



HAL
open science

Analyse des performances des athlètes ayant participé aux Jeux Paralympiques et Championnats du monde en fonction de l'origine du handicap : inné versus acquis

Alexis François

► To cite this version:

Alexis François. Analyse des performances des athlètes ayant participé aux Jeux Paralympiques et Championnats du monde en fonction de l'origine du handicap : inné versus acquis. Santé publique et épidémiologie. 2021. dumas-03377118

HAL Id: dumas-03377118

<https://dumas.ccsd.cnrs.fr/dumas-03377118>

Submitted on 14 Oct 2021

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.

MASTER 2 BIOSTATISTIQUE
PROMOTION 2020-2021

**Analyse des performances des athlètes ayant participé
aux Jeux Paralympiques et Championnats du monde en
fonction de l'origine du handicap : inné versus acquis**

Laboratoire IRMES



Rapporteur François Alexis
Encadrants Le Toquin Bryan (ingénieur recherche) / Schipman Julien (doctorant)

Du 01/03/2021 au 28/08/2021

Remerciements

Je tiens à remercier Mr le professeur Jean-François Toussaint de m'avoir accueilli au sein de son laboratoire et partagé ses conseils et expertise.

Je remercie aussi Mr Bryan Le Toquin et Mr Julien Schipman pour leur encadrement précieux sur l'ensemble de ces 6 mois, pour leur expertise et partage sur la thématique complexe du paralympisme.

Je remercie Mr Robin Genuer d'avoir été mon référent de l'Université de Bordeaux.

Je remercie Mr Quentin Delarochelambert, Mr Imad Hamri, Mr Geoffroy Berthelot, Mme Cécile Proust-Lima pour leurs aides et conseils aux différentes étapes de ce stage.

Enfin, je remercie l'ensemble des stagiaires pour les discussions et moments passés ensembles.

Table des matières

1	Introduction	9
2	L'INSEP et le Laboratoire IRMES	9
3	Contexte de l'étude	11
3.1	Le handicap	11
3.2	Le mouvement paralympique	12
3.3	Le système de classification	13
3.4	Objectifs de l'étude	14
4	Matériels et méthodes	15
4.1	Origine des données	15
4.2	Construction de la variable d'intérêt : origine du handicap	15
4.3	Indicateur de performance	17
4.4	Relation âge et performance : les fonctions Moore & IMAP	17
4.5	Analyses statistiques	19
4.5.1	Statistiques descriptives	19
4.5.2	Analyses longitudinales	20
5	Résultats	23
5.1	Données brutes, bases Résultats	24
5.2	Données longitudinales, bases Ranking	28
6	Discussion	34
7	Conclusion	39
7.1	Conclusion générale	39
7.2	Conclusion personnelle	39
	Bibliographie	40
	Annexes	43

Table des abréviations

CIP : Comité International Paralympique

CNSA : Caisse Nationale de Solidarité pour l'Autonomie

CNSD : Centre National des Sports de la Défense

CPSF : Comité Paralympique et Sportif Français

CREPS : Centre de Ressources, d'Expertise et de la Performance Sportive

EMPOWER : Exploring Menstrual Periods Of Women athletes to Escalate Ranking

FFH : Fédération Française Handisport

IMAP : Integrative Model of Age-Performance

INRIA : Institut National de Recherche en Informatique et Automatique

INSEP : Institut National du Sport, de l'Expertise et de la Performance

IRMES : Institut de Recherche bio-Médicale et d'Épidémiologie du Sport

JO : Jeux Olympiques

JP : Jeux Paralympiques

OMS : Organisation Mondiale de la Santé

PSH : Personne en Situation de Handicap

Liste des tableaux

1	Caractéristiques des athlètes en natation en fonction de l'origine du handicap. Base Résultats (n=1239).	24
2	Résultats modèle mixte IMAP. Analyse univariable avec l'origine du handicap. Natation.	30
3	Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Natation.	31
4	Caractéristiques des athlètes en athlétisme en fonction de l'origine du handicap. Base Résultats (n=2771).	43
5	Caractéristiques des athlètes en natation. Données répétées.	48
6	Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de course.	48
7	Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de lancer.	48
8	Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de saut.	48
9	Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Natation.	49
10	Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de course.	50
11	Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de lancer.	51
12	Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de saut.	52
13	Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de course.	59
14	Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de lancer.	59
15	Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de saut.	60

Table des figures

1	Diagramme de flux concernant la sélection des performances en natation et athlétisme.	23
2	Distribution de % de la meilleure performance en fonction de l'origine du handicap pour chaque sexe en natation. Handicap important (n=986).	26
3	Distribution de % de la meilleure performance en fonction de l'origine du handicap et du type d'épreuve. Athlétisme, hommes (n=3481).	27
4	Exemple de trajectoires de performance maximale à chaque âge en natation.	28
5	Trajectoires selon IMAP à partir des estimations des paramètres individuels du modèle mixte. Natation.	32
6	Méthode des moindres carrés pour l'estimation des paramètres des fonctions Moore et IMAP.	44
7	Distribution du % de la meilleure performance en fonction de l'origine du handicap. Natation (n=5034).	45
8	Distribution du % de la meilleure performance en fonction de l'origine du handicap et du type d'épreuve. Toutes sévérités, hommes. Natation (n=2702).	45
9	Distribution du % de la meilleure performance en fonction de l'origine du handicap et du type d'épreuve. Toutes sévérités, femmes. Natation (n=2332).	46
10	Distribution du % de la meilleure performance en fonction de l'origine du handicap. Athlétisme (n=6097).	46
11	Distribution du % de la meilleure performance en fonction de l'origine du handicap et de la sévérité. Athlétisme, épreuves debout, hommes (n=2614).	47
12	Distribution du % de la meilleure performance en fonction de l'origine du handicap et de la sévérité. Athlétisme épreuves fauteuil, hommes (n=1304).	47
13	Trajectoires de performance en natation. Échantillon de 7 ou plus performances annuelles. Courbe bleue : smoothing méthode "loess".	53
14	Trajectoires de performance en athlétisme. Échantillon de 7 ou plus performances annuelles. Courbe bleue : smoothing méthode "loess".	53
15	Code modèle SAEMIX.	54
16	Distributions des effets aléatoires. Modèle mixte multivariable. Natation.	55
17	Predictions versus observations. Modèle mixte multivariable. Natation.	56
18	Visual predictive check (vpc). Modèle mixte multivariable. Natation.	57
19	Distribution des erreurs. Modèle mixte multivariable. Natation.	57
20	Exemples de fit individuels. Modèle mixte multivariable. Natation.	58
21	Distribution des maximums de performance théorique en fonction de l'origine du handicap et de sa sévérité. Natation, hommes (n=779).	61
22	Distribution des maximums de performance théorique en fonction de l'origine du handicap et de sa sévérité. Natation, femmes (n=621).	61

Résumé

L'optimisation de la performance sportive de très haut niveau pour les personnes en situation de handicap est régie par un système de classification complexe et remis en question. Ce dernier ne considère pas l'origine du handicap (congénital ou acquis au cours de la vie). Cette étude a donc pour objectif d'étudier les performances des athlètes en fonction du caractère inné ou acquis de leur déficience en natation et athlétisme.

L'origine du handicap a été collectée à partir des données biographiques des athlètes. Les performances sont réparties en deux bases : une base de résultats sur les compétitions considérées et une base de "*ranking*" répertoriant les meilleures performances annuelles des athlètes. Une analyse brute a été réalisée sur cette première base et une analyse longitudinale par trajectoires de performance maximale et modèle non-linéaire mixte a été réalisée sur la base de ranking. La relation âge/performance a été modélisée par les fonctions de Moore et IMAP.

Les résultats montrent des échantillons très proches avec des performances qui sont plus élevées en natation pour le groupe acquis et inversement en athlétisme. Les âges au pic de performance sont plus précoces pour le groupe congénital chez les hommes et chez les femmes.

Ces résultats offrent des pistes de réflexion sur la thématique de l'origine du handicap en lien avec la performance et pourraient avoir des implications importantes sur le système de classification. De plus, l'impact sur les trajectoires de performance des athlètes pourrait orienter les choix des entraîneurs et des fédérations sportives.

Mots-clés : handicap, congénital, acquis, performance, trajectoires, modèle non-linéaire mixte

Abstract

The optimization of high-level sport performance for people with disabilities is ruled by a complex and questionable classification system. This system does not consider the origin of the disability (congenital or acquired during life). The aim of this study is therefore to investigate the performance of athletes according to the congenital or acquired nature of their disability in swimming and athletics.

The origin of the disability was collected from the athletes' biographical data. The performances are divided into two datasets : a result dataset on the competitions considered and a "*ranking*" dataset listing the athletes' best annual performances. A raw data analysis was carried out on this first dataset and a longitudinal analysis by maximum performance trajectories and non-linear mixed model was carried out on the ranking dataset. The age/performance relationship was modelled by the Moore and IMAP functions.

The results show very close samples with performances that are higher in swimming for the acquired group and conversely in athletics. The ages at peak performance are earlier for the congenital group in men and women.

These results provide new insights on the issue of the origin of disability in relation to performance and could have important implications for the classification system. Furthermore, the impact on the performance trajectories of athletes could guide choices of coaches and sports federations.

Keywords : disability, congenital, acquired, performance, trajectories, non-linear mixed model

Préambule

Par souci de transparence, des représentations graphiques supplémentaires ont été ajoutées sur un Drive au lien suivant :

https://drive.google.com/drive/folders/16FqpG6O7xYxMj0-JAlbzuUATS9_Dqmy7?usp=sharing

Tout problème d'accessibilité est à signaler à l'adresse mail suivante :

alexisfrancois001@gmail.com

1 Introduction

Dans le cadre de mon Master 2 mention Biostatistique, j'ai réalisé un stage au sein du laboratoire IRMES de l'INSEP à Vincennes (Paris, Île-de-France).

J'ai choisi ce stage qui s'inscrit dans la continuité de mon stage de Master 1 (intitulé *Épidémiologie des accidents traumatiques au cours de la pratique sportive*) du fait de mon projet d'allier la santé et les statistiques avec le milieu sportif. En effet, c'est un milieu passionnant et en pleine émergence avec une demande croissante de l'utilisation de l'outil statistique par les différents acteurs de la haute performance sportive. Ayant réalisé un mémoire de licence sur la reprise d'un projet de vie chez les personnes amputées, j'ai eu la chance de pouvoir mener des entretiens de patients dans le centre de référence de rééducation fonctionnelle de Nouvelle-Aquitaine, à savoir la Tour de Gassies. Cela m'a permis de développer une sensibilité vis-à-vis du handicap et d'avoir une première expérience de terrain autour de cette problématique, ce qui s'inscrit pleinement dans la thématique de ce stage autour de la performance des athlètes en situation de handicap aux Jeux Paralympiques (JP) et Championnats du Monde (CM).

De plus, étant compétiteur régulier dans la pratique du badminton depuis maintenant sept ans, j'ai une grande sensibilité concernant la compréhension de la performance en milieu sportif, ses déterminants ainsi que ses limites.

J'ai réalisé ce stage du 1er mars au 28 août 2021.

Dans un premier temps, je vais décrire la structure dans laquelle j'ai effectué mon stage. Après avoir fourni des éléments de contexte et exposé les objectifs de ma mission, je présenterai les tâches que j'ai réalisées à travers la méthode de travail ainsi que les résultats trouvés. Enfin, ce rapport s'achèvera sur une discussion sur la mission réalisée ainsi qu'une conclusion.

2 L'INSEP et le Laboratoire IRMES

Dirigé par Abdelghani YALOUZ depuis 2017, l'Institut National du Sport, de l'Expertise et de la Performance (INSEP) est un Établissement Public à caractère Scientifique, Culturel et Professionnel (EPSCP) placé sous la tutelle du ministre des Sport [1].

L'INSEP se voit comme *la* structure de référence du sport de haut niveau en France grâce au rayonnement européen et international de ses athlètes sur une multitude de disciplines sportives. En effet, à l'exception du football, une grande partie des équipes de France y est basée. Jouissant d'un domaine de 28 hectares en plein coeur du bois de Vincennes à Paris, l'Institut bénéficie d'un nombre important d'infrastructures (sportives, scolaires, universitaires, médicales) accueillant, selon le rapport d'activités 2018, pas moins de 810 athlètes pour 21 disciplines Olympiques et Paralympiques ainsi que 150 entraîneurs permanents.

Promoteur du réseau *Grand INSEP*, organisation collaborative visant à diffuser les exigences et connaissances de la haute performance sportive en France au sein des différentes structures territoriales (CREPS, CNDS, écoles nationales etc), les missions de l'INSEP sont multiples :

- "Accompagner et soutenir les fédérations sportives pour l'amélioration des performances des Équipes de France dans les compétitions de référence (Jeux Olympiques et Paralympiques, Championnats du monde et d'Europe) [...]",
- Assurer la synergie d'un réseau constitué de structures dédiées au sport de haut niveau (Grand INSEP),
- Assurer "la mise en place d'une offre de formation initiale ou continue permettant d'élever le niveau de compétence et d'expertise des acteurs du développement du sport, et notamment celui des cadres des fédérations sportives inscrites dans une logique de haut niveau",
- Assurer "le positionnement de l'établissement sur l'échiquier international du sport de haut niveau par le développement d'actions de coopération avec les différents acteurs du champ (recherche, enseignement ou expertise),

- Assurer "la conduite de travaux de recherche scientifique, médicale, technologique prioritairement orientés vers la production de la performance sportive".

La place de la recherche à l'INSEP est donc d'une importance capitale pour répondre à ces missions. Cet établissement est constitué de deux grands laboratoires de recherche : le SEP (Sport Expertise et Performance, non présenté ici) ainsi que l'IRMES [2].

Dirigé par le professeur Jean-François TOUSSAINT, l'Institut de Recherche bio-Médicale et d'Épidémiologie du Sport (IRMES, créé en 2006) vise à développer l'apport scientifique autour des quatre axes de recherche suivants : la physiopathologie du sport, l'épidémiologie de la performance, la physiologie expérimentale et la prévention par les activités physiques ou sportives [2]. Ses compétences constituent non seulement un atout concernant la quantification et l'optimisation de la performance sportive comme outil d'aide à la décision pour les différents acteurs du milieu sportif (entraîneurs, préparateurs physiques, professionnels de santé, enseignants) mais aussi concernant l'amélioration du suivi médical des athlètes [2].

Actuellement, le laboratoire est composé de 30 membres dont : 1 directeur, 4 chercheurs, 8 ingénieurs de recherche, 4 doctorants, 3 enseignants chercheurs, 1 assistante de gestion financière et administrative et 9 stagiaires répartis sur différents projets [2].

Les projets en cours s'inscrivent dans la démarche d'identification et de développement des déterminants de la haute performance sportive en vue de la préparation et participation des athlètes français aux JO et JP de Paris 2024. Ils sont au nombre de quatre et couvrent chacun une thématique bien distincte [2] :

- **Detect**, ce projet a pour but de mettre en place des outils d'estimation de potentiels et de médaillabilité pour les athlètes français puis internationaux. L'objectif étant de placer l'athlète dans son contexte concurrentiel et de fournir aux acteurs des méthodes de visualisation graphique de ces estimations afin d'orienter les stratégies de préparation et approches d'une compétition.
- **PerfAnalytics**, en collaboration avec l'INRIA, ce projet a pour but d'améliorer l'utilisation de l'outil qui représente l'analyse vidéo dans l'analyse de la performance sportive. L'objectif étant d'automatiser la reconnaissance de caractéristiques biomécaniques (ex : type de coup porté pour les sports de combat) afin de les "traduire en identificateurs de performance pour caractériser la performance de l'athlète, diagnostiquer l'efficacité des gestes individuels et modéliser une stratégie réussie de séquence d'actions en fonction du contexte environnemental (équipement, adversaires, arbitres, etc.)".
- **EMPOWER**, ce projet a pour but de "de maximiser les performances des athlètes féminines élites en optimisant leurs réponses à l'entraînement par des charges de travail adaptées en synergie avec leur physiologie et leur cycle menstruel".
- **Paraperf**, ce projet, retenu par le Programme Prioritaire de Recherche « Sport de Très Haute Performance » vise à comprendre les enjeux spécifiques de la très haute performance paralympique selon les trajectoires de performances individuelles, l'optimisation du matériel et les environnements humains pour mieux accompagner les athlètes jusqu'aux podiums des Jeux de Paris 2024.

Le travail réalisé au sein de cette structure pendant six mois s'inscrit dans le cadre du projet Paraperf, et plus particulièrement dans le lot de travail numéro 1 qui vise à caractériser les trajectoires de progression des athlètes Paralympiques [3].

3 Contexte de l'étude

3.1 Le handicap

Selon le rapport mondial sur le handicap de l'Organisation Mondiale de la Santé (OMS), plus d'un milliard de personnes seraient en situation de handicap dont 253 millions auraient une déficience visuelle ou encore 75 millions seraient obligées d'utiliser un fauteuil roulant quotidiennement.

La loi française du 11 février 2005 définit le handicap comme "toute limitation d'activité ou restriction de participation à la vie en société subie dans son environnement par une personne en raison d'une altération substantielle, durable ou définitive d'une ou plusieurs fonctions physiques, sensorielles, mentales, cognitives ou psychiques, d'un polyhandicap ou d'un trouble de santé invalidant".

Le concept de handicap est donc évolutif. Alors qu'il se basait uniquement sur un diagnostic médical, le caractère social et environnemental s'y est intégré de par les barrières qu'il érige contre une pleine participation et intégration de l'individu à la société [4]. En effet, toute personne en situation de handicap (PSH) doit pouvoir s'appuyer sur un environnement qui lui est adapté, ce qui, en dépit de cela, l'empêchera de se réaliser pleinement. Les aménagements pour PSH subissent les mêmes problématiques liées aux inégalités de développement où l'accès aux soins ou à des infrastructures adaptées présentent de fortes disparités entre les territoires. Cela se retrouve bien évidemment dans le milieu sportif où la nécessité d'un matériel adapté à la discipline sportive pratiquée vient s'ajouter aux exigences environnementales.

Les types de handicap sont multiples. Il n'en sera pas dressé une liste exhaustive ici mais les handicaps considérés dans cette étude sont énoncés en 3.3.

