC-MS/MS enabled assignment of a metabolic fingerprint to each patient. Combining this with a statistical model allowed the construction of a NCCAH diagnostic score with 100% sensitivity and specificity in our cohort. The most significant variables were 21-DB, 17-OHP, and 21-DF. If this score was routinely implemented, use of the ACTH test could be restricted to screening patients for AI, which, based on the prevalence of NCCAH, would be nearly one in ten children. Genetic analysis of *CYP21A2* would remain essential in NCCAH patients to identify severe mutations. Further studies are needed for the validation of this score, particularly with the use of a larger multicenter prospective cohort.

CRedit authorship contribution statement

Hélène Agnani: Conceptualization, Methodology, Writing – review & editing. **Guillaume Bachelot:** Data curation, Data analysis. **Thibaut Eguether:** Visualization, Investigation. **Bettina Ribault:** mass spectrometry analysis. **Jean Fiet:** Methodology, Writing – review & editing. **Yves Le Bouc:** Methodology, Writing – review & editing. **Irène Netchine:** Methodology, Writing. **Muriel Houang:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Antonin Lamazière:** Conceptualization, Methodology, Data analysis, Writing – review & editing.

Acknowledgments

We thank Dominique Farabos and Nicolas Hélin for technical supports.

Disclosure Summary

The authors have nothing to disclose.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.jsbmb.2022.106085](https://doi.org/10.1016/j.jsbmb.2022.106085).