L'origine du handicap peut être **congénitale**, on parle alors de handicap de naissance ou inné. Un handicap de naissance peut être dû à une anomalie du développement prénatal et peut être de cause génétique (ex : anomalie chromosomique, non forcément héréditaire), toxique (ex : consommations à risque de la mère), environnementale (ex : exposition à un environnement nocif) ou encore traumatique (ex : strangulation par cordon ombilical).

Par opposition, le handicap peut être **acquis** au cours de la vie. Il peut être dû à une blessure traumatique comme les amputations liées aux blessures de guerre ou accidentelle comme en témoignent les dommages corporels associés aux accidents de la route ou encore aux phénomènes de chute (ex : amputations, paralysies).

Le caractère évolutif d'une maladie est à prendre en compte lorsque l'on s'intéresse à l'origine du handicap afin qu'il n'y ait pas d'ambiguïté entre une maladie qui peut être congénitale mais qui déclenche un handicap plus tard au cours de la vie. En effet, il est facilement considéré dans le langage courant que le caractère "congénital" ou "acquis" désigne le mécanisme qui mène au handicap sans tenir compte de son moment d'apparition. Par exemple, la rétinite pigmentaire est une maladie génétique dégénérative qui entraîne une perte progressive de la vision jusqu'à la cécité. Son origine est donc congénitale mais l'origine du handicap doit être considérée comme acquise du fait que l'individu a pu jouir d'une vision "normale" pendant une partie de sa vie. Cette problématique fera l'objet de précisions dans la partie 4 de ce rapport.

Enfin, l'origine du handicap, qu'elle soit congénitale ou acquise influence différemment le vécu et le rapport au handicap [5]. En effet, un individu né avec un handicap développe directement ses capacités motrices et intellectuelles en fonction des interactions avec un environnement qui lui apparaît tel qu'il est [5]. Un individu ayant acquis un handicap au cours de sa vie voit ses interactions avec l'environnement qu'il connaissait totalement bouleversées. En fonction de l'âge d'acquisition du handicap, le processus de réhabilitation de l'individu passe par la capacité à réapprendre à évoluer dans un environnement où les automatismes établis ne sont plus peut s'avérer être un réel défi [5].

Plusieurs questions se posent alors dans le milieu du sport concernant l'origine du handicap : la technicité d'une discipline sportive développée par un individu ayant un handicap depuis la naissance le rend-il moins compétitif qu'un individu qui a pu acquérir une certaine technicité avant d'être touché par le handicap ? Quel est l'impact de la déficience sur la dimension motrice et plus particulièrement sur des paramètres comme l'équilibre ou la coordination ? Sur la dimension mentale ?

3.2 Le mouvement paralympique

La genèse du mouvement paralympique et plus généralement du sport pour les personnes en situation de handicap (PSH) est apparu au cours de la Seconde Guerre Mondiale avec la prise en charge des soldats blessés au combat. Le docteur Ludwig Guttman (1899-1980), neurochirurgien allemand, est souvent référencé comme le créateur du mouvement de part son implication dans le processus de réhabilitation des PSH à travers la pratique de l'activité physique et sportive. Ce dernier est à l'origine de la création d'un centre initialement prévu pour les vétérans lésés vertébraux à l'hôpital de Stoke Mandeville (Londres) où il y organisa les premiers événements sportifs pour PSH et, en 1948, la première compétition nationale [6]. Gagnant rapidement en popularité au fil des années, on considère que les premiers Jeux Paralympiques (JP) d'été eurent lieu à Rome en 1960 alors que les premiers JO "modernes" d'été virent le jour en 1896. Il faudra attendre 1976 pour assister à ceux d'hiver, en Suède (1924, premiers JO d'hiver) [6].

Ainsi, d'un point de vue historique, le mouvement bénéficie d'une antériorité moindre sur le devant de la scène sportive internationale. Le rapport au handicap, les moyens mis à disposition des individus ayant un handicap et souhaitant s'investir dans une carrière sportive, mais aussi la sous-médiatisation des événements associés constituent un frein à sa croissance.

Les JP de Londres en 2012 ont permis une croissance considérable de l'intérêt porté sur la réalisation de performances sportives par des PSH grâce à une importante stratégie de communication. L'augmentation des moyens investis, les améliorations technologiques apportées aux athlètes et aux infrastructures marquent un tournant vers une meilleure "reconnaissance" de ce type d'évènement [7]. Enfin, alors que l'on dénombrait une participation de 400 athlètes représentant 23 nations en 1960, les derniers Jeux Paralympiques de Rio en 2016 ont accueilli plus de 4000 athlètes pour 159 nations ce qui montre que le mouvement paralympique est en pleine phase ascendante [6].

Cependant, hormis les JP qui ont lieu tous les 4 ans, les autres types d'évènements séduisent moins. Par exemple, en France en 2018, les PSH sont présentes dans 0,9% des programmes sportifs selon les données de la Caisse Nationale de Solidarité pour l'Autonomie (CNSA). Toutefois, des initiatives émergent comme le programme français "Jouons Ensemble" de médiatisation du para-sport, tenu du 17 au 23 mai 2021. Cela a pour but "d'améliorer la représentation du handicap dans les médias et l'accessibilité des programmes" en prévision des JO de Tokyo 2021 de part la diffusion de reportages et de portraits d'athlètes. Les enjeux soulevés sont importants afin de valoriser les performances des para-athlètes internationaux et ainsi de mettre en lumière grâce aux témoignages que le handicap n'est pas un frein à la pratique d'une activité physique et sportive. En effet, les bénéfices du sport sur la santé mentale et physique sont connus et participent à l'autonomisation des PSH mais aussi au regain de confiance et d'estime de soi [8, 9].

Cet intérêt croissant porté sur les athlètes en situation de handicap s'accompagne de nouvelles demandes des fédérations nationales et internationales sur la recherche scientifique. L'objectif est d'apporter des connaissances spécifiques et d'objectiver les performances des athlètes en situation de handicap afin de mieux les accompagner vers leurs objectifs sportifs tout en développant la détection de jeunes talents ainsi que les chances de médailles.

Au niveau international, l'organisation à la tête du mouvement paralympique est le Comité International Paralympique (CIP), créé en 1989. Le besoin d'une unique unité gouvernante pour guider le mouvement a mené à sa création. Originellement, des fédérations internationales avaient été créées pour représenter les différents types de handicap (exemple : Fédération Internationale des Sports pour Personnes Aveugles). La mission du CIP est de "promouvoir les valeurs du paralympisme et de créer des opportunités sportives pour toutes personnes avec un handicap, qu'il soit amateur ou professionnel". Au total, 182 Comités Nationales Paralympiques participent au mouvement. Enfin, l'organisme a aussi la charge de fournir un ensemble de directives à destination des fédérations sportives internationales concernant la mise en place d'un protocole de catégorisation des para-athlètes en fonction de la sévérité de leur handicap.

3.3 Le système de classification

Garantir l'égalité des chances est un enjeu majeur d'attractivité pour toute discipline sportive qui se veut compétitive. En effet, le caractère trop unilatéral d'une confrontation peut être vecteur de désintérêt [10]. Ainsi, tout sport est régi par un système de classification qui peut être basé sur les performances ou encore sur des critères de sélection spécifiques (critères de performances, critères de poids) [10].

Il semble évident qu'un système basé sur la performance ne peut en aucun cas assurer l'égalité des chances des para-athlètes du fait de la multiplicité des degrés de sévérité du handicap. Initialement, le système de classification était basé sur un diagnostic médical de l'individu lésé et ceci indépendamment de la discipline sportive. Par exemple, les individus amputés formaient une catégorie bien distincte des individus ayant des troubles neurologiques. À partir de la fin des années 1970, le diagnostic médical a été remplacé par un diagnostic fonctionnel spécifique à la discipline sportive exercée. Ainsi, un athlète n'est plus classé en fonction de son handicap mais en fonction de l'impact du handicap sur la technicité minimale requise pour une discipline. Ce système fût pour la première fois instauré aux Jeux Paralympiques de Barcelone en 1992 [10].

Chaque sport a son propre système de classification déterminé par les fédérations internationales associées. Celles-ci ont la responsabilité d'allouer une catégorie au sein de laquelle l'athlète peut exercer sa discipline grâce à des "classificateurs". Les classificateurs constituent des groupes d'experts médicaux et sportifs qui déterminent si le handicap de l'athlète le rend éligible pour une catégorie donnée. Ce processus est constitué d'un ensemble d'examen cliniques au cours desquels l'athlète va être évalué sur ses capacités de force musculaire et de mouvement dans l'espace (spécifiques et non-spécifiques à la discipline sportive) [10, 11]. Dans le cas des handicaps dits "progressifs", une nouvelle classification peut être demandée.

Le système de classification est découpé en trois grandes catégories, elles-mêmes découpées en sous-catégories numérotées selon le degré de sévérité du handicap (plus le numéro de la catégorie est faible, plus le degré de sévérité du handicap est élevé) [11]. Le numéro de la classification est précédé d'une lettre pour préciser la nature de la discipline (S pour "swimming" ou encore "T" et "F" en athlétisme pour les épreuves sur piste et terrain respectivement). On retrouve ainsi les catégories suivantes :

- Handicaps physiques (aussi appelés biomécaniques ou neuro-musculo-squelettiques) : cette catégorie regroupe les athlètes touchés par une perte de force musculaire ou de capacité de mouvement, la perte ou dimension anormale d'un ou plusieurs membres, la petite taille, l'hypertonie, l'ataxie ou encore l'athétose.
- Déficients visuels : cette catégorie regroupe les athlètes touchés au niveau de la structure oculaire ou du cortex visuel cérébral entraînant une diminution partielle ou totale, progressive ou non de la vision.
- Déficients intellectuels : cette catégorie regroupe les athlètes qui éprouvent des limites de fonctionnement des capacités mentales et comportementales.

Supposé promouvoir l'équité et attribuer les performances d'un athlète au développement de ses capacités grâce à l'entraînement, le système est soumis à controverse, suscite un manque de confiance en son fonctionnement de la part des différents acteurs et montre certaines limites [12]. Des failles ont pu être identifiées, comme le démontre le scandale autour de la médaille d'Or de l'équipe d'Espagne de Basketball aux JP de Sydney en 2000 où 10 des 12 athlètes ne présentaient aucun handicap. L'actuel système de catégorisation des déficients visuels est fortement remis en cause du fait qu'il se base sur un examen médical commun à toutes les disciplines sportives : les athlètes sont classifiés en fonction de leur acuité et champ visuel uniquement. La capacité à détecter les mouvements ou encore la sensibilité aux contrastes ne sont pas prises en compte alors qu'elles peuvent être des déterminants de la performance [13]. De plus, minimiser l'impact du handicap sur l'issue d'une compétition par l'attribution d'une catégorie constitue une lourde tâche lorsque les tests supposent une coopération totale de l'athlète alors que celui-ci peut intentionnellement aggraver la sévérité de son handicap pour

obtenir une "meilleure" classification [10].

Enfin, le système est complexe et en constante évolution. La recherche scientifique joue un rôle majeur pour (i) développer un ensemble de tests objectifs nécessaires à la classification des athlètes et (ii) permettre de développer un équilibre autour des performances des athlètes en fonction de la sévérité de leur handicap sans pour autant les pénaliser de leurs prouesses [10].

3.4 Objectifs de l'étude

Dans le cadre du programme prioritaire de recherche Paraperf, l'objectif de ce stage est d'analyser les performances des athlètes ayant participé aux grandes compétitions que sont les Jeux Paralympiques et les Championnats du Monde et plus particulièrement d'analyser ces performances en fonction de l'origine du handicap de l'athlète : de naissance (congénital/inné) ou acquis au cours de la vie. Dans un second temps, l'objectif sera d'étudier les trajectoires de performance des athlètes en fonction de l'âge qui est un déterminant majeur de la performance et de l'origine du handicap.

L'analyse des performances sportives en fonction de l'origine du handicap sur plusieurs disciplines sportives n'a pas encore fait l'objet de publication scientifique. Actuellement, une étude s'intéresse à cette problématique en para-natation et para-athlétisme uniquement chez les déficients visuels [14]. Entre autres, les études comparant handicap inné versus acquis en milieu sportif s'intéressent à la psychologie, aux aspirations et motivations de ces deux populations plutôt qu'aux performances (étude en basketball [15], étude en athlétisme [16]).

L'investigation des différences entre les deux groupes étudiés pourrait ainsi permettre d'apporter des connaissances pour améliorer le système de classification mais aussi pour l'accompagnement des athlètes en fonction du stade de leur carrière sportive ou encore l'identification de jeunes talents.

4 Matériels et méthodes

△ : Pour la suite de ce rapport, il est nécessaire de préciser le sens de trois termes utilisés : *discipline*, *évènement* et *trajectoire*. Le terme *discipline* désigne le couple distance + épreuve (Ex : 100m nage libre). Le terme *évènement* désigne l'association d'une épreuve pour une distance pour une classification et sexe donné (Ex : 100m nage libre hommes classification S10). Enfin, une *trajectoire* est définie dans cette étude comme l'ensemble des meilleures performances annuelles d'un athlète dans un évènement précis sur sa durée de suivi (1 perf./âge).

Les analyses ont été réalisées à partir du logiciel R version 3.6.0.

4.1 Origine des données

Les données à disposition ont été extraites de *Gracenote Sports*, qui "recueille, conserve et fournit des données sportives approfondies, notamment les calendriers, les scores, les statistiques et les informations sur les équipes et les joueurs des principales ligues et événements professionnels du monde.". À la suite de cette extraction, non réalisée au cours de ce stage, deux bases de données "Résultats" ont été constituées concernant la natation et l'athlétisme. La base de données natation regroupe les quatre nages : dos, papillon, brasse et nage libre (hors épreuves sur 25 mètres) avec, en plus, l'épreuve de "medley" (ensemble des quatre nages dans une épreuve). La base de données athlétisme regroupe les épreuves de course (allant du 100m au marathon), de saut et de lancer (javelot, disque, massue et poids). Ces bases répertorient l'ensemble des résultats (classement et temps, distances pour les disciplines métriques ou encore points) des athlètes sur différents types de compétition à savoir les championnats d'Europe, championnats du monde (CM) et les Jeux Paralympiques (JP). Ont été retenus uniquement les résultats aux CM et JP entre 2012 et 2019. Les compétitions retenues sont les suivantes :

- JP été : Londres (2012) et Rio (2016)
- CM été : Lyon (2013), Doha (2015), Londres (2017) et Dubaï (2019)

De plus, les bases de "Ranking" sur les disciplines citées ci-dessus sont venues compléter les bases Résultats. Elles regroupent les meilleures performances annuelles des athlètes dans les disciplines auxquelles ces derniers ont concouru (avec compétitions hors JP/CM). Une jointure entre les bases Résultats et Ranking a été réalisée afin de ne conserver uniquement les athlètes présents dans les bases Résultats. Un athlète peut donc avoir plusieurs "meilleures" performances annuelles s'il a concouru dans plusieurs disciplines (ex : 100m nage libre et 100m brasse). Ces deux types de bases vont permettre respectivement d'analyser les performances brutes des athlètes sur les compétitions d'intérêt et de les placer dans un contexte longitudinal. La base Ranking bénéficie d'un recul datant de 2009.

L'échantillon d'athlètes sur lequel repose cette étude est un échantillon d'athlètes sélectionnés parmi une population plus vaste pour participer aux grandes compétitions (CM, JP) du fait de leurs performances.

Un "nettoyage" de ces bases a dû être réalisé. En effet, le nom d'un même athlète constituant les bases de données mises à disposition peut varier, ce qui engendre une part importante de doublons avec les noms avec/sans accents, les noms complets versus noms abrégés (ex : athlètes hispaniques). Il y a donc une part importante de *data management* pour nettoyer les données et permettre ainsi une homogénéisation entre les bases. Un regard attentif doit être porté sur cette problématique afin de ne pas surestimer la variable d'intérêt qui est l'origine du handicap des athlètes. Cette phase a constitué une part importante du temps de travail.

4.2 Construction de la variable d'intérêt : origine du handicap

Dans un premier temps, il était nécessaire de récupérer l'origine du handicap de l'athlète, variable d'intérêt manquante dans les données. Pour cela, une méthode de "web-scraping" a été mise en place

grâce aux bibliothèques `Rvest`, `purrr` et `wdman` du logiciel R. La méthode consiste à parcourir automatiquement un site web et y récupérer sur la page d'intérêt l'information souhaitée afin de la stocker sous forme de variable par la détection du code CSS/HTML associé. L'origine du handicap (en plus d'un texte libre décrivant ce handicap) a donc été récupérée à partir des biographies des athlètes présents dans les bases sur deux sites du CIP. Pour cela, les URLs associées à chaque athlète ont été récupérées automatiquement du fait de leur simplicité de construction : une partie commune + "prénom-nom". Cette première étape a été réalisée sur le premier site du CIP. Si l'information n'était pas disponible sur ce site, une deuxième étape, plus complexe, a dû être considérée. En effet, les URLs sur le second site du CIP ne pouvaient pas être récupérées de manière automatique, une émulation de navigateur (Chrome) a été nécessaire afin de simuler une navigation sur les différents onglets du site considéré dans le but d'accéder à la biographie de l'athlète en question. Cette méthode consiste à rentrer automatiquement le nom de l'athlète dans la barre de recherche du site et simuler un clic sur le bouton "Entrer".

Tous les athlètes ne bénéficient pas de la même qualité biographique sur les sites du CIP. Bien que, comme énoncé précédemment, l'échantillon d'étude est constitué d'athlètes dont le niveau de performance les rend éligibles à la participation aux grands événements sportifs, les informations disponibles sur ces derniers dépendent en majeure partie de leur médiatisation et donc de leurs résultats. Dans le cas où l'origine du handicap est manquante pour un athlète mais que la description de son handicap est disponible, une première imputation « à la main » a été réalisée. Le choix sur l'origine du handicap a été déterminé à partir de mots clefs explicites indiquant la temporalité liée à la déficience (ex : « est né avec... », « a eu ... à 15 ans » etc). Dans les cas où une maladie particulière est citée, des recherches ont été réalisées afin de déterminer si elles entraînent dans tous les cas une déficience dès la naissance ou si, au contraire, la maladie entraîne une déficience tardive ou progressive. Si aucune information de temporalité d'acquisition du handicap peut être déterminée à partir de la description fournie, l'origine du handicap a été considérée comme manquante. Les URLs associées à chaque athlète ont été conservées dans une variable "Source".