References

- [1] S.E. Oberfield, A.B. Sopher, A.T. Gerken, Approach to the girl with early onset of pubic hair, *J. Clin. Endocrinol. Metab.* 96 (2011) 1610–1622, <https://doi.org/10.1210/jc.2011-0225>.
- [2] P.W. Speiser, B. Dupont, P. Rubinstein, A. Piazza, A. Kastelan, M.I. New, High frequency of nonclassical steroid 21-hydroxylase deficiency, *Am. J. Hum. Genet.* 37 (1985) 650–667.
- [3] R.C. Wilson, S. Nimkarn, M. Domic, J. Obeid, M.R. Azar, M. Azar, et al., Ethnic-specific distribution of mutations in 716 patients with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency, *Mol. Genet. Metab.* 90 (2007) 414–421, <https://doi.org/10.1016/j.ymgme.2006.12.005>.
- [4] E. Carmina, D. Dewailly, H.F. Escobar-Morreale, F. Kelestimur, C. Moran, S. Oberfield, et al., Non-classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency revisited: an update with a special focus on adolescent and adult women, *Hum. Reprod. Update* 23 (2017) 580–599, <https://doi.org/10.1093/humupd/dmx014>.
- [5] M. Finkelstein, J.M. Shafer, Inborn errors of steroid biosynthesis, *Physiol. Rev.* 59 (1979) 353–406, <https://doi.org/10.1152/physrev.1979.59.2.353>.
- [6] C. Moran, R. Azziz, E. Carmina, D. Dewailly, F. Fruzzetti, L. Ibañez, et al., 21-Hydroxylase-deficient nonclassical adrenal hyperplasia is a progressive disorder: a multicenter study, *Am. J. Obstet. Gynecol.* 183 (2000) 1468–1474, <https://doi.org/10.1067/mob.2000.108020>.
- [7] K.J. Pijnenburg-Kleizen, G.F. Born, B.J. Otten, D.A. Schott, E.L. van den Akker, W. H. Stokvis-Brantsma, et al., Absence of clinically relevant growth acceleration in untreated children with non-classical congenital adrenal hyperplasia, *Horm. Res. Paediatr.* 77 (2012) 164–169.
- [8] O. Eyal, Y. Tenenbaum-Rakover, S. Shalitin, S. Israel, N. Weintrob, Adult height of subjects with nonclassical 21-hydroxylase deficiency, *Acta Paediatr.* 102 (2013) 419–423, <https://doi.org/10.1111/apa.12147>.
- [9] M.G. Forest, V. Tardy, M. Nicolino, M. David, Y. Morel, 21-hydroxylase deficiency: an exemplary model of the contribution of molecular biology in the understanding and management of the disease, *Ann. d'Endocrinologie* 66 (2005) 225–232, [https://doi.org/10.1016/S0003-4266\(05\)81754-8](https://doi.org/10.1016/S0003-4266(05)81754-8).
- [10] S.F. Witchel, Congenital adrenal hyperplasia, *J. Pediatr. Adolesc. Gynecol.* 30 (2017) 520–534, <https://doi.org/10.1016/j.jpag.2017.04.001>.
- [11] P.W. Speiser, W. Arlt, R.J. Auchus, L.S. Baskin, G.S. Conway, D.P. Merke, et al., Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an endocrine society clinical practice guideline, *J. Clin. Endocrinol. Metab.* 103 (2018) 4043–4088, <https://doi.org/10.1210/jc.2018-01865>.
- [12] F. Kuttann, P. Couillin, F. Girard, L. Billaud, M. Vincens, C. Boucekkine, et al., Late-onset adrenal hyperplasia in hirsutism, *N. Engl. J. Med.* 313 (1985) 224–231, <https://doi.org/10.1056/NEJM198507253130404>.
- [13] K.R. van der Ploeg, B.G. Wolthers, G.T. Nagel, M. Volmer, N.M. Drayer, The diagnosis of 21-hydroxylase deficiency in a prematurely born infant on the basis of the urinary steroid excretion pattern, *Clin. Chim. Acta* 120 (1982) 341–353, [https://doi.org/10.1016/0009-8981\(82\)90375-8](https://doi.org/10.1016/0009-8981(82)90375-8).
- [14] J.W. Honour, C.G. Brook, Clinical indications for the use of urinary steroid profiles in neonates and children, *Ann. Clin. Biochem.* 34 (Pt 1) (1997) 45–54, <https://doi.org/10.1177/000456329703400107>.
- [15] J.M. Hines, I. Bancos, C. Bancos, R.D. Singh, A.V. Avula, W.F. Young, et al., High-resolution, accurate-mass (HRAM) mass spectrometry urine steroid profiling in the diagnosis of adrenal disorders, *Clin. Chem.* 63 (2017) 1824–1835, <https://doi.org/10.1373/clinchem.2017.271106>.
- [16] I. Bancos, A.E. Taylor, V. Chortis, A.J. Sitch, C. Jenkinson, C.J. Davidge-Pitts, et al., Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the EURINE-ACT study: a prospective test validation study, *Lancet Diabetes Endocrinol.* 8 (2020) 773–781, [https://doi.org/10.1016/S2213-8587\(20\)30218-7](https://doi.org/10.1016/S2213-8587(20)30218-7).
- [17] S.A. Wudy, M. Hartmann, J. Homoki, Hormonal diagnosis of 21-hydroxylase deficiency in plasma and urine of neonates using benchtop gas chromatography-mass spectrometry, *J. Endocrinol.* 165 (2000) 679–683, <https://doi.org/10.1677/joe.0.1650679>.
- [18] J.W. Honour, G. Rumsby, Problems in diagnosis and management of congenital adrenal hyperplasia due to 21-hydroxylase deficiency, *J. Steroid Biochem. Mol. Biol.* 45 (1993) 69–74, [https://doi.org/10.1016/0960-0760\(93\)90124-f](https://doi.org/10.1016/0960-0760(93)90124-f).
- [19] C.Z. Minutti, J.M. Lacey, M.J. Magera, S.H. Hahn, M. McCann, A. Schulze, et al., Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia, *J. Clin. Endocrinol. Metab.* 89 (2004) 3687–3693, <https://doi.org/10.1210/jc.2003-032235>.
- [20] N. Janzen, M. Peter, S. Sander, U. Steuerevald, M. Terhardt, U. Holtkamp, et al., Newborn screening for congenital adrenal hyperplasia: additional steroid profile using liquid chromatography-tandem mass spectrometry, *J. Clin. Endocrinol. Metab.* 92 (2007) 2581–2589, <https://doi.org/10.1210/jc.2006-2890>.
- [21] E. Schwarz, A. Liu, H. Randall, C. Haslip, F. Keune, M. Murray, et al., Use of steroid profiling by UPLC-MS/MS as a second tier test in newborn screening for congenital adrenal hyperplasia: the Utah experience, *Pediatr. Res.* 66 (2009) 230–235, <https://doi.org/10.1203/PDR.0b013e3181aa3777>.
- [22] J. Fiet, Y. Le Bouc, J. Guéchet, N. Hélin, M.A. Maubert, D. Farabos, et al., A liquid chromatography/tandem mass spectrometry profile of 16 serum steroids, including 21-deoxycortisol and 21-deoxycorticosterone, for management of congenital adrenal hyperplasia, *J. Endocr. Soc.* 1 (2017) 186–201, <https://doi.org/10.1210/je.2016-1048>.
- [23] A. Crinò, R. Schiaffini, P. Ciampalini, S. Spera, L. Beccaria, F. Benzi, et al., Genetic obesity study group of Italian Society of pediatric endocrinology and diabetology (SIEDP), hypogonadism and pubertal development in Prader-Willi syndrome, *Eur. J. Pediatr.* 162 (2003) 327–333, <https://doi.org/10.1007/s00431-002-1132-4>.
- [24] Michel. Sempé Pédrón, Guy., Roy-Pernot, Marie-Paule, Auxologie: méthode et séquences, Theraplix, Paris, 1979.
- [25] M.F. Rolland-Cachera, T.J. Cole, M. Sempé, J. Tichet, C. Rossignol, A. Charraud, Body Mass Index variations: centiles from birth to 87 years, *Eur. J. Clin. Nutr.* 45 (1991) 13–21.
- [26] W.W. Greulich, S.I. Pyle (Eds.), Radiographic Atlas of Skeletal Development of the Hand and Wrist, 2, repr, Stanford Univ. Press [u.a.], Stanford, Calif, 2011.
- [27] M. Pérez-Enciso, M. Tenenhaus, Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach, *Hum. Genet.* 112 (2003) 581–592, <https://doi.org/10.1007/s00439-003-0921-9>.
- [28] M. Tenenhaus, La Régression PLS: Théorie Et Pratique, Editions Technip, 1998.
- [29] S. Wiklund, E. Johansson, L. Sjöström, E.J. Mellerowicz, U. Edlund, J.P. Shockcor, et al., Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, *Anal. Chem.* 80 (2008) 115–122, <https://doi.org/10.1021/ac0713510>.
- [30] R. Balducci, B. Boscherini, A. Mangiantini, M. Morellini, V. Toscano, Isolated precocious pubarche: an approach, *J. Clin. Endocrinol. Metab.* 79 (1994) 582–589, <https://doi.org/10.1210/jcem.79.2.8045980>.
- [31] C. Dacou-Voutetakis, M. Dracopoulou, High incidence of molecular defects of the CYP21 gene in patients with premature adrenarche, *J. Clin. Endocrinol. Metab.* 84 (1999) 1570–1574, <https://doi.org/10.1210/jcem.84.5.5683>.
- [32] A. Stoupa, L. Gonzalez-Briceno, G. Pinto, D. Samara-Boustani, C. Thalassinou, I. Flechtner, et al., Inadequate cortisol response to the tetracosactide (synacthen) test in non-classic congenital adrenal hyperplasia: an exception to the rule? *Horm. Res. Paediatr.* 83 (2015) 262–267, <https://doi.org/10.1159/000369901>.
- [33] J.-B. Armengaud, M.-L. Charkaluk, C. Trivin, V. Tardy, G. Bréart, R. Brauner, et al., Precocious pubarche: distinguishing late-onset congenital adrenal hyperplasia from premature adrenarche, *J. Clin. Endocrinol. Metab.* 94 (2009) 2835–2840, <https://doi.org/10.1210/jc.2009-0314>.
- [34] R. Bello, Y. Leventhal, L. Lazar, S. Shalitin, A. Tenenbaum, M. Phillip, et al., Basal 17-hydroxyprogesterone cannot accurately predict nonclassical congenital adrenal hyperplasia in children and adolescents, *Acta Paediatr.* 106 (2017) 155–160, <https://doi.org/10.1111/apa.13630>.
- [35] U. Ambrozjak, A. Kępczyńska-Nyk, A. Kuryłowicz, E.M. Małunowicz, A. Wójcicka, P. Miśkiewicz, et al., The diagnosis of nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency, based on serum basal or post-ACTH stimulation 17-hydroxyprogesterone, can lead to false-positive diagnosis, *Clin. Endocrinol. (Oxf.)* 84 (2016) 23–29, <https://doi.org/10.1111/cen.12935>.
- [36] W.L. Miller, Congenital adrenal hyperplasia: time to replace 17OHP with 21-deoxycortisol, *Horm. Res. Paediatr.* 91 (2019) 416–420, <https://doi.org/10.1159/000501396>.
- [37] K. Sarafoglou, K. Banks, J. Kylo, S. Pittock, W. Thomas, Cases of congenital adrenal hyperplasia missed by newborn screening in Minnesota, *JAMA* 307 (2012) 2371–2374, <https://doi.org/10.1001/jama.2012.5281>.
- [38] F. Gatelais, J. Berthelot, F. Beringue, P. Descamps, D. Bonneau, J.-M. Limal, et al., Effect of single and multiple courses of prenatal corticosteroids on 17-hydroxyprogesterone levels: implication for neonatal screening of congenital adrenal hyperplasia, *Pediatr. Res.* 56 (2004) 701–705, <https://doi.org/10.1203/01.PDR.0000142733.50918.6E>.
- [39] P. Cavarzere, D. Samara-Boustani, I. Flechtner, M. Dechaux, C. Elie, V. Tardy, et al., Transient hyper-17-hydroxyprogesteronemia: a clinical subgroup of patients diagnosed at neonatal screening for congenital adrenal hyperplasia, *Eur. J. Endocrinol.* 161 (2009) 285–292, <https://doi.org/10.1530/EJE-09-0145>.
- [40] M.I. New, F. Lorenzen, A.J. Lerner, B. Kohn, S.E. Oberfield, M.S. Pollack, et al., Genotyping steroid 21-hydroxylase deficiency: hormonal reference data, *J. Clin. Endocrinol. Metab.* 57 (1983) 320–326, <https://doi.org/10.1210/jcem-57-2-320>.
- [41] S.R. Dahl, I. Neramo, I. Brønstad, E.S. Husebye, K. Løvås, P.M. Thorsby, Assay of steroids by liquid chromatography-tandem mass spectrometry in monitoring 21-hydroxylase deficiency, *Endocr. Connect.* 7 (2018) 1542–1550, <https://doi.org/10.1530/EC-18-0453>.
- [42] J. Fiet, B. Gueux, M.C. Raux-DeMay, F. Kuttann, P. Vexiau, J.L. Brerault, et al., Increased plasma 21-deoxycorticosterone (21-DB) levels in late-onset adrenal 21-hydroxylase deficiency suggest a mild defect of the mineralocorticoid pathway, *J. Clin. Endocrinol. Metab.* 68 (1989) 542–547, <https://doi.org/10.1210/jcem-68-3-542>.
- [43] A.F. Turcu, J. Rege, R. Chomic, J. Liu, H.K. Nishimoto, T. Else, et al., Profiles of 21-carbon steroids in 21-hydroxylase deficiency, *J. Clin. Endocrinol. Metab.* 100 (2015) 2283–2290, <https://doi.org/10.1210/jc.2015-1023>.
- [44] J. Fiet, B. Gueux, M. Gourmelen, F. Kuttann, P. Vexiau, P. Couillin, et al., Comparison of basal and adrenocorticotropic-stimulated plasma 21-deoxycortisol and 17-hydroxyprogesterone values as biological markers of late-onset adrenal hyperplasia, *J. Clin. Endocrinol. Metab.* 66 (1988) 659–667, <https://doi.org/10.1210/jcem-66-4-659>.
- [45] M. Bidet, C. Bellané-Chantelot, M.-B. Galand-Portier, V. Tardy, L. Billaud, K. Laborde, et al., Clinical and molecular characterization of a cohort of 161 unrelated women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency and 330 family members, *J. Clin. Endocrinol. Metab.* 94 (2009) 1570–1578, <https://doi.org/10.1210/jc.2008-1582>.