Remarque : ces données sont des données publiques déclaratives distribuées sous une licence libre et ouverte.

Une problématique observée autour de ces données est l'absence de standardisation autour de l'origine du handicap. Par exemple, il est commun, notamment chez les déficients visuels, que deux athlètes ayant la même pathologie aient une origine du handicap pour l'un acquise et l'autre congénitale. Comme il a pu être énoncé en introduction, une attention particulière a été portée sur le caractère progressif de maladies congénitales. Il a pu être observé que de nombreuses origines du handicap étaient considérées comme congénitales alors que la description faisait état d'une maladie dégénérative avec un handicap pouvant se déclarer plus tard dans la vie. Pour les cas d'origine congénitale du handicap avec évolution, l'origine du handicap a été considérée comme acquise si un développement normal des fonctions motrices pouvait être supposé, c'est-à-dire si aucun handicap n'a été déclaré jusqu'au 6ème mois post-natal [17, 18]. Le développement de mouvements volontaires chez le nourrisson a lieu autour des 20 semaines après la naissance [17, 18]. On part du principe que le handicap est acquis s'il y a eu un temps de latence sans trouble(s) entre le développement théorique des mouvements moteurs généraux chez le nouveau-né et l'âge de déclaration du handicap dans la description même s'il demeure une incertitude du fait du caractère déclaratif de la description. Sans aucun indicateur de temporalité et en l'absence d'une description explicite du handicap, l'origine était considérée, une nouvelle fois, comme manquante.

Enfin, une variable de sévérité du handicap a été ajoutée en fonction de la classification des athlètes sur les différentes épreuves. Le découpage a été réalisé en fonction des données de la littérature [19, 20] et des effectifs présents dans les classes. En natation, six catégories ont été créées : handicaps physiques sévères, importants, modérés, minimes, déficients visuels et déficients intellectuels. En athlétisme, huit catégories ont été créées : lésés cérébraux, lésés cérébraux fauteuil, tétraplégiques, paraplégiques, déficients membres inférieurs, déficients membres supérieurs, déficients visuels et déficients intellectuels. Les athlètes exerçant en position debout ont été différenciés des athlètes en fauteuil roulant.

Les parcours de vie des athlètes en situation de handicap varient nécessairement en fonction de l'origine du handicap mais aussi de sa sévérité, il est donc important de considérer ces variables [21].

4.3 Indicateur de performance

Concernant les performances, un indicateur a été utilisé : le pourcentage de la meilleure performance. Cet indicateur a déjà été utilisé sur ce type de données [22]. Il existe d'autres indicateurs similaires comme le pourcentage de record olympique [23] ou encore le pourcentage de record du monde [24]. A partir des résultats (temps, métriques etc) réalisés par les athlètes dans chaque discipline et pour chaque classification, on parlera alors d'évènement, la meilleure performance a été récupérée. Chaque résultat pour chaque évènement a donc été comparé à la meilleure performance associée. Pour cela, quand le résultat est un temps, celui-ci est converti en vitesse (mètres/sec) à partir de la formule

$$V(ms) = \frac{distance(m)}{temps(sec)}$$

Pour les résultats métriques, aucune conversion a été appliquée. Le pourcentage de la meilleure performance correspond donc à : meilleure performance = 100% et

$$\%perf_{ik} = \frac{résultat_{ik}}{meilleurrésultat_k} * 100$$

- $\forall i$ allant de 1 au nombre de performances pour chaque évènement.
- k représente l'évènement auquel est associée la meilleure performance.

Cet indicateur normalisé permet des comparaisons inter et intra-classes. Le package `lubridate` a été utilisé pour les conversions de type vitesse.

Remarque : les échantillons d'études étant constitués d'athlètes participant aux JP et CM, le pourcentage de la meilleure performance est, dans une majeure partie des cas, équivalent au pourcentage de record du monde (autre indicateur utilisé dans les études).

4.4 Relation âge et performance : les fonctions Moore & IMAP

L'évolution de la performance en fonction de l'âge est basée sur les travaux de recherches modélisant les processus de croissance pour séries temporelles. Benjamin Gompertz (1779-1865) et Ludwig Von Bertalanffy (1901-1972) sont à l'origine de fonctions mathématiques largement utilisées dans les processus de modélisation de la mortalité humaine et dynamique de population (modèles proies-prédateurs) [25], évolution des tumeurs [26] ou encore concernant la croissance des organismes biologiques [27].

Fonction de Moore (1975) [28]

Dan H Moore a réalisé la première étude modélisant la relation entre la performance sportive et l'âge en athlétisme et plus particulièrement sur des épreuves de course et de lancer. Cette étude avait pour but de déterminer la dynamique de détérioration de la performance en fonction de l'âge, du sexe et du type d'épreuve. L'équation proposée est la suivante :

$$P(t) = a \left(1 - e^{-bt}\right) + c \left(1 - e^{-dt}\right)$$

- $P(t)$ représente la performance à un âge. t
- les coefficients a, b, c, d sont strictement positifs.

L'équation peut être décrite comme la somme de deux fonctions de croissance de Von Bertalanffy où $P(t) = A(t) + B(t)$. $A(t)$ représente le processus exponentiel croissant et $B(t)$ le processus décroissant. La courbe obtenue est en forme de U inversé, elle commence par une phase ascendante forte à la jeunesse du sportif atteignant un maximum puis redescend plus lentement du fait de la diminution

des fonctions motrices et musculaires dûe au vieillissement et tombe à zéro. Pour chaque évènement, le pic exact est calculé et correspond à l'âge où la performance est maximale [29].

Dans cette étude, un cinquième paramètre e a été rajouté à l'équation car à âge nul, la fonction est construite de manière que la performance soit égale à 0. La courbe estimée sur les données passera donc forcément par 0. Les données à disposition sont éloignées du point (0,0), le cinquième paramètre e permet donc de décaler la courbe afin de mieux s'adapter aux données car il évite ce passage par l'origine.

Cette équation développée par Moore a été construite d'un point de vue statistique et empirique. En effet, les paramètres n'incluent aucune hypothèse biologique ou physiologique.

Fonction Integrative Model of Age-Performance, IMAP (2019) [30]

La fonction IMAP (développée à l'IRMES) est basée sur une approche déterministe de dynamique de population cellulaire, sénescence réplivative¹ et perte de fonctionnalité. Elle constitue une première approche pour mettre en parallèle les dynamiques de performance avec les dynamiques cellulaires. L'objectif étant d'inclure les aspects biologiques dans la modélisation de la relation asymétrique entre l'âge et la performance en se basant sur l'unité biologique que représente la cellule. Soit le modèle populationnel :

$$\begin{cases} \frac{dN(t)}{dt} = \alpha(t)N(t) \\ P(t) = \beta(t)N(t) \end{cases}$$

- $N(t)$, représente une population de cellules à un instant t .
- $\alpha(t) = \alpha_0 e^{-\alpha_r t}$, représente le taux de duplication par cellule par unité de temps.
- $\beta(t) = \beta_0 (1 - e^{\beta_r (t-t_d)})$, représente le déclin associé à la perte fonctionnelle des cellules.

Ci-dessous la résolution du problème de Cauchy :

$N'(t) = \alpha(t)N(t)$ admet une solution générale de la forme $N(t) = C e^{\int_0^t \alpha(t) dt}$ où C est une constante.

$$\text{On a donc : } \int_0^t \alpha(t) dt = \frac{\alpha_0}{\alpha_r} (1 - e^{-\alpha_r t}) \Rightarrow N(t) = C \cdot \frac{\alpha_0}{\alpha_r} (1 - e^{-\alpha_r t})$$

$$\text{Soit la condition initiale } N(t_0) = N_0 \Rightarrow C = N_0 \cdot \frac{\alpha_r}{\alpha_0} \text{ d'où } N(t) = N_0 \cdot \frac{\alpha_0}{\alpha_r} (1 - e^{-\alpha_r t})$$

On a donc la formulation IMAP suivante :

$$P(t) = \beta_0 N_0 \cdot e^{\frac{\alpha_0}{\alpha_r} (1 - e^{-\alpha_r t})} \cdot (1 - e^{\beta_r (t - t_d)})$$

Où :

- $P(t)$, la performance à un âge t .
- $\beta_0 N_0$, la performance à t_0 l'âge d'entrée dans la compétition.
- α_0 , la performance maximale théorique (pic de performance).
- α_r , la vitesse de croissance au pic de performance, plus le coefficient est faible plus la croissance est rapide. Le pic de performance est paramétré par α_0 , α_r et β_r et ces différents paramètres modulent l'âge au pic.
- β_r , la pente du déclin, plus le coefficient est élevé, plus le déclin est tardif. Le coefficient a pour effet de repousser la date du déclin et de l'âge où la performance est maximale. Il influence aussi le niveau de la performance au pic (plus β_r est grand plus le pic augmente).
- t_d , l'âge au décès (âge auquel $P(t)=0$).

1. Campisi J. The biology of replicative senescence. European Journal of Cancer. 1 avr 1997 ;33(5) :703-9.

Remarque : par simplification, les paramètres IMAP seront respectivement appelés a, b, c, d et e .

Remarque 2 : il est difficile de donner une interprétation unique aux paramètres qui, en fait, interagissent tous ensembles.

Remarque 3 : ces formules suggèrent l'utilisation de maximums de performance.

Les paramètres des fonctions Moore et IMAP sont estimés par la méthode des moindres carrés. La fonction `R` utilisée est explicitée en Figure 6. Cette méthode a pour but de déterminer les coefficients qui minimisent la somme des carrés des erreurs du modèle employé (les erreurs sont déterminées à partir de la distance qui sépare une valeur observée de sa valeur prédite). La fonction `nlminb` utilisant la méthode "quasi-Newton" permet de réaliser l'optimisation. Des contraintes sur les paramètres doivent être ajoutées comme il a été expliqué précédemment (paramètres positifs ou encore âge au décès borné). Enfin, cette méthode permet de déterminer notamment un âge au pic de performance.

4.5 Analyses statistiques

4.5.1 Statistiques descriptives

Dans un premier temps, les différents échantillons sur lesquels se base cette étude ont été décrits afin d'en identifier le contenu. Les résultats seront présentés en fonction de l'origine du handicap qui est donc une variable à trois modalités : congénital (inné), acquis ou manquant.

Les variables qualitatives ont été comparées par un test du Chi-deux, les résidus ajustés ou résidus d'Haberman ont été utilisés afin d'identifier le sens des interactions significatives (sous H_0 , l'absence de lien, les résidus suivent asymptotiquement une loi normale centrée réduite).

En l'absence des hypothèses de normalité des distributions étudiées et d'égalités des variances (test de Levene), les variables quantitatives (% de la meilleure performance notamment) ont été comparées par un test de Wilcoxon ou de Kruskal-Wallis (KW) en fonction du nombre de k échantillons comparés. Le raisonnement est le suivant : en supposant k échantillons aléatoires indépendants $(X_{1,1}, \dots, X_{1,n1}), \dots, (X_{k,1}, \dots, X_{k,nk})$ et k séries d'observations $(x_{1,1}, \dots, x_{1,n1}), \dots, (x_{k,1}, \dots, x_{k,nk})$. Soit $L_i(X)$ la loi de la variable aléatoire X dans la sous-population d'ordre i avec $1 \leq i \leq k$, alors, les hypothèses testées sont les suivantes :

- H_0 : $L_1(X) = \dots = L_i(X) = L_k(X)$. Les lois sont identiques dans les k sous-populations étudiées. Il n'y a pas de différence de performance entre les groupes étudiés.
- H_1 : Au moins deux lois diffèrent. Au moins deux groupes présentent une différence de performance.

Dans le cas d'un test de KW significatif, il est intéressant d'identifier quelles paires de groupes présentent des différences. Un test post-hoc de Dunn avec ajustement de Bonferroni a donc été réalisé.

Remarque : du fait du nombre important d'observations, une légère différence peut être considérée significative par le test.

L'ensemble des performances sur la base Résultats ont été comparées en fonction de l'origine du handicap puis dans les groupes de sévérité du handicap ou encore selon les types d'épreuves chez les hommes et chez les femmes séparément. Une limite à cela est la participation répétée des athlètes aux différentes compétitions étudiées sur le même évènement (ex : 100m dos S13 2012, 2016 et 2019). En effet, Il faut tenir compte du poids que représentent ces athlètes dans l'analyse afin de ne pas accorder trop d'importance à un athlète qui aurait, par exemple, d'excellentes performances récurrentes dans un même évènement (en plus de la corrélation entre les observations). Ainsi, plutôt que d'analyser l'ensemble des performances, une analyse sur les "performeurs" a été réalisée. L'analyse sur des performeurs s'intéressent à la meilleure performance réalisée dans un évènement par un athlète. Un athlète qui aurait plusieurs performances sur un même évènement en aura alors plus qu'une par évènement réalisé. Les performances d'un même athlète dans deux disciplines différentes sont considérées indépendantes.

Remarque : les analyses longitudinales n'ont pas été réalisées sur cette base car elle ne retrace pas les carrières des athlètes considérés mais uniquement les performances aux JP/CM.

4.5.2 Analyses longitudinales

△ : Pour ces analyses, ont été retenus uniquement les athlètes pour lesquels l'origine du handicap ne présente pas de valeur manquante. De plus, le choix a été fait de sélectionner une sous-population de ces individus, à savoir ceux ayant au moins 7 mesures de meilleure performance annuelle [24]. Il est plausible de considérer que sur 7 ans de suivi et plus, les performances des athlètes observent une tendance assimilable à une trajectoire de performance asymétrique (type Moore/IMAP).

Les analyses longitudinales ont été réalisées en deux temps.

Dans un premier temps, les fonctions de Moore et IMAP ont été utilisées pour modéliser les trajectoires maximales de performance des athlètes en fonction de l'âge. Ainsi, à partir des maximums de performance (i.e % meilleure perf.) à chaque âge, chez les hommes et chez les femmes, les trajectoires de performances ont été estimées en fonction de l'origine du handicap, initialement sur l'ensemble des maximums de performance de l'échantillon total puis dans chaque groupe de sévérité du handicap séparément. Un âge au pic a été calculé au maximum de la fonction estimée. La modélisation par la fonction IMAP nécessite d'initialiser le paramètre "e" (l'âge au décès) à un âge maximal. Le paramètre a donc été initialisé à 122 ans, ce qui représente la valeur de longévité maximale répertoriée à nos jours (Jeanne Louise Calment).

Les qualités d'ajustement des trajectoires de Moore et IMAP aux maximums de performances ont été évaluées à partir du coefficient de détermination R^2 . Un R^2 proche de 1 définit une bonne qualité d'ajustement du modèle de régression aux données. La formule utilisée est la suivante :

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

Où :

- \hat{Y}_i correspond à la valeur prédite du % meilleure perf. pour une performance i
- Y_i correspond à la valeur observée du % meilleure perf. pour une performance i
- \bar{Y} correspond à la moyenne empirique du % meilleure perf.

En parallèle des trajectoires de performances, une approche par modèle mixte a été considérée.

Le modèle mixte est particulièrement adapté à l'étude des données répétées dans le temps afin de suivre l'évolution d'un marqueur tout en tenant compte de la variabilité inter et intra individuelle. En effet, les observations d'un même individu ne sont pas indépendantes, elles constituent un "groupe". On parle alors de modèle à effets fixes et à effets aléatoires. Afin de prendre en compte la multi-participation des athlètes aux épreuves sportives (Ex : course & brasse en natation ou encore 100m & 400m en sprint sur piste), une trajectoire individuelle correspond à la trajectoire d'un athlète dans un événement donné. Ainsi, l'individu i communément appelé dans les notations usuelles, est représenté ici par le couple événement/athlète. Un même athlète peut donc avoir plusieurs trajectoires dans la modélisation, considérées indépendantes. Le choix de regrouper les performances des athlètes en prenant la meilleure performance à chaque âge quel que soit l'événement n'a pas été fait car une performance dans un événement peut être totalement différente d'une performance dans un autre événement pour un même athlète du fait de la stratégie et des exigences variées d'une discipline à l'autre (par exemple en natation, un athlète peut être bon sur les épreuves de dos mais moins performant sur les épreuves de brasse qui sollicitent différemment les membres).

Après avoir brièvement rappelé la structure d'un modèle linéaire mixte, il sera présenté ci-après la structure du modèle non-linéaire.

Le modèle linéaire mixte

Le modèle s'écrit sous formulation générale de la manière suivante :

$$Y_{ij} = X_{ij}^\top \beta + Z_{ij}^\top \eta_i + e_{ij}, \forall i = 1 \dots N, \forall j = 1 \dots n_i$$

Où :

- Y_{ij} correspond à la $j^{\text{ème}}$ observation de l'individu i .
- N correspond au nombre total d'individus.
- n_i correspond au nombre d'observations du sujet i .
- X_{ij} est le vecteur des p variables explicatives du modèle. β correspond au vecteur de dimension p des coefficients estimés pour les effets fixes. Le couple $X_{ij}^\top \beta$ constitue les effets populationnels, on parle de variation de la variable réponse Y en moyenne.
- Z_{ij} est sous-vecteur de X_{ij} de dimension q avec $q < p$. η_i correspond au vecteur de dimension q des effets aléatoires. Le couple $Z_{ij}^\top \eta_i$ correspond aux déviations individuelles par rapport à la moyenne.
- $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ et $\eta_i \sim \mathcal{N}(0, \Omega)$ avec Ω matrice de variance-covariance des effets aléatoires.
- Les erreurs sont supposées indépendantes : $e_{ij} \perp e_{ij'}$.
- Les erreurs sont supposées indépendantes des effets aléatoires : $e_{ij} \perp \eta_i$.