- [46] N. Skordis, C. Shamma, E. Efstathiou, K. Kaffe, V. Neocleous, L.A. Phylactou, Endocrine profile and phenotype-genotype correlation in unrelated patients with non-classical congenital adrenal hyperplasia, *Clin. Biochem.* 44 (2011) 959–963, <https://doi.org/10.1016/j.clinbiochem.2011.05.013>.
- [47] S. Livadas, M. Dracopoulou, A. Dastamani, A. Sertedaki, M. Maniati-Christidi, A.-M. Magiakou, et al., The spectrum of clinical, hormonal and molecular findings in 280 individuals with nonclassical congenital adrenal hyperplasia caused by mutations of the CYP21A2 gene, *Clin. Endocrinol. (Oxf.)* 82 (2015) 543–549, <https://doi.org/10.1111/cen.12543>.
- [48] C. Minutolo, A.D. Nadra, C. Fernández, M. Taboas, N. Buzzalino, B. Casali, et al., Structure-based analysis of five novel disease-causing mutations in 21-Hydroxylase-Deficient patients, *PLoS ONE* 6 (2011), e15899, <https://doi.org/10.1371/journal.pone.0015899>.
- [49] C. Kamrath, L. Wettstaedt, C. Boettcher, M.F. Hartmann, S.A. Wudy, Androgen excess is due to elevated 11-oxygenated androgens in treated children with congenital adrenal hyperplasia, *J. Steroid Biochem. Mol. Biol.* 178 (2018) 221–228, <https://doi.org/10.1016/j.jsbmb.2017.12.016>.
- [50] M.R. Lasarev, E.R. Bialk, D.B. Allen, P.K. Held, Application of principal component analysis to newborn screening for congenital adrenal hyperplasia, *J. Clin. Endocrinol. Metab.* 105 (2020) e2930–e2940, <https://doi.org/10.1210/clinem/dgaa371>.
- [51] E.H. Wilkes, G. Rumsby, G.M. Woodward, Using machine learning to aid the interpretation of urine steroid profiles, *Clin. Chem.* 64 (2018) 1586–1595, <https://doi.org/10.1373/clinchem.2018.292201>.
- [52] G. Eisenhofer, C. Durán, C.V. Cannistraci, M. Peitzsch, T.A. Williams, A. Riestler, et al., Use of steroid profiling combined with machine learning for identification and subtype classification in primary aldosteronism, *JAMA Netw. Open* 3 (2020), e2016209, <https://doi.org/10.1001/jamanetworkopen.2020.16209>.