Le modèle non-linéaire mixte

La dynamique d'un processus biologique, comme l'évolution de la performance en fonction de l'âge, est souvent non-linéaire et peut être modélisée par un modèle à effets aléatoires. Le package SAEMIX² du logiciel R a été utilisé pour cette analyse. Il a été développé et implanté dans le logiciel MONOLIX utilisé pour l'estimation de modèles en pharmacocinétique basés sur des équations différentielles. Ce package montre de bonnes propriétés de convergence comparé aux packages nlme et lme4 et permet de quantifier l'impact de l'origine du handicap sur les paramètres des fonctions Âge/Performance. Le modèle général est le suivant :

$$Y_{ij} = f(x_{ij}, \psi_i) + g(x_{ij}, \psi_i, \xi) \varepsilon_{ij}, \forall i = 1 \dots N, \forall j = 1 \dots n_i$$

Où :

- x_{ij} correspond à la variable de régression à savoir le temps (âge dans l'étude).
- ψ_i correspond au vecteur des paramètres individuels, à estimer et initialisé par la méthode des moindres carrés citée en 4.4. $\psi_i = h(C_i \mu + \eta_i)$ avec
 - h est une transformation du vecteur gaussien. Il est possible de spécifier 3 transformations, à savoir : log-normale, logit ou probit (par défaut normale).
 - C_i est un vecteur de taille p le nombre de covariables intégrées dans le modèle.
 - μ le vecteur des effets fixes.
 - $\eta_i \sim \mathcal{N}(0, \Omega)$, les effets aléatoires individuels.
- g définit la structure des erreurs du modèle. Trois choix sont possibles pour cette structure : constante, proportionnelle ou combinée. *Un quatrième choix, exponentielle, est possible mais nécessite une transformation des données.*
 - ξ est un vecteur de paramètres associés au choix de la structure des erreurs. $\xi = a$ pour une structure des erreurs constante (où $g = a$), $\xi = b$ pour une structure proportionnelle (où $g = bf$), $\xi = (a, b)$ pour une structure combinée (où $g = a + bf$).
 - $\varepsilon_{ij} \sim \mathcal{N}(0, 1)$ constitue les erreurs du modèle.
- Les erreurs sont supposées indépendantes des effets aléatoires : $\varepsilon_{ij} \perp \eta_i$.

2. Comets E, Lavenu A, Lavielle M. Parameter Estimation in Nonlinear Mixed Effect Models Using saemix, an R Implementation of the SAEM Algorithm. Journal of Statistical Software. 29 août 2017 ;80(1) :1-41.

L'objectif étant d'estimer le vecteur $\theta=(\mu,\Omega,\xi)$. Cela est réalisé par maximisation de la vraisemblance par l'algorithme EM.

Les spécificités techniques du package dépassent largement le cadre de cette étude mais sont bien documentées.

Stratégie de modélisation

Dans un premier temps, le "meilleur modèle vide" a été investigué, c'est-à-dire avec uniquement la relation variable réponse \sim temps. La stratégie a été réalisée en considérant la modélisation du temps par la fonction de Moore puis par la fonction IMAP. La procédure de sélection de modèle réalisée au cours de cette analyse est la suivante :

- choix de la distribution des paramètres : normale, log-normale, logit ou probit.
- choix de la structure des effets aléatoires : afin de vérifier la convergence des modèles, les effets aléatoires ont été introduit un à un, puis l'ensemble des combinaisons possibles ont été testées. Le choix de la structure des effets aléatoires se base sur les critères *AIC* et *BIC* qui sont à minimiser.
- choix de la structure de corrélation des effets aléatoires : à partir de la structure des effets aléatoires, différentes structures de corrélation ont été testées. La structure retenue est celle minimisant les critères *AIC* et *BIC*.
- choix de la structure des erreurs, 4 choix possibles : constantes, combinées, proportionnelles et exponentielles.
- sur le modèle "vide" optimal sélectionné, une analyse univariante a été réalisée. Pour chaque covariable intégrée dans le modèle, les effets fixes sur les paramètres ont été testés. Ont été retenus dans un premier temps uniquement les effets fixes présentant une *p - valeur* $\leq 25\%$.
- une analyse multivariante a été réalisée couplée à une procédure de sélection pas à pas descendante.
- l'adéquation du modèle final a été investiguée avec les différents graphes de diagnostic : distribution des résidus et des effets aléatoires, prédictions vs observations, visual predictive check³, fit individuels.

A partir des estimations des 5 paramètres individuels avec le modèle mixte réalisé, il est possible de créer une matrice de paramètres de taille $(n \times k)$ où n =nombre de trajectoires par évènement/athlète et $k=5$ le nombre de coefficients (*a,b,c,d* et *e*). Les trajectoires individuelles peuvent alors être tracées. L'optimisation des fonctions liées aux trajectoires individuelles permet de quantifier un âge au pic de performance ainsi qu'un maximum de performance théorique. Ainsi, les âges au pic et les maximums de performance individuels ont été comparés en fonction de l'origine du handicap puis dans chaque catégorie de sévérité du handicap par un test de Wilcoxon.

3. Mats O. Karlsson, Nick Holford. A Tutorial on Visual Predictive Checks. Disponible à l'url : <https://www.page-meeting.org/?abstract=1434>

5 Résultats

Les bases utilisées en natation et athlétisme sont les bases "Résultats" et "Ranking" (Figure 1). L'objectif de ce travail étant de s'intéresser aux performances des athlètes aux JP et CM, 10 038 et 11 608 performances sont retenues pour chacune de ces deux grandes disciplines sur les bases Résultats. Les performances ne remplissant pas certains critères sont exclues de l'analyse (disqualifications, résultats non-valides). Les déficients intellectuels sont exclus des échantillons d'étude du fait du pourcentage élevé de valeurs manquantes sur l'origine du handicap (~75%). L'échantillon total en natation est donc composé de 9 332 performances, soit 35% de l'échantillon initiale, et 11 005 en athlétisme (54%).

Concernant les bases Ranking, 32 730 performances constituent les trajectoires de performance en natation et 28 781 en athlétisme.

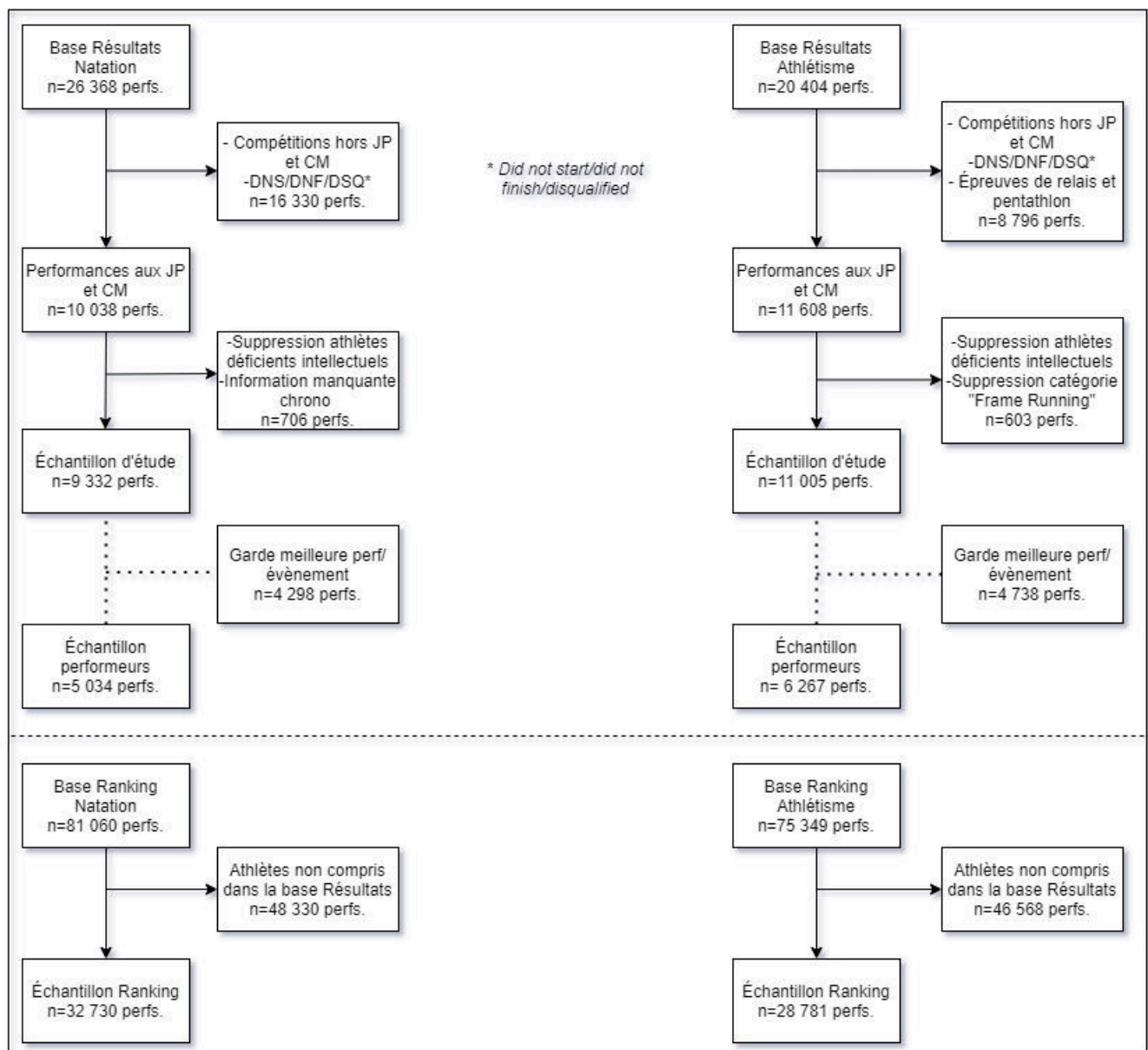


FIGURE 1 – Diagramme de flux concernant la sélection des performances en natation et athlétisme.

5.1 Données brutes, bases Résultats

Les caractéristiques individuelles des athlètes retenus dans ces analyses sont à retrouver Table 1 pour la natation et Table 4 pour l'athlétisme.

Sur l'échantillon "natation", 1 239 athlètes distincts correspondent aux 9 332 performances. Parmi ces athlètes, 41% ont une origine du handicap congénitale et 37% acquise, ce qui constitue un peu plus de trois quarts de l'ensemble des athlètes considérés. Il y a 22% de données manquantes. L'échantillon est légèrement plus composé d'hommes que de femmes (57%), ce qui se retrouve dans les catégories d'origine du handicap. De plus, l'origine du handicap est indépendante du sexe ($p=0,25$). La sévérité du handicap semble liée à son origine ($p<0,001$). La moitié des athlètes qui ont une origine du handicap congénitale ont une sévérité du handicap modérée à minime avec une sur-représentation des handicaps minimes. La catégorie des déficients visuels est sur-représentée chez les athlètes qui ont un handicap acquis au cours de la vie (un quart) mais est aussi la catégorie la plus pourvoyeuse de données manquantes (30%). Un peu plus de la moitié des athlètes ont participé à la fois aux JP et aux CM. Les athlètes qui ont un handicap de naissance sont significativement plus jeunes (28 ± 7 ans) que ceux l'ayant acquis (32 ± 9 ans).

TABLE 1 – Caractéristiques des athlètes en natation en fonction de l'origine du handicap. Base Résultats (n=1239).

Variables	Origine du handicap									p^3	
	Congénital (n=502)			Acquis (n=456)			Manquant (n=281)				
	n	% ¹	Moy(ET) ²	n	% ¹	Moy(ET) ²	n	% ¹	Moy(ET) ²		
Sexe										<i>ns</i>	
	hommes	271	54,0		268	58,8		165	58,7		
	femmes	231	46,0		188	41,2		116	41,3		
Sévérité										***	
	sévère	73	14,5		89	19,5		46	16,4		
	importante	109	21,7		84	18,4		52	18,5		
	modérée	120	23,9		89	19,5		57	20,3		
	minime	133	26,5		74	16,2		41	14,6		
	visuelle	67	13,3		120	26,3		85	30,2		
Discipline⁴										<i>ns</i>	
	nage libre	424	28,6		368	28,6		221	32,4		
	brasse	233	15,7		224	17,4		120	17,6		
	dos	291	19,7		270	21,0		147	21,6		
	papillon	244	16,5		204	15,8		95	13,9		
	medley ⁵	288	19,5		222	17,2		99	14,5		
Compétition										***	
	JP	63	12,5		91	20,0		61	21,7		
	CM	171	34,1		115	25,2		132	47,0		
	les deux	268	53,4		250	54,8		88	31,3		
Âge⁶ (années)		501		27,5(6,8)	456		32,1(9,0)	280		29,1(7,7)	***
Nb.perf/athlète		502		8,5(6,6)	456		7,8(6,2)	281		4,9(3,9)	***

1 : Les pourcentages sont présentés en colonne. 2 : Moyenne(Ecart-Type).

3 : seuil $\alpha=5\%$. * = $<0,05$, ** = $<0,01$, *** = $<0,001$, ns=non significatif.

4 : Un même athlète peut participer à plusieurs épreuves.

5 : L'épreuve de Medley est constituée des 4 nages (ou 3 en fonction de la sévérité du handicap, le papillon étant retiré pour les handicaps les plus sévères).

6 : L'âge ici correspond à l'âge actuel pour permettre une première comparaison. Une performance est associée à un âge.

En athlétisme, l'échantillon d'étude est constitué de 2 771 athlètes distincts pour 11 005 performances. Parmi ces athlètes, 42% ont un handicap acquis au cours de la vie et 26% de naissance. Les données manquantes constituent ici 32% de l'échantillon. Les athlètes masculins sont plus nombreux (57%) et sont sous-représentés dans le groupe de handicap congénital et sur-représentés dans le groupe "manquant". Huit catégories de sévérité du handicap sont considérées. Les athlètes cérébraux lésés constituent un quart des athlètes ayant un handicap de naissance et 50% des athlètes ayant un handicap acquis au cours de la vie sont atteints de paraplégie ou de déficience visuelle. La majorité des athlètes concourent en position debout quelle que soit l'origine du handicap. Les épreuves d'athlétisme sont divisées en trois catégories : course, lancer et saut. Il n'y a pas de différence significative entre les participations à ces trois épreuves en fonction de l'origine du handicap mais les épreuves de course sont les plus pourvoyeuses de performances. Les athlètes ayant un handicap de naissance sont significativement plus jeunes (29 ± 7 ans) que ceux l'ayant acquis au cours de leur vie (36 ± 9 ans).

L'analyse sur les performeurs s'intéresse aux meilleures performances dans chaque évènement auxquels les athlètes ont concouru sur les compétitions des JP et CM entre 2012 et 2019.

L'analyse en natation ne montre pas de différence de pourcentage de la meilleure performance entre un handicap de naissance ou acquis au cours de la vie, que ce soit chez les hommes ou chez les femmes. Le groupe des données manquantes ont des performances significativement plus faibles quel que soit le sexe (Figure 7). En étudiant les performances en fonction des différents groupes de sévérité du handicap, seulement les hommes ayant un handicap déterminé "important" ainsi que les déficients visuels présentent une différence significative entre les origines innées versus acquises avec une différence de médiane de performance de respectivement, 2,6% et 2,0% en faveur du groupe acquis (Figure 2).

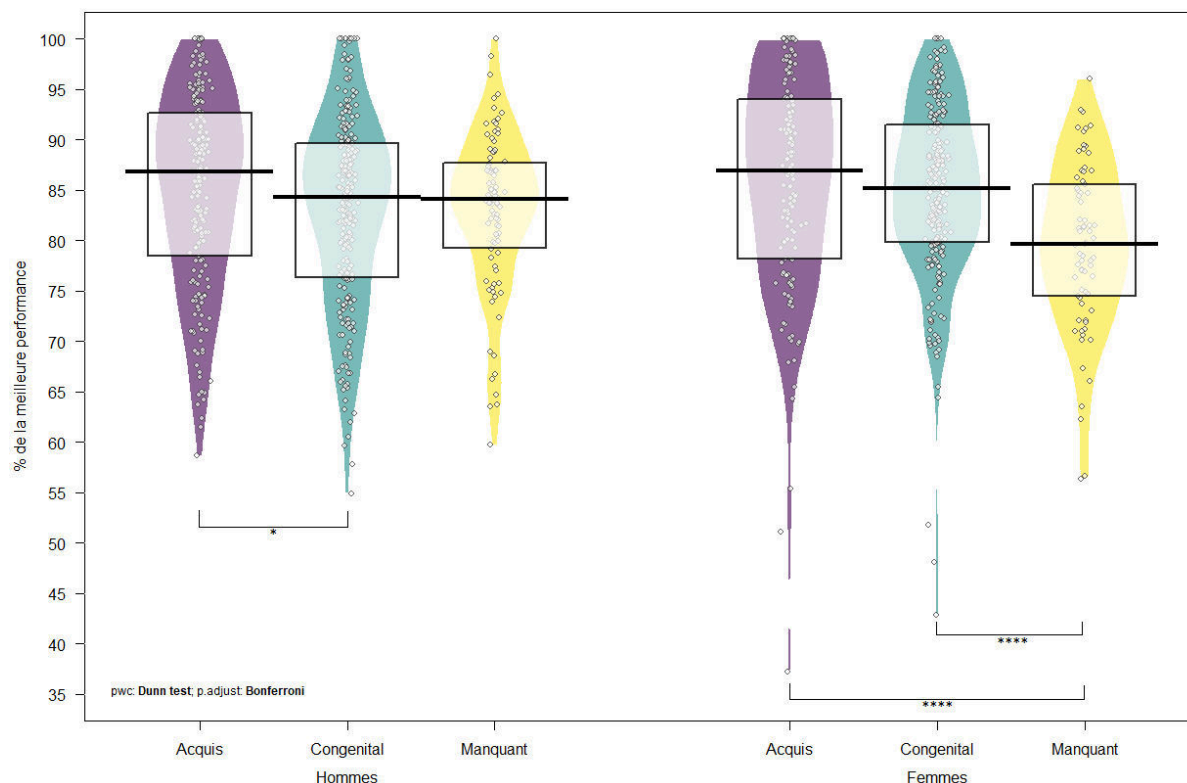


FIGURE 2 – Distribution de % de la meilleure performance en fonction de l'origine du handicap pour chaque sexe en natation. Handicap important (n=986).

Quelle que soit la sévérité du handicap, les athlètes pour lesquels l'origine du handicap est manquante détiennent une médiane de performance plus faible ou dans une moindre mesure équivalente à ce qui peut être observé sur les données complètes (différence allant jusqu'à 10% chez les femmes ayant un handicap sévère). Enfin, concernant les types d'épreuves étudiés à savoir la nage libre, le dos, la brasse, le papillon ou encore les quatre nages (medley), aucune différence de performance n'est observée entre les groupes inné et acquis quel que soit le sexe (Figure 8 et Figure 9). Toutefois, le groupe acquis présente dans la grande majorité des cas des médianes de performance supérieures au groupe congénital.

Remarque : en athlétisme, les athlètes de petite taille n'ont pas été pris en compte dans les analyses car cette catégorie contient forcément des athlètes ayant un handicap congénital.

L'athlétisme montre des résultats différents. Concernant l'ensemble des performeurs observés séparément pour chaque sexe, les hommes ayant un handicap de naissance présentent des performances significativement supérieures vis-à-vis de ceux ayant un handicap acquis au cours de la vie avec une différence de médiane de performance égale à 1,8%. Quel que soit le sexe, les athlètes pour lesquels aucune donnée n'est disponible concernant l'origine du handicap présentent des médianes de performance significativement plus faibles que les groupes inné et acquis allant de 2 à 4% de différence (Figure 10). L'étude en fonction de la sévérité du handicap va dans le sens des résultats sur l'ensemble des performeurs. Les hommes lésés cérébraux, mais aussi ceux présentant une déficience aux membres inférieurs

ou supérieurs d'origine congénitale ont des médianes de performance significativement plus élevées comparées aux handicaps acquis avec 2,5 à 6% de différence (Figure 11). Les catégories qui viennent d'être citées voient les athlètes évoluer debout. Concernant les sévérités de handicap nécessitant l'utilisation d'un fauteuil (cérébraux lésés fauteuil, paraplégiques, tétraplégiques), les hommes atteints de paraplégie ou tétraplégie présentent une différence significative de performance en fonction de l'origine du handicap : les athlètes paraplégiques ou tétraplégiques d'origine congénitale présentent une médiane de performance plus élevée de respectivement 2,4% et 3,9% (Figure 12). Chez les femmes aucune différence n'est observée entre les groupes congénitaux versus acquis. Enfin, les résultats concernant les types d'épreuves étudiés, à savoir la course, le saut ou encore le lancer, sont présentés selon la position exercée (debout ou fauteuil). Ils montrent que les performeurs hommes ayant un handicap d'origine congénitale performant significativement sur les épreuves de course et de lancer debout (Figure 3). Chez les femmes, les performeuses aux épreuves de course et lancer fauteuil ayant un handicap acquis performant mieux que celles ayant un handicap inné.

Aucune différence de performance n'est observée sur les épreuves de saut.

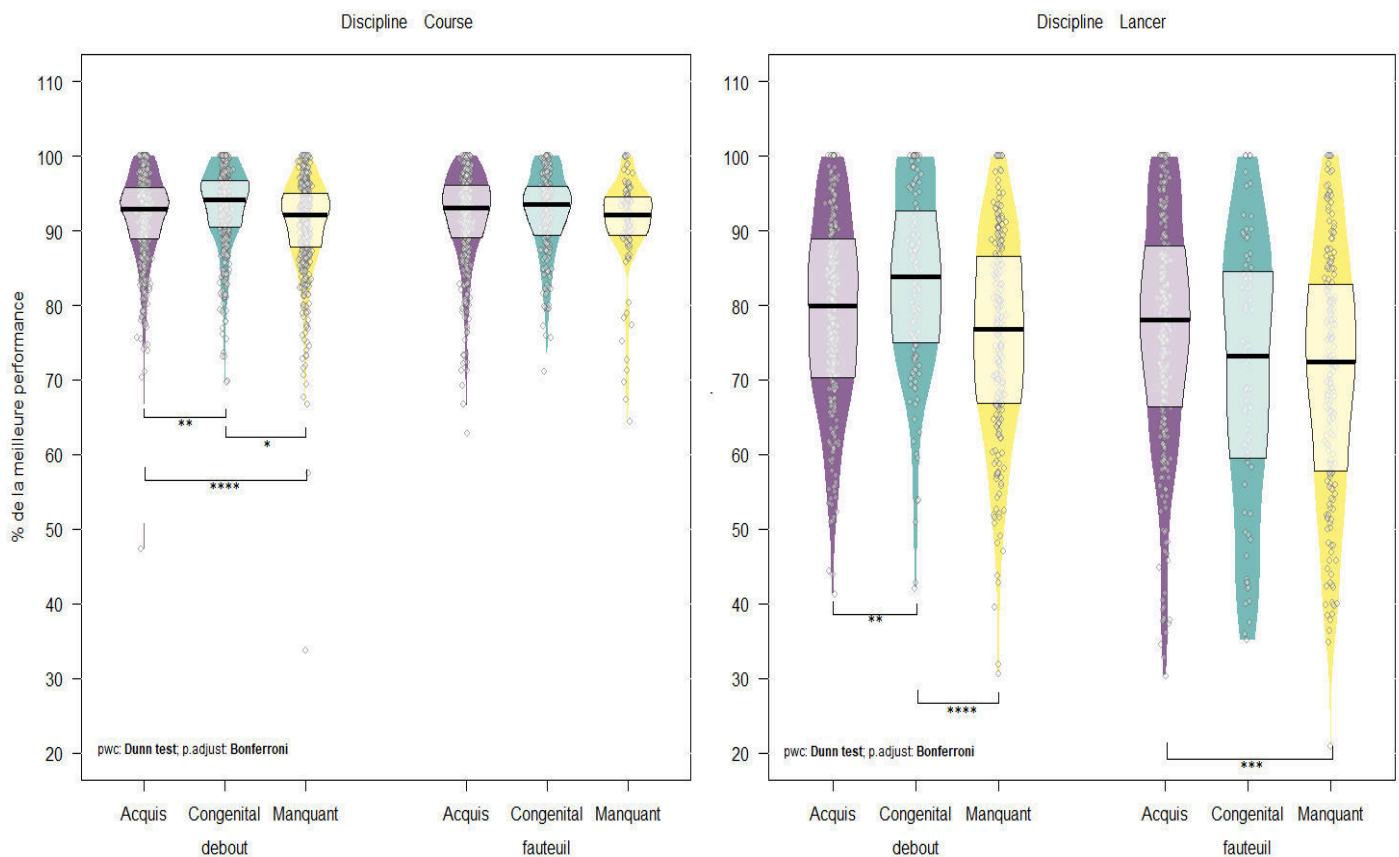


FIGURE 3 – Distribution de % de la meilleure performance en fonction de l'origine du handicap et du type d'épreuve. Athlétisme, hommes (n=3481).

Deux tendances apparaissent sur ces premiers résultats où les disciplines en milieu aquatique se différencient des disciplines terrestres. En effet, d'une part, en natation, les athlètes ayant un handicap acquis au cours de la vie semblent performer mieux que leurs homologues ayant un handicap inné et d'autre part, en athlétisme, la relation semble inverse sur les différences observées. Dans la grande majorité des cas, les athlètes pour lesquels l'origine du handicap n'est pas renseignée ont des médianes de performance plus faibles. Une hypothèse se vérifie concernant la qualité d'observance d'un athlète (disponibilité de l'information sur l'origine du handicap) qui dépendrait en partie de ses résultats.

5.2 Données longitudinales, bases Ranking

Les caractéristiques des données pour les analyses longitudinales en natation et athlétisme sont résumées Table 5, Table 6, Table 7 et Table 8. Sur l'ensemble des trajectoires de performances entre les hommes et les femmes, les athlètes qui ont un handicap d'origine congénitale sont plus jeunes que ceux ayant un handicap acquis. 50% des échantillons de natation et d'athlétisme ont 4 à 7 ou plus mesures de performances par trajectoire.

La modélisation des trajectoires de performance maximale en fonction de l'âge par les fonctions de Moore et IMAP mène aux résultats suivants. Au sens du R^2 , la fonction de Moore offre dans la quasi totalité des observations une meilleure adaptation aux données des maximums de performance quel que soit le sexe, l'origine du handicap, sa sévérité ou encore le type de sport considéré.

Les âges au pic de performance estimés en natation sont plus jeunes chez les athlètes ayant un handicap d'origine congénitale par rapport au groupe acquis quel que soit le sexe ou le choix de la fonction de relation (20,8 vs 23,6 ans pour les hommes et 19,2 vs 21,4 ans pour les femmes avec Moore). Les athlètes ayant un handicap sévère d'origine congénitale (Figure 4) ont la plus importante différence d'âge au pic (8,6 ans chez les hommes et 9,6 ans chez les femmes) par rapport au groupe acquis et cet écart tend à diminuer en même temps que la sévérité du handicap (différence moyenne de 2,9 ans sur les autres catégories). De plus (sauf handicap important), les âges au pic de performance sont plus précoces à mesure que le handicap diminue pour les handicaps physiques. Cette tendance ne se vérifie pas forcément chez les femmes mais ces dernières ont des âges au pic plus jeunes, à observations équivalentes (sévérité et origine du handicap), que les hommes (Table 9).

En athlétisme, et plus particulièrement sur les épreuves de course et lancer, les âges au pic de performance estimés sont aussi plus jeunes concernant le groupe inné chez les hommes et chez les femmes (respectivement 25,9 vs 32,9 ans et 18,5 vs 26,1 ans). Concernant les épreuves de saut, la relation est plutôt inverse (Table 10, Table 11 et Table 12). Les âges au pic en fonction de l'origine du handicap et de sa sévérité en athlétisme sont plus élevés comparés à la natation avec des âges se situant aux alentours des 30 ans.

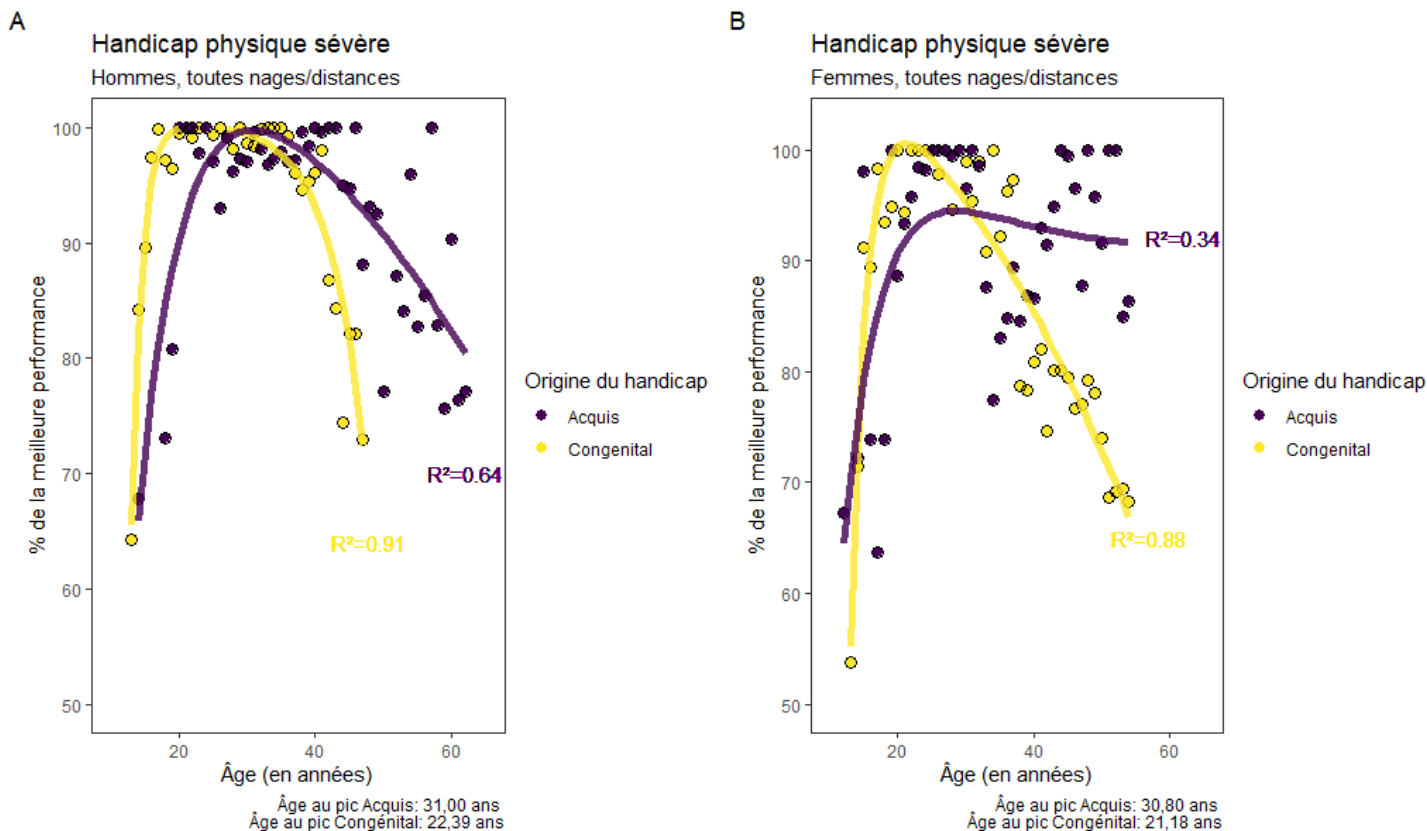


FIGURE 4 – Exemple de trajectoires de performance maximale à chaque âge en natation.

Les femmes observent majoritairement des âges au pic plus précoces que les hommes. Dans la quasi-totalité des trajectoires observées, les athlètes ayant un handicap de naissance voient leurs courbes de performance maximale décliner plus tôt que le groupe acquis.

Bien que récurrente dans les résultats de la littérature modélisant âge et performance, cette méthode ne tient pas compte des trajectoires de performance individuelles et présente ses limites.

La modélisation par modèle mixte s'avère donc intéressante.

Résultats modèles non-linéaires mixtes

Le package SAEMIX n'étant pas usuel, un exemple de code commenté est fourni Figure 15 afin de montrer comment est défini le modèle. Les variables catégorielles à plus de deux modalités n'étant pas implémentées dans le package, les covariables considérées sont l'origine du handicap, le sexe, l'âge d'entrée dans la compétition et, uniquement pour l'athlétisme (course et lancer), la position (debout/fauteuil).

L'estimation rendue par le modèle utilisant la fonction de Moore ne respecte pas les hypothèses sur les paramètres (paramètre estimé d négatif, écarts-types démesurés et fits individuels éloignés des observations). Les résultats sont présentés à partir de la relation âge/performance modélisée par IMAP qui pour rappel est construite sur la base de processus biologiques contrairement à Moore.

L'équation du modèle est la suivante :

$$\widehat{P_i(t)} = (a + a_i) \cdot e^{\frac{(b+b_i)}{(c+c_i)}(1-e^{-(c+c_i)Age_{ij}})} \cdot (1 - e^{(d+d_i)(Age_{ij}-(e+e_i))}) + g(x_{ij}, \psi_i, \xi)\varepsilon_{ij}$$

avec avec a, b, c, d les effets fixes et a_i, b_i, c_i, d_i et e_i les effets aléatoires de l'évènement/athlète i .

Natation

11 876 performances permettent la modélisation de 1 400 trajectoires (≥ 7 mesures annuelles) soit 15% des trajectoires de l'échantillon total (Figure 13).

Les critères AIC et BIC sont minimisés avec des effets aléatoires corrélés considérés sur les paramètres b, c, d et e avec une transformation log-normale. Cela considère que l'ajout d'une déviation individuelle par rapport à la valeur moyenne initiale de performance n'est pas nécessaire. De même, le meilleur modèle considère une structure des erreurs homoscedastique où $Var(\varepsilon_{ij}) = a^2$ ($BIC=69\,644,64$ modèle par défaut $\Rightarrow 63\,584,10$ modèle "vide" final).

TABLE 2 – Résultats modèle mixte IMAP. Analyse univariable avec l'origine du handicap. Natation.

Paramètres	Effets fixes (ET% ¹)	Variance effets aléatoires (ET% ¹)	p
a	5,49(6,7%)	-	-
b	0,59(3,5%)	0,07(8,7%)	-
c	0,20(1,1%)	0,08(7,3%)	-
d	0,02(5,5%)	0,57(36,3%)	-
e	122(-)	0,17(29,4%)	-
a. ²	2,43(0,7%)	-	-
$\beta_{handicap,b}$	0,02(47,9%)	-	0,018
$\beta_{handicap,d}$	-0,19(34,8%)	-	0,002

1 : Coefficient de variation $\Rightarrow ET\% = (\text{Écart-type}/\text{Estimation}) * 100$.

2 : Constante d'erreur.

L'analyse univariable avec l'origine du handicap après stratégie de sélection de modèle donne les résultats Table 2 ci-dessus. Le groupe acquis est considéré comme modalité de référence. Le paramètre e est fixé et n'est pas estimé. Ici, l'effet fixe représentant l'influence de l'origine du handicap sur les paramètres b et d d'IMAP est significatif selon un test de Wald. On a donc : $\log(b_i) = \log(b_{pop,acquis}) + \beta_{congenital} + \eta_{i,2}$, respectivement pour $\log(d_i)$.

Il est difficile d'envisager d'interpréter ces paramètres ensembles car l'effet d'un déclin plus précoce des performances pour un groupe donné influence directement sur le pic de performance au sens de la fonction IMAP. Ici, bien que le groupe congénital décline en moyenne plus tôt de par le $\beta_{handicap,d}$ estimé négatif, il n'est pas possible d'affirmer que ces derniers ont un pic de performance en moyenne plus élevé même si $\beta_{handicap,b}$ est significatif et positif. Ainsi, il est plus correct de considérer que le groupe congénital présente en moyenne un paramètre b plus élevé et un paramètre d moins élevé que le groupe acquis.

Les résultats des analyses univariées pour les autres covariables ne seront pas présentés ici.