Table des illustrations

Figure 1 : Répartition schématique des origines/causes de l'infertilité dans le couple.....	15
Figure 2 : Algorithme décisionnel d'après les recommandations de la SFE pour l'exploration des hyperandrogénies, à partir d'un résultat pathologique (élevé) de testostérone plasmatique.	21
Figure 3 : Principales causes d'infertilités féminines – places des troubles ovulatoires et du SOPK, première cause d'infertilité chez la femme.	23
Figure 4 : Algorithme diagnostique pour le SOPK, d'après les recommandations internationales 2023	24
Figure 5 : Place centrale de l'optimisation pré-conceptionnelle du mode de vie dans algorithme de prise en charge de l'infertilité dans un contexte de SOPK, avant tout autre traitement en AMP.	25
Figure 6 : Hypothèse du continuum et de recouvrement des situations entre le SOPK et l'infertilité inexplicée.	26
Figure 7 : Vue d'ensemble de la stéroïdogénèse surrénalienne et la modulation périphérique.	29
Figure 8 : Principe général du fonctionnement d'un spectromètre de masse avec une source d'ionisation par electrospray	32
Figure 9 : Structure du système de chromatographie liquide couplé à la spectrométrie de masse en tandem (LC-MS/MS).	33
Figure 10 : Spectres générés par la LC-MS/MS.	34
Figure 11 : Deux approches différentes : programmation traditionnelle et apprentissage automatique.	37
Figure 12 : Fonctionnement d'un modèle de classification.	38

Résumé :

L'infertilité représente un enjeu majeur de santé publique, affectant au moins 50 millions de couples à travers le monde. Les troubles métaboliques jouent un rôle central dans l'infertilité inexpliquée et le syndrome des ovaires polykystiques (SOPK). Ces conditions, souvent liées à des modes de vie modernes, caractérisés par une alimentation inadéquate et un manque d'activité physique, peuvent entraîner des conséquences métaboliques affectant la fertilité. Les approches thérapeutiques visant à améliorer le mode de vie, l'alimentation et l'activité physique pourraient donc être pertinentes seules, ou en complément de l'assistance médicale à la procréation (AMP), dans ces contextes. L'objectif principal de ce travail est d'explorer ces conditions complexes de l'infertilité, difficiles à identifier et à caractériser, dans lesquelles le(s) métabolisme(s) joue(nt) un rôle essentiel dans la physiopathologie. Notre but est de développer des outils diagnostiques et de suivi en utilisant des approches omiques et des méthodes basées sur l'intelligence artificielle, dans le but à plus long terme d'opérer un changement de sémiologie pour les troubles métaboliques et endocriniens de la fertilité. Deux applications ont été menées en parallèles : la première vise à développer un modèle d'apprentissage automatique pour identifier les facteurs nutritionnels et métaboliques liés à l'infertilité inexpliquée, à partir de données du PHRC ALIFERT. Le modèle utilise des profils "couples" et des méthodes d'apprentissage pour distinguer les couples fertiles des infertiles, sur la base d'une signature obtenue à partir d'une sélection de biomarqueurs métaboliques et nutritionnels. La deuxième application se concentre sur la voie de biosynthèse des stéroïdes pour explorer les causes de l'hyperandrogénie chez les femmes, dont le syndrome des ovaires polykystiques (SOPK) et l'hyperplasie congénitale par bloc en 21-hydroxylase (NC21OHD). Des techniques de modélisation mathématique et des analyses biochimiques par chromatographie en phase liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS) ont été associées pour simplifier le diagnostic du SOPK, notamment dans le diagnostic différentiel SOPK / NC21OHD. Nous avons montré, à travers deux applications, que le traitement de données hétérogènes par des outils d'intelligence artificielle innovants permettait d'améliorer ou de simplifier l'identification des patients ou des couples infertiles. Ces outils pourraient potentiellement redéfinir l'infertilité inexpliquée à long terme en remettant en question la notion d'absence de diagnostic et en démontrant l'existence d'une composante métabolique importante, réduisant ainsi progressivement la part des situations inexpliquées. Par ailleurs, la nature hétérogène du SOPK imposera l'application de méthodes de modélisation multi-blocs, sophistiquées, en vue d'une meilleure identification des sous-groupes de patientes et de la prise en compte des comorbidités.

Mots clés : infertilité, intelligence artificielle, métabolisme, spectrométrie de masse, stéroïdes

Using artificial intelligence methods to explore metabolic and endocrine disorders in reproductive functions and infertility.

Abstract :

Infertility is a major public health issue, affecting at least 50 million couples worldwide. Metabolic disorders play a central role in unexplained infertility and polycystic ovary syndrome (PCOS). These conditions, often related to modern lifestyles involving inadequate diet and lack of physical activity, can lead to metabolic consequences affecting fertility. Therapeutic approaches aimed at improving lifestyle, diet and physical activity could therefore be relevant alone, or as a complement to medically assisted reproduction (ART), in these contexts. The main objective of this work is to explore these complex conditions of infertility, difficult to identify and characterize, in which metabolism(s) play(s) an essential role in pathophysiology. Our aim is to develop diagnostic and monitoring tools using omics approaches and artificial intelligence-based methods, with the longer-term goal of changing the semiology of metabolic and endocrine fertility disorders. Two applications have been conducted in parallel: the first aims to develop a machine learning model to identify nutritional and metabolic factors linked to unexplained infertility, based on data from the ALIFERT PHRC. The model uses "couple" profiles and learning methods to distinguish between fertile and infertile couples, based on a signature obtained from a selection of metabolic and nutritional biomarkers. The second application focuses on the steroid biosynthesis pathway to explore the causes of hyperandrogenism in women, including polycystic ovary syndrome (PCOS) and congenital 21-hydroxylase block hyperplasia (NC21OHD). Mathematical modelling techniques and biochemical analyses using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) have been combined to simplify the diagnosis of PCOS, particularly in the differential diagnosis of PCOS / NC21OHD. Using two applications, we have shown that the processing of heterogeneous data by innovative artificial intelligence tools can improve or simplify the identification of infertile patients or couples. These tools could potentially redefine long-term unexplained infertility by challenging the notion of absence of diagnosis and demonstrating the existence of an important metabolic component, thus progressively reducing the proportion of unexplained situations. Furthermore, the heterogeneous nature of PCOS will require the application of sophisticated, multi-block modeling methods to better identify patient subgroups and take comorbidities into account.

Keywords : infertility, artificial intelligence, metabolism, mass spectrometry, steroids