Les résultats de l'analyse multivariée sont présentés ci-dessous :

TABLE 3 – Résultats modèle mixte IMAF après sélection de modèle. Analyse multivariée. Natation.

Paramètres	Effets fixes (ET% ¹)	Variance effets aléatoires (ET% ¹)	<i>p</i>
<i>a</i>	5,38(7,2%)	-	-
<i>b</i>	0,62(3,3%)	0,06(9,1%)	-
<i>c</i>	0,20(1,4%)	0,07(7,6%)	-
<i>d</i>	0,02(4,9%)	0,62(35,3%)	-
<i>e</i>	122(-)	0,21(26,2%)	-
<i>a</i> . ²	2,40(0,7%)	-	-
$\beta_{handicap,b}$	-0,01(14,7%)	-	<0,001
$\beta_{sexe,b}$	-0,05(29,9%)	-	<0,001
$\beta_{age-initial,b}$	-0,02(9,3%)	-	<0,001
$\beta_{sexe,c}$	-0,06(27,3%)	-	<0,001
$\beta_{age-initial,c}$	-0,02(12,2%)	-	<0,001
$\beta_{age-initial,e}$	0,01(15,1%)	-	<0,001

1 : Coefficient de variation $\Rightarrow ET\% = (\acute{E}cart\text{-}type/Estimation) * 100$.

2 : Constante d'erreur.

Les femmes ainsi que le groupe acquis représentent les modalités de référence. La variable d'âge d'entrée dans la compétition a été rapportée arbitrairement à 15 ans dans un souci d'interprétation (âge première performance - 15). Ainsi, par exemple, les femmes ayant un handicap acquis d'âge d'entrée dans la compétition égal à 15 ans auront en moyenne un coefficient *b* d'IMAF égal à 0,62.

Après ajustement sur le sexe et l'âge d'entrée dans la compétition, l'origine du handicap impacte significativement et uniquement le paramètre *b* de la fonction IMAF ($p < 0,001$). Le coefficient $\beta_{handicap,b}$ étant négatif, les athlètes ayant un handicap congénital de sexe et d'âge d'entrée dans la compétition égaux auront **en moyenne** un pic de performance (en % de la meilleure performance) plus faible que les acquis. Ces différences seront discutées par la suite. De plus, les estimations des autres paramètres β ne seront pas interprétées du fait que, comme énoncé précédemment, il est difficile d'interpréter un coefficient significatif indépendamment des autres associés à la même variable dans la modélisation (sexe sur *b* et *c*, âge d'entrée sur *b*, *c* et *e*).

L'adéquation du modèle final vérifie les hypothèses du modèle. Les distributions des effets aléatoires considérés sur les paramètres *b*, *c*, *d* et *e* suivent bien une loi normale (graphiques quantile-quantile, Figure 16). Il en est de même pour la distribution des erreurs du modèle (Figure 19). Le graphique des observations versus prédictions ne montre pas de tendance particulière (Figure 17). Le "visual predictive check" montre que le modèle prédit correctement l'évolution médiane de la performance (Figure 18). Des exemples de fit individuels sont fournis Figure 20.

Athlétisme

Sur les épreuves de **course**, 26% des trajectoires de l'échantillon de départ est pris en compte dans l'analyse, ce qui représente 1 173 trajectoires de performance Figure 14. Les femmes, le groupe acquis, la position debout ainsi qu'un âge d'entrée dans la compétition égal à 15 ans représentent les modalités de référence.

Le même schéma de modélisation que la natation est considéré comme le plus adapté pour la modélisation de ces trajectoires, à savoir : structure des effets aléatoires corrélés sur les paramètres *b*, *c*, *d* et *e* avec une transformation log-normale et une structure des erreurs constante ($BIC=61\ 740,5$ modèle par défaut $\Rightarrow 51\ 164,9$ modèle "vide" final).

L'analyse univariée avec l'origine du handicap ne permet pas de conclure à des différences de performance entre le groupe congénital et le groupe acquis malgré le fait que les coefficients des effets fixes de l'origine du handicap sur les paramètres *b* et *c* du modèle soient significativement supérieurs à

0. En effet, comme énoncé en méthode, une augmentation du paramètre b permet la croissance de la valeur du pic de performance et une augmentation du paramètre c participe à sa diminution. Il en est de même pour l'analyse multivariable après ajustement sur le sexe, l'âge d'entrée dans la compétition ainsi que la position d'exercice de la discipline (Table 13). Toutefois, les résultats montrent que l'augmentation de l'âge d'entrée dans la compétition diminue la performance moyenne avec un coefficient b négatif et c positif.

Il est donc préférable de s'intéresser aux trajectoires individuelles prédites par le modèle, ce qui a été fait par la suite. Il n'y a pas de violation des hypothèses d'adéquation du modèle.

Les épreuves de **lancer** sont modélisées par 517 trajectoires soit 25% de l'échantillon de départ. Les résultats sont à retrouver Table 14, ($BIC=28\,472,20$ modèle par défaut $\Rightarrow 27\,573,50$ modèle "vide" final). Les effets fixes associés à la variable de l'origine du handicap sont significativement différents de 0 sur les paramètres b , c , d et e .

Aucun résultat n'apparaît concernant les épreuves de **saut**, ce qui fait écho aux premiers résultats sur l'analyse des performeurs (Table 15). L'échantillon est constitué de 119 trajectoires soit 18% de l'échantillon de départ avec un seuil du nombre de mesures annuelles abaissé à 5 ou plus du fait de la faible quantité de données.

Analyses sur les trajectoires individuelles issues du modèle mixte

Les estimations des paramètres d'IMAP permettent de reconstruire les trajectoires individuelles et d'estimer un pic de performance théorique en fonction de l'origine du handicap, ce qui n'était pas possible à partir de la stratégie de modélisation utilisant les maximums de performance à chaque âge. La Figure 5 représente un exemple de quatre trajectoires d'athlètes en natation :

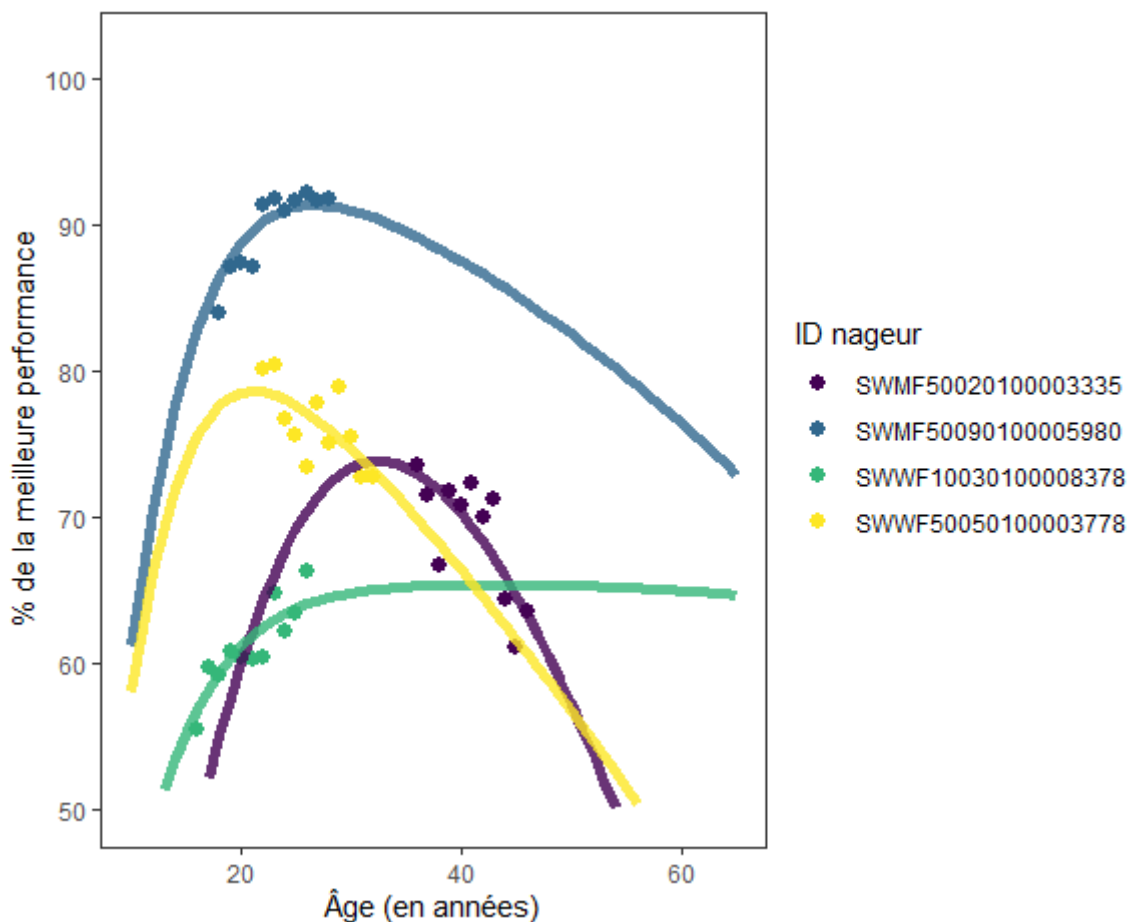


FIGURE 5 – Trajectoires selon IMAP à partir des estimations des paramètres individuels du modèle mixte. Natation.

Les maximums de performance théoriques individuels en natation sont significativement plus élevés chez les athlètes ayant acquis leur handicap au cours de la vie ($p < 0,01$, différence 1,23% chez les hommes et 1,25% chez les femmes). Ces résultats se retrouvent pour les sévérités "modérée", "minime" et "visuelle" chez les hommes et pour les sévérités "sévère" et "importante" chez les femmes (Figure 21, Figure 22). De plus ces athlètes sont significativement plus âgés pour un âge médian de 28 ans contre 26 pour le groupe congénital.

En athlétisme, sur l'ensemble des trajectoires des épreuves de course, il n'y a pas de différence de performance chez les hommes selon le test de Wilcoxon ($p = 0,785$), cependant la médiane des athlètes du groupe congénital est légèrement plus élevée que le groupe acquis (93,31% contre 92,94%). De plus, les femmes ayant acquis leur handicap performant significativement mieux avec une médiane de maximum de performance théorique de 92,91% contre 91,47% ($p < 0,01$). La discrétisation en fonction de la sévérité du handicap n'est pas étudiée une à une mais en fonction de la position d'exercice de la discipline. Les résultats suivent ceux énoncés ci-dessus avec, chez les hommes, aucune différence significative observée sur les positions debout (0,6%) et fauteuil (0,2%) mais une distribution de la performance en faveur du groupe congénital et chez les femmes une distribution significative en faveur du groupe acquis que la position soit fauteuil (2%) ou debout (3%). Enfin, concernant les épreuves de lancer masculines, les hommes ayant un handicap inné ont une médiane de performance supérieure de 2% comparée au groupe acquis. Dans tous les cas, l'âge au pic estimé est antérieur pour le groupe congénital (différences entre 6 et 8 ans).

L'ensemble des résultats issus de la modélisation par modèles mixtes appuient les résultats des analyses préliminaires sur les performeurs ainsi que les trajectoires par maximums de performance.

6 Discussion

L'objectif de cette étude était d'analyser les performances des athlètes ayant participé aux grandes compétitions internationales (JP et CM) en fonction de l'origine de leur handicap, inné ou acquis. Les performances ont été comparées à partir de deux approches : une appelée "brute" en tenant compte des meilleures performances des athlètes dans l'ensemble des événements auxquels ils ont concouru et l'autre dite "longitudinale" visant à qualifier l'évolution de la performance en fonction de l'âge et l'origine du handicap. Ce rapport tente de mettre l'accent sur l'évolution graduelle des analyses qui sont d'une part globales (performeurs, uniquement sur les grandes compétitions) puis tendent vers de l'individualisation en prenant en compte l'ensemble de la carrière des athlètes (trajectoires maximales, modèle mixte). Les analyses ont révélé que les athlètes ayant acquis leur handicap ont de meilleures performances en natation, mais ceci ne s'est pas vérifié sur les différentes épreuves d'athlétisme. Les athlètes ayant un handicap de naissance ont des carrières qui commencent plus tôt avec un potentiel déclin plus rapide des performances.

La spécificité des performances régies par un système de classification impose une normalisation des résultats des athlètes du fait que l'on s'intéresse à l'ensemble des catégories de handicap et non pas à une en particulier. Cela permet de comparer les performances dans des catégories de sévérité du handicap différentes.

Une particularité des performances des athlètes en situation de handicap est leur multi-participation aux différentes disciplines [22]. En effet, le nombre important d'événements pour une population moins nombreuse d'athlètes au niveau international comparée aux valides pourrait augmenter la possibilité d'atteindre les podiums [22]. De plus, les athlètes qui participent à plusieurs disciplines sont ceux qui ont des performances plus élevées [22]. Cela justifie le choix de l'analyse initiale sur les *performeurs* en fonction des trois groupes associés à l'origine du handicap : "congénital", "acquis" et "manquant". Sur cette analyse discrétisée selon la sévérité, les types d'épreuves ou encore le sexe, peu de différences significatives de performance sont observées. Cependant les distributions de performance sont légèrement plus élevées chez les athlètes ayant acquis leur handicap sur les épreuves de natation. La tendance semble être la même chez les femmes sur les épreuves d'athlétisme alors que chez les hommes le groupe congénital observe des distributions de performance légèrement plus élevées. Ceci peut être expliqué par le fait que seule la meilleure performance dans chaque événement sur les compétitions internationales (JP et CM) considérées ait été gardée. Une hypothèse peut être émise selon laquelle les athlètes étant éligibles à ce type de compétition ne présentent pas tous d'importantes disparités sur leurs meilleures performances sans tenir compte de l'évolution de leur carrière.

Actuellement, deux études à travers la méthode Delphi⁴ [13, 35] discutent de l'impact de l'origine du handicap sur les épreuves de natation et d'athlétisme mais uniquement chez les déficients visuels. Les experts, au nombre de 16, ont atteint un consensus (100%) en natation où les athlètes ayant acquis leur handicap bénéficieraient d'un avantage car la vision permettrait au nageur d'acquérir plus facilement des compétences motrices générales ainsi que certaines techniques de nage. Il est possible de se demander si l'acquisition de mécanismes moteurs pré-handicap pourrait être favorable à de meilleures performances après une période de réhabilitation et d'entraînement intensif pour les athlètes ayant acquis leur handicap (dans le cas des handicaps physiques). Toutefois, aucun consensus n'a cependant été trouvé en athlétisme et aucune autre étude ne mentionne l'origine du handicap sur cette discipline et permettrait de discuter les résultats observés.

L'âge est un déterminant incontournable de la performance, qu'elle soit sportive ou non, avec un développement des fonctions musculaires durant les premières étapes de la vie (enfance, adolescence) puis un déclin progressif observé après maturation [20]. Dans un premier temps, et pour prendre ainsi en compte les parcours sportifs des athlètes, une analyse a été réalisée à partir des trajectoires maximales de performance modélisées par les fonctions de Moore et d'IMAP.

Les résultats ont montré un âge au pic plus précoce pour le groupe congénital mais aussi des âges au pic plus précoces sur la natation comparée à l'athlétisme. Les différences d'âge au pic entre le groupe

4. La méthode Delphi consiste à réunir un panel d'experts afin de trouver un consensus sur une problématique précise.

congénital et le groupe acquis en natation peuvent être expliquées par le fait que les athlètes ayant un handicap de naissance entrent dans la compétition plus tôt du fait que leur apprentissage de la discipline se fait potentiellement depuis un jeune âge comme cela a pu être observé chez les déficients visuels [13]. En effet, la natation est une discipline accessible et fortement recommandée dans les processus de rééducation fonctionnelle et physique et d'appréhension du handicap. Le milieu aquatique permet d'assouplir les contraintes liées à l'apesanteur mais aussi au matériel et ainsi de diminuer l'impact des déficiences sur la capacité à se mouvoir tout en renforçant les capacités existantes. Cela rend possible la réalisation de mouvements qui seraient alors plus difficiles sur terre et permettrait à l'individu de se réaliser [31]. D'un autre côté, les athlètes ayant acquis leur handicap, deviendraient compétiteurs plus tardivement. En cause, un temps d'adaptation variable face aux contraintes liées au handicap et à son âge d'acquisition et, potentiellement, l'apprentissage à partir des bases de la discipline. Les résultats ont montré que cette différence est d'autant plus importante chez les handicaps sévères. Ceci peut être dû au fait que les athlètes qui concourent dans ces catégories évoluent principalement en fauteuil roulant dans leur vie quotidienne, bénéficient d'un usage limité de leurs mains, jambes ou de leur tronc, ont des problèmes de coordination modérés à sévères, et/ou un certain degré de perte au niveau de leurs membres (pouvant atteindre les quatre) [23]. Ces nageurs ont généralement besoin d'assistance pour le transport, l'accès aux installations ou encore le départ des courses. Il en résulte que ces athlètes disposent de beaucoup moins de temps d'entraînement, ce qui peut avoir un impact négatif sur la progression de leurs performances par rapport aux nageurs ayant une fonction physique plus élevée [23]. Pour les athlètes ayant acquis leur handicap, la familiarisation avec la mise en place d'une logistique pour répondre à ces problématiques peut repousser l'âge au pic. Ceci est bien évidemment dépendant, encore une fois, de l'âge d'acquisition du handicap.

La différence d'âge au pic de performance entre la natation et l'athlétisme pourrait être expliquée par les exigences propres à chaque discipline. En natation et comme cela vient d'être explicité ci-dessus, le milieu est par essence moins traumatique. Sur 22 sports étudiés au JP de Londres, la natation se situe au 16e rang des blessures alors que l'athlétisme occupe la 6e place [32]. En athlétisme, les phénomènes de compensation sur l'appareil musculo-squelettique dus à la perte d'un membre mais aussi les charges de travail intenses encaissées par les membres supérieurs des athlètes évoluant en fauteuil ainsi que les contraintes de l'appareillage pourraient être déterminants dans le risque de blessures [33]. Une adaptation potentiellement plus longue peut être associée à cette discipline, de par les sollicitations qu'elle demande sur le plan physique (risque de blessures) mais aussi sur le plan matériel (adaptation aux fauteuils, matériels orthopédiques [20]). De plus, l'étude des trajectoires maximales de performance a permis de révéler un déclin plus précoce et abrupte des performances pour le groupe inné. Une hypothèse avancée après réflexions internes à l'IRMES avec son directeur est que cela pourrait être dû aux pathologies mêmes qui peuvent être plus handicapantes avec l'avancée en âge (et évolutives) pour des athlètes qui ont commencé le niveau professionnel très jeune. En opposition à cela, les athlètes ayant acquis leur handicap sont plus touchés par des traumatismes (forte prévalence d'amputations), souvent non évolutifs et commencent leur carrière plus tard ce qui peut expliquer leur longévité (la majorité des athlètes au-delà des 50 ans ont un handicap acquis). Ainsi, des analyses en fonction des pathologies pourraient apporter des informations supplémentaires sur les parcours de vie dans le sport de haut niveau pour les sportifs en situation de handicap [21].

Les résultats en athlétisme sur l'âge au pic de performance vont dans le sens des études précédentes [20] avec un âge au pic plus tardif chez les athlètes performant en fauteuil qui sont sur-représentés dans la catégorie des handicaps acquis au cours de la vie.

Toutefois, ces analyses ne contrôlent pas le fait que, potentiellement, une chute du maximum de performance à un âge donné peut être dû au fait qu'un athlète performant ait arrêté sa carrière là où il contribuait à tous les maximums. De plus, une très bonne performance peut surestimer une trajectoire dans une catégorie donnée.

L'utilisation d'un modèle mixte amène un niveau de démonstration supplémentaire dans cette étude. L'observation de trajectoires de performance à partir des meilleures performances annuelles par évènement/athlète permettent de tenir compte dans l'analyse de la variabilité individuelle ce qui n'était pas le cas avec la méthode des maximums de performance. L'utilisation des fonctions de

Moore et d'IMAP pour représenter la relation âge/performance se justifie par deux points. D'une part, des études ont pu mettre en évidence que l'utilisation de fonctions quadratiques (dans le cas présent relation modélisée par $\text{âge} + \text{âge}^2$ par exemple) rapportait des estimations erronées de l'âge au pic de performance du fait que ce type de fonction ne tient pas compte de la relation asymétrique entre l'âge et la performance [30]. D'autre part, l'ajustement de modèles à partir de fonctions splines ou encore polynômes fractionnaires ne permet pas de produire des paramètres interprétables même si elles s'adaptent mieux aux données [34]. De plus, ces fonctions peuvent être influencées par des valeurs particulières ce qui peut être constaté de manière récurrente dans les données de performance des para-athlètes qui peuvent être très variables d'une année sur l'autre. La stratégie adoptée par l'utilisation de SAEMIX offre donc une nouvelle approche de la modélisation de la performance sportive par modèle mixte proposée par Hogart [24] (utilisation de splines de régression).

Dans cette étude, l'ajout de covariables comme l'origine du handicap a permis de mettre en valeur des performances plus élevées chez les para-nageurs ayant acquis leur handicap après ajustement sur les autres variables du modèle (coefficient b d'IMAP significatif) ainsi que des différences de performances sur les trajectoires individuelles légèrement supérieures à 1%. Ces résultats peuvent seulement être comparés à l'étude réalisée chez les déficients visuels [14] et apportent donc de nouvelles informations. Les âges au pic de performance entre les hommes et les femmes vont dans le sens de l'analyse sur les trajectoires maximales de performance ainsi que les résultats des études en natation [14, 24] et athlétisme [20]. En athlétisme, il apparaît difficile d'interpréter l'effet simultané de l'origine du handicap sur plusieurs paramètres de la fonction d'IMAP. En effet, l'étude des variations de la fonction d'IMAP (point où la dérivée s'annule) Annexe 12 permet de montrer que les paramètres b , c , d et e influencent l'âge au pic et le maximum de performance. La sélection d'un modèle robuste permet donc de pallier cela et de s'intéresser aux trajectoires individuelles prédites. Les différences observées sont cohérentes entre les analyses. Toutefois, il peut être noté que la construction du modèle est sensible à l'initialisation des paramètres d'IMAP et à la forme des données (trajectoires pouvant être en "dents de scie"). Sur toutes les trajectoires prédites, les âges au pic estimés ainsi que les maximums théoriques étaient plus élevés comparés à la méthode des maximums (âges atteignant les 40 ans et plus, maximums dépassant les 100%). En particulier, l'initialisation du paramètre e d'IMAP (i.e l'âge où la performance égale 0, soit le décès) qui a été fixé influe sur la courbe prédite (par les estimations individuelles). Il est évident que l'objectif n'est pas de prédire un âge au décès et que le choix est difficile entre déterminer un âge où tous les athlètes ont une performance égale à 0 ou laisser de la flexibilité avec des effets aléatoires. Le choix a été fait de garder des effets aléatoires mais cela a pour effet d'allonger les trajectoires dans les âges élevés sans diminuer drastiquement l'indicateur de performance. Une analyse de sensibilité pourrait être conduite en testant différentes initialisations des paramètres de la fonction d'IMAP.

Une autre dimension pourrait être explorée et relève de l'aspect mental des athlètes en situation de handicap. En effet, les athlètes ont besoin d'avoir confiance en leurs capacités lors des grandes compétitions internationales et notamment les JP qui sont les plus médiatisés. Un meilleur sentiment d'auto-efficacité⁵ et accompagnement psychologique permettraient l'amélioration des performances [36]. Une étude s'intéresse à l'analyse du bien-être psychologique d'une population d'athlètes participant aux épreuves de natation et d'athlétisme notamment [9]. Les résultats rapportent une meilleure estime de soi ainsi qu'une anxiété moindre chez les athlètes ayant acquis le handicap qui ont dû développer de meilleures capacités d'adaptation pour confronter et accepter leur déficience [9]. Toutefois, l'échantillon étudié est faible et le manque de littérature sur la question suggère que ces résultats sont à prendre sous réserve.

5. croyance qu'a un individu en sa capacité à réaliser une tâche

Limites

Le web-scraping permet de récupérer de l'information sur des sites officiels mais au-delà il est très difficile de compléter les données. La première limite à laquelle se heurte ce projet est donc la disponibilité des informations relatives à l'origine du handicap. En effet, il a pu être constaté que la proportion de données manquantes s'articule autour de 20-30% en fonction de la discipline sportive étudiée. De plus, les variables à disposition sont dans la totalité des variables liées à la performance (rang, temps/mètres/points) ou aux caractéristiques de l'évènement dans lequel l'athlète a concouru (distance, classification, type d'épreuve, lieux etc). L'origine du handicap est la variable d'intérêt principale dans l'explication des variations de l'indicateur de performance qui a été énoncé en 4.3. Il semble donc difficile d'imputer l'origine du handicap à partir de variables de performances. En effet, aucune information relative aux antécédents familiaux (ex : maladies héréditaires) ou caractéristiques biologiques de l'athlète ont été récoltées. Ceci peut s'expliquer bien évidemment par la sensibilité de ces informations mais aussi par l'ensemble d'un processus complexe et coûteux permettant de les récolter. Toutefois deux stratégies d'imputation ont été considérées en plus de l'imputation "à la main" énoncée en 4.2. Ces méthodes sont basées sur (i) la régression logistique et (ii) les forêts aléatoires à partir, respectivement, des packages `MICE` et `missForest` de R. L'hypothèse considérée sur le type de données manquantes par ces modèles est l'hypothèse MAR (données manquantes aléatoirement). Dans cette étude, il est correct de considérer que la probabilité d'observer une donnée manquante sur l'origine du handicap dépend uniquement de variables observées comme le rang (plus un athlète sera haut classé plus il aura de chance d'être médiatisé), mais aussi les types d'épreuves (ex : la nage du papillon est très technique et ainsi peu représentée à l'inverse des épreuves de 100m). C'est une hypothèse forte. Une analyse de sensibilité concernant ces méthodes a été réalisée : sur le jeu de données complet concernant l'origine du handicap, 10%, 20%, 30% puis 50% de données manquantes ont été introduites aléatoirement sur cette variable d'intérêt. Les deux méthodes ont ainsi été appliquées afin de comparer les prédictions avec les données connues à partir du pourcentage de mauvaise classification. Les résultats n'ont pas été concluants avec des proportions de mauvaises classifications de 45% avec la régression logistique et de 37% avec les forêts aléatoires ce qui s'explique par le faible pouvoir discriminant de variables comme l'âge, le sexe ou la sévérité du handicap. En parallèle de cela, l'imputation "à la main" à partir des descriptions est délicate et nécessite de bien comprendre les maladies lorsqu'elles sont mentionnées.

Concernant les analyses longitudinales, la difficulté a été dans un premier temps d'homogénéiser les noms des athlètes entre les bases Résultats et Ranking. Une méthode de détection de chaîne de caractères a été implémentée (distance de Jaro-Winkler) mais présente elle-même ses limites lorsque le nom d'un athlète dans une base diffère en grande partie dans l'autre base (noms de famille différents, surnoms etc). De plus, l'utilisation des maximums pour construire les trajectoires maximales de performance peut biaiser les analyses sachant que la performance maximale observée peut ne pas être représentative de l'échantillon d'un âge donné. Le modèle mixte permet donc d'appuyer les résultats de cette analyse. Cependant, la modélisation par les fonctions de Moore ou IMAP souffre d'un manque de données sur la phase décroissante de la courbe de performance en fonction de l'âge. Il y a une opposition entre les trajectoires d'athlètes "jeunes" qui sont majoritaires et observent surtout une phase croissante et les trajectoires d'athlètes plus âgés observant plutôt une phase descendante. Il est donc difficile d'observer une pleine trajectoire de performance car les données ne couvrent pas (encore) l'ensemble de la carrière des athlètes. Le suivi maximal est de 11 ans ce qui représente 6% des trajectoires en natation. La fonction IMAP notamment considérée pour le modèle mixte est peut-être ainsi trop complexe et les trajectoires individuelles seraient à prendre sous réserve.

Enfin, il est difficile de récupérer l'âge d'acquisition du handicap chez les athlètes l'ayant acquis au cours de la vie. Il apparaît évident que l'acquisition d'un handicap à l'âge de 5 ans n'aura pas le même impact sur la trajectoire de performance que l'acquisition d'un handicap à 15 ans. Une analyse différenciant les âges d'acquisition du handicap devrait être plus robuste. Si dans les échantillons à disposition la majorité des athlètes ont eu leur handicap tôt, l'âge au pic de performance risque d'être antérieur que si l'âge d'acquisition du handicap est majoritairement plus tardif ce qui n'est quasiment pas vérifiable à partir des descriptions fournies.

Perspectives

Une analyse complémentaire a été initiée et peut être ajoutée à la méthodologie. Il est possible de synthétiser l'information fournie par les 5 paramètres estimés du modèle mixte grâce à une réduction de dimension, obtenue par une analyse factorielle (ACP) couplée à une méthode de classification non-supervisée. L'intérêt de cela est de pouvoir identifier des clusters, groupements d'individus similaires ayant des cinétiques de performance, de progression propre. Ainsi l'identification de clusters et des caractéristiques des individus qui les composent pourrait apporter de l'information supplémentaire afin de déterminer des trajectoires types. Actuellement réalisée sur les données de natation, cette analyse ne fournit pas de discrétisation en fonction de l'origine du handicap mais de sa sévérité.

De plus, les JP de Tokyo se déroulant du 24 août au 5 septembre viendront enrichir les données existantes sur les athlètes en situation de handicap. Il serait donc intéressant de fournir de nouvelles analyses et de pouvoir comparer les différentes "paralympiades". Actuellement, une étude s'intéresse à la comparaison des performances entre les classifications (de 1992 à 2012) [23]. Une attention pourrait être portée sur une nouvelle comparaison de ces performances notamment sur les paralympiades de Rio et Tokyo afin de déterminer si le système de classification maintient son objectif d'équité entre les athlètes.

Il apparaît que l'origine du handicap pourrait être un déterminant de la performance sportive. La poursuite des investigations des différences entre les athlètes ayant un handicap de naissance et ceux l'ayant acquis semble nécessaire et pourrait apporter un niveau de précision supplémentaire dans le choix de la classification des athlètes. Ceci pourrait être un outil à la détection de jeunes talents. Toutefois, le manque d'informations actuel et notamment sur les âges d'acquisition du handicap constitue un frein à une meilleure individualisation des trajectoires de performance.

7 Conclusion

7.1 Conclusion générale

Actuellement, les athlètes en situation de handicap performant dans des classes qui ne tiennent pas compte de l'origine et de l'âge d'acquisition du handicap. L'étude qui a été réalisée montre des différences de performance au niveau international entre les athlètes qui ont acquis leur handicap et les athlètes qui ont un handicap de naissance. Les analyses ont été discrétisées en fonction de la sévérité du handicap et montrent des résultats propres à la discipline sportive étudiée. Ces résultats sont les premiers concernant l'analyse des performances des para-athlètes sur plusieurs disciplines sportives en fonction de l'origine et du type de handicap (physique ou visuel). Il apparaît qu'en natation, les athlètes ayant acquis leur handicap et plus particulièrement les hommes pourraient avoir de meilleures performances à l'inverse des analyses en athlétisme. Les différences sont faibles et moins probantes chez les femmes. De plus, les âges au pic de performance seraient plus précoces pour le groupe congénital. L'origine du handicap a donc son importance dans les performances mais il apparaît que la sévérité du handicap soit plus déterminante dans l'observation des différences de performance entre les athlètes. Bien que cette étude vienne compléter les premiers résultats de l'étude de l'origine du handicap sur les performances chez les déficients visuels [14], il semble nécessaire de bénéficier d'un recul plus conséquent afin d'infirmer ou de confirmer les différences observées. Toutefois, cette première analyse globale pourrait initier des réflexions concernant la détection de jeunes talents et la mise en place d'objectifs de performance adaptés pour les potentiels performeurs aux futures paralympiques en fonction de leur âge et trajectoires de performances actuelles.

Enfin, le projet Paraperf est constitué de trois lots de travail. Cette étude est en lien avec le lot numéro 1. Il serait intéressant de croiser ces résultats avec ceux du lot de travail numéro 3 qui a pour objectif "d'identifier les freins et leviers socio-environnementaux à la haute performance". Cependant, à l'heure actuelle aucune étude concernant l'origine du handicap autre que celle-ci n'a de premiers résultats.

Ce rapport s'achèvera sur cette citation de Sir Phillip Craven à la cérémonie de clôture des JP 2016 : *"Paralympiens, vos performances exceptionnelles ont focalisé le monde sur vos capacités sensationnelles. Les gens ont été émerveillés par ce que vous pouviez faire et ont oublié ce qu'ils croyaient que vous ne pouviez pas faire. Vous avez montré au monde qu'avec une attitude positive, le corps humain, et surtout le cœur et l'esprit humain, ne connaît pas de limites et qu'absolument tout est possible. Vous avez défié les attentes, réécrit le livre des records et transformé une pitié mal fondée en fierté. Vous êtes désormais des héros et des modèles pour une nouvelle génération de fans de sport du monde entier.*

7.2 Conclusion personnelle

Ces six mois au sein du laboratoire IRMES à l'INSEP m'ont permis de découvrir le monde de la recherche scientifique au sein d'une équipe particulièrement dynamique. J'ai eu la possibilité de découvrir un lieu mythique de l'excellence de la performance sportive en France, de visiter ses infrastructures et de rencontrer ses athlètes olympiques. J'ai aussi eu la chance de ne pas avoir eu de télétravail imposé, grâce à cela j'ai pu m'imprégner de l'environnement dans lequel j'étais, ce que je n'ai pas pu faire au cours de mon stage de M1. Ce stage fut donc ma première expérience professionnelle dans le cadre de mes études et directement en lien avec la thématique sur laquelle je souhaite poursuivre mes efforts : l'activité physique et sportive. J'ai découvert au sein du laboratoire la multitude de projets sur lesquels se concentrent les équipes, l'enjeu de leurs réflexions sur la compréhension de la performance sportive à maintenant 3 ans des Jeux de Paris 2024 mais surtout l'entraide constante qui permet de faire avancer ce tout. Les réunions de recherche auxquelles j'ai pu assister permettent de suivre les projets de chacun et représentent une véritable émulsion d'idées pour diriger au mieux la recherche. Enfin, j'ai développé des compétences que je n'avais pas (web-scraping) et apporté mes connaissances sur un sujet jusqu'ici peu étudié. Ce travail ainsi que la méthodologie développée pourront servir de base de réflexion pour le futur.

Bibliographie

- [1] Ministère des Sports, Institut National du Sport de l'Expertise et de la Performance. INSEP Rapport d'activités 2018. Disponible à l'url :
<https://www.insep.fr/sites/default/files/2020-01/ra-2018-web.pdf>
- [2] Laboratoires de recherche IRMES. INSEP [cité 23 mars 2021]. Disponible à l'url :
<https://labos-recherche.insep.fr/fr/linstitut-de-recherche-bio-medicale-et-depidemiologie-du-sport>
- [3] Lancement du projet de recherche Paraperf | Handisport.org [Internet]. [cité 17 mai 2021]. Disponible à l'url :
<https://www.handisport.org/lancement-du-projet-de-recherche-paraperf/>
- [4] O'Young B, Gosney J, Ahn C. The Concept and Epidemiology of Disability. *Physical Medicine and Rehabilitation Clinics of North America*. nov 2019;30(4) :697-707.
- [5] Bogart KR. The role of disability self-concept in adaptation to congenital or acquired disability. *Rehabilitation Psychology*. fév 2014;59(1) :107-15.
- [6] Legg, D. Paralympic Games : History and Legacy of a Global Movement. *Physical Medicine and Rehabilitation Clinics of North America*. 2018;29(2),417-425.
- [7] Channel 4. Research : Paralympics coverage helped change attitudes to disability [Internet]. [cité 8 juin 2021]. Disponible à l'url :
<https://www.channel4.com/press/news/research-paralympics-coverage-helped-change-attitudes-disability>
- [8] Marin-Urquiza A, Ferreira JP, Van Biesen D. Athletic identity and self-esteem among active and retired Paralympic athletes. *European Journal of Sport Science*. 3 juill 2018;18(6) :861-71.
- [9] Campbell E. Psychological well-being of participants in wheelchair sports : comparison of individuals with congenital and acquired disabilities. *Perceptual and Motor Skills*. oct 1995;81(2) :563-8.
- [10] Tweedy SM, Vanlandewijck YC. International Paralympic Committee position stand-background and scientific principles of classification in Paralympic sport. *Br J Sports Med*. avr 2011;45(4) :259-69.
- [11] International Paralympic Committee (IPC). Explanatory Guide to Paralympic Classification, paralympic summer sports. Sep 2015.
- [12] Powis B, Macbeth J. "We know who is a cheat and who is not. But what can you do?" : Athletes' perspectives on classification in visually impaired sport. *International Review for the Sociology of Sport*. 3 Fév 2019.
- [13] Ravensbergen HJC, Genee AD, Mann DL. Expert Consensus to Guide the Classification of Paralympic Swimmers With Vision Impairment : A Delphi Study. *Front Psychol*. 17 oct 2018;9 :1756.
- [14] Le Toquin B, Schipman J, De Laroche Lambert Q, Sauliere G, Toussaint JF. Is the visual impairment origin a performance factor ? Analysis of international level Para swimmers and Para athletes [Under Review]. 2021.
- [15] Skordilis EK, Skafida FA, Chrysagis N, Nikitaras N. Comparison of sport achievement orientation of male wheelchair basketball athletes with congenital and acquired disabilities. *Percept Mot Skills*. Déc 2006;103(3) :726-32.

- [16] Page SJ, O'Connor EA, Wayda VK. Exploring Competitive Orientation in a Group of Athletes Participating in the 1996 Paralympic Trials. *Percept Mot Skills*. oct 2000 ;91(2) :491-502.
- [17] Guillaume Collet. Rapport de Conférence - Les Mouvements Généraux - Évaluation Neurologique du Nourrisson de moins d'un an. 30 Mar 2020.
- [18] Prechtl HFR. General movement assessment as a method of developmental neurology : new paradigms and their consequences. *Developmental Medicine & Child Neurology*. 2001 ;43(12) :836-42.
- [19] Fulton SK, Pyne DB, Hopkins WG, Burkett B. Variability and progression in competitive performance of Paralympic swimmers. *Journal of Sports Sciences*. 1 Mar 2009 ;27(5) :535-9.
- [20] Schipman J, Gallo P, Marc A, Antero J, Toussaint J-F, Sedeaud A, et al. Age-Related Changes in Para and Wheelchair Racing Athlete's Performances. *Front Physiol*. 19 mars 2019 ;10 :256.
- [21] Patatas JM, Bosscher VD. Understanding Paralympic Athletes' Pathways : a preliminary investigation. In : *VISTA Conference 2017*. 2017. p. 102-3.
- [22] Schipman J, Saulière G, Le Toquin B, Marc A, Forstmann N, Toussaint J-F, et al. Involvement in Multiple Race Events Among International Para and Non-disabled Swimmers. *Front Sports Act Living*. 2020 ;2 :608777.
- [23] Burkett B, Payton C, Van de Vliet P, Jarvis H, Daly D, Mehrkuehler C, et al. Performance Characteristics of Para Swimmers : How Effective Is the Swimming Classification System? *Phys Med Rehabil Clin N Am*. mai 2018 ;29(2) :333-46.
- [24] Hogarth L, Nicholson V, Payton C, Burkett B. Modelling the age-related trajectory of performance in Para swimmers with physical, vision and intellectual impairment. *Scand J Med Sci Sports*. Avr 2021 ;31(4) :925-35.
- [25] Milevsky MA. Calibrating Gompertz in reverse : What is your longevity-risk-adjusted global age? *Insurance : Mathematics and Economics*. mai 2020 ;92 :147-61.
- [26] Laird AK. Dynamics of Tumor Growth. *Br J Cancer*. sept 1964 ;18(3) :490-502.
- [27] Lee L, Atkinson D, Hirst AG, Cornell SJ. A new framework for growth curve fitting based on the von Bertalanffy Growth Function. *Sci Rep*. 14 mai 2020 ;10(1) :7953.
- [28] Moore DH. A study of age group track and field records to relate age and running speed. *Nature*. janv 1975 ;253(5489) :264-5.
- [29] Berthelot G, Len S, Hellard P, Tafflet M, Guillaume M, Vollmer J-C, et al. Exponential growth combined with exponential decline explains lifetime performance evolution in individual and human species. *AGE*. Août 2012 ;34(4) :1001-9.
- [30] Berthelot G, Bar-Hen A, Marck A, Foulonneau V, Douady S, Noirez P, et al. An integrative modeling approach to the age-performance relationship in mammals at the cellular scale. *Sci Rep*. Déc 2019 ;9(1) :418.
- [31] Handisport au fil de l'eau - Le dossier - Handisport Mag. N°162 [Internet]. [cité 20 août 2021]. Disponible à l'url :
http://www.handisport-lemag.org/_archive/162/catalogues/catalogue_accessible/le-dossier/handisport-au-fil-de-leau.html
- [32] Willick SE, Webborn N, Emery C, Blauwet CA, Pit-Grosheide P, Stomphorst J, et al. The epidemiology of injuries at the London 2012 Paralympic Games. *Br J Sports Med*. mai 2013 ;47(7) :426-32.
- [33] Lucie Caudal. Épidémiologie des blessures en athlétisme handisport : étude de cohorte sur 4 compétitions nationales. *Médecine humaine et pathologie*. 2018. ffdumas-02072520f.

- [34] Botton J, Scherdel P, Regnault N, Heude B, Charles M-A, EDEN Mother-Child Cohort Study Group. Postnatal Weight and Height Growth Modeling and Prediction of Body Mass Index as a Function of Time for the Study of Growth Determinants. *Ann Nutr Metab.* 2014 ;65(2-3) :156-66.
- [35] Allen PM, Dolan R, Croxall H, Ravensbergen RHJC, Brooks A, Zenk F, et al. Evidence-based Classification in Track Athletics for Athletes with a Vision Impairment : A Delphi Study. *Optom Vis Sci.* nov 2020 ;97(11) :984-94.
- [36] Martin J. Mental Preparation for the 2014 Winter Paralympic Games. *Clinical Journal of Sport Medicine.* janv 2012 ;22(1) :70-3.

Annexes

Annexe 1

TABLE 4 – Caractéristiques des athlètes en athlétisme en fonction de l'origine du handicap. Base Résultats (n=2771).

Variables	Origine du handicap										p^3
	Congénital (n=711)			Acquis (n=1161)			Manquant (n=899)				
	n	% ¹	Moy(ET) ²	n	% ¹	Moy(ET) ²	n	% ¹	Moy(ET) ²		
Sexe											***
	hommes	434	61,0		793	68,3		627	69,7		
	femmes	277	39,0		368	31,7		272	30,3		
Sévérité											***
	CP ⁴	196	27,6		95	8,2		218	24,2		
	CP fauteuil	73	10,3		37	3,2		96	10,7		
	membres inf.	67	9,4		222	19,1		61	6,8		
	membres sup.	102	14,3		118	10,2		98	10,9		
	paraplégie	81	11,4		323	27,8		123	13,7		
	tétraplégie	10	1,4		90	7,8		12	1,3		
	petite taille	93	13,1		0	0,0		0	0,0		
	visuelle	89	12,5		276	23,8		291	32,4		
Position											***
	debout	547	76,9		711	61,2		668	74,3		
	fauteuil	164	23,1		450	38,8		231	25,7		
Discipline⁵											ns
	course	443	52,4		661	50,1		566	47,4		
	lancer	260	30,8		465	35,3		435	36,4		
	saut	142	16,8		193	14,6		194	16,2		
Compétition											***
	JP	70	9,9		184	15,8		191	21,2		
	CM	258	36,2		413	35,6		453	50,4		
	les deux	383	53,9		564	48,6		255	28,4		
Âge⁶ (années)		711		29,1(7,0)	1161		35,7(9,4)	897		32,0(8,2)	***
Nb.perf/ath		711		4,8(3,9)	1161		4,3(3,8)	899		2,9(2,6)	***

1 :Les pourcentages sont présentés en colonne.

2 :Moyenne(Ecart-Type).

3 :seuil $\alpha=5\%$. *** = <0.001 , ns=non significatif.

4 : CP= lésés cérébraux.

5 :Un même athlète peut participer à plusieurs disciplines ce qui explique que les effectifs sont supérieurs aux effectifs totaux des athlètes en fonction de l'origine du handicap.

6 : L'âge ici correspond à l'âge actuel pour permettre une première comparaison sur l'échantillon avec une ligne par athlète.

```

# Fonctions
IMAP1=function(t,p){
  p[1]*exp((p[2]/p[3])*(1-exp(-p[3]*t)))*(1-exp(p[4]*(t-p[5])))}
moore2= function(t,p){
  p[1]*(1-exp(-p[2]*t))+p[3]*(1-exp(p[4]*t))+p[5]}

MMC2=function(age,performance,methode,nbpara,precision=1000,borne=0,
initial=runif(nbpara)){

  SQ_methode=function(p){ # Permet de calculer la somme carre des residus
    sum((performance-methode(age,p))^2) }

  o=rep(10^6,precision) # o permet de stocker la fonction objectif
  r=rep(0,precision) # r permet de sotcker le R2
  para=rep(10^6,precision*(nbpara+1)) # stockage de tous les parametres

  for (i in seq(nbpara+1,precision*(nbpara+1),nbpara+1)){
    # optimisation de SQ_methode, initilisation avec runif
    sol=nlminb(start=initial + rnorm(nbpara,0,sqrt(abs(initial))),SQ_methode)
    # nouvelle optimisation, borne=122 > IMAP e=age décès 122 max...
    res= nlminb(start = sol$par,SQ_methode,lower=c(0,0,0,0,80),
    upper = c(Inf,Inf,Inf,Inf,122))
    para[i:(i+nbpara-1)]=res$par # stockage des paramètres calculés
    o[i/(nbpara+1)]=res$objective} # résultat

  compt=which.min(o) # minimum de la fonction a minimiser
  parametre=rep(0,nbpara)
  # stockage des meilleurs parametres
  parametre=para[((nbpara+1)*compt):(((nbpara+1)*compt)+(nbpara-1))]

  fit=methode(age,parametre) # regression
  R2=sum((fit-mean(vitesse))^2)/sum((vitesse-mean(vitesse))^2) #R2
  # R2a=R2 - (nbpara*(1-R2))/(length(age)-nbpara-1) R2 ajusté

  plot(age,vitesse,xlab="Age",ylab="% de la meilleure performance")
  lines(sort(age),methode(sort(age),parametre),col=round(runif(1,1,3),0))
  peak=max((methode(sort(age),parametre)))
  age_peak=sort(age)[methode(sort(age),parametre)==max((methode(sort(age),parametre)))]
  x=seq(min(age),max(age),0.01)
  ap=x[methode(x,parametre)==max(methode(x,parametre))]

  # paramètres, R2 et âge au pic de performance estimés
  return(list(parametre=parametre,objectif=min(o),R2=R2,peak=peak,age_peak2=ap))}

# Exemple d'utilisation, initialise p[5] à max âge
MMC_IMAP <- MMC2(data$Age, data$Perf, IMAP1,5,precision = 300,
initial = c(runif(4),122))
# pour moore, changer dans nlminb bornes : param a,b,c,d positifs [0,Inf], e[-Inf, Inf]

```

FIGURE 6 – Méthode des moindres carrés pour l'estimation des paramètres des fonctions Moore et IMAP.

Annexe 3

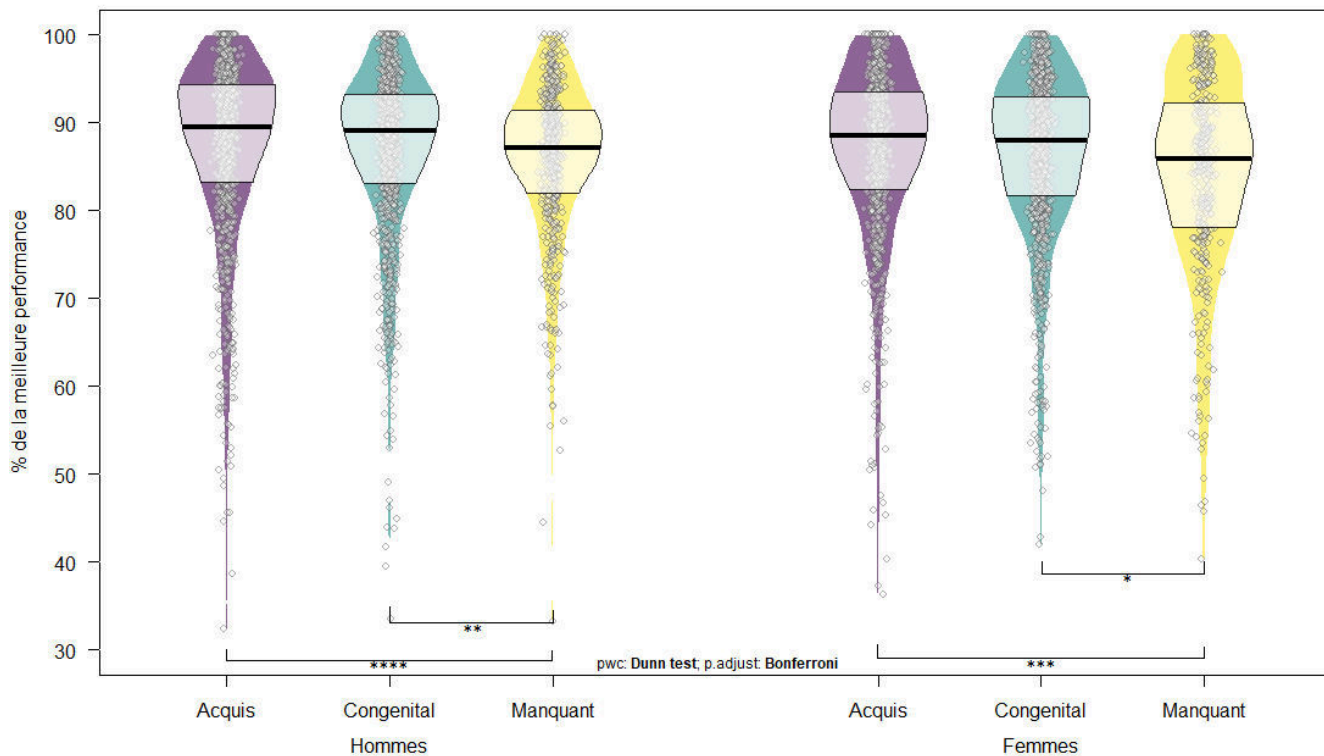


FIGURE 7 – Distribution du % de la meilleure performance en fonction de l’origine du handicap. Natation (n=5034).

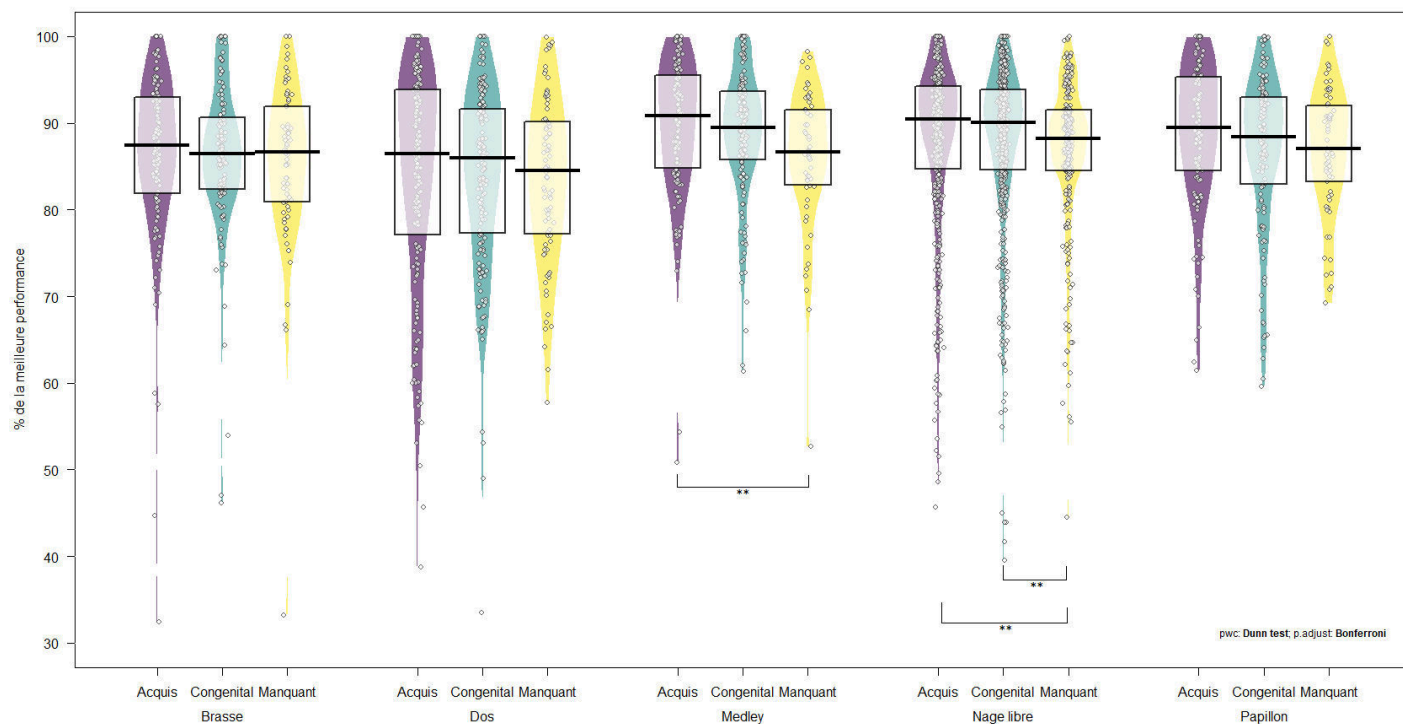


FIGURE 8 – Distribution du % de la meilleure performance en fonction de l’origine du handicap et du type d’épreuve. Toutes sévérités, hommes. Natation (n=2702).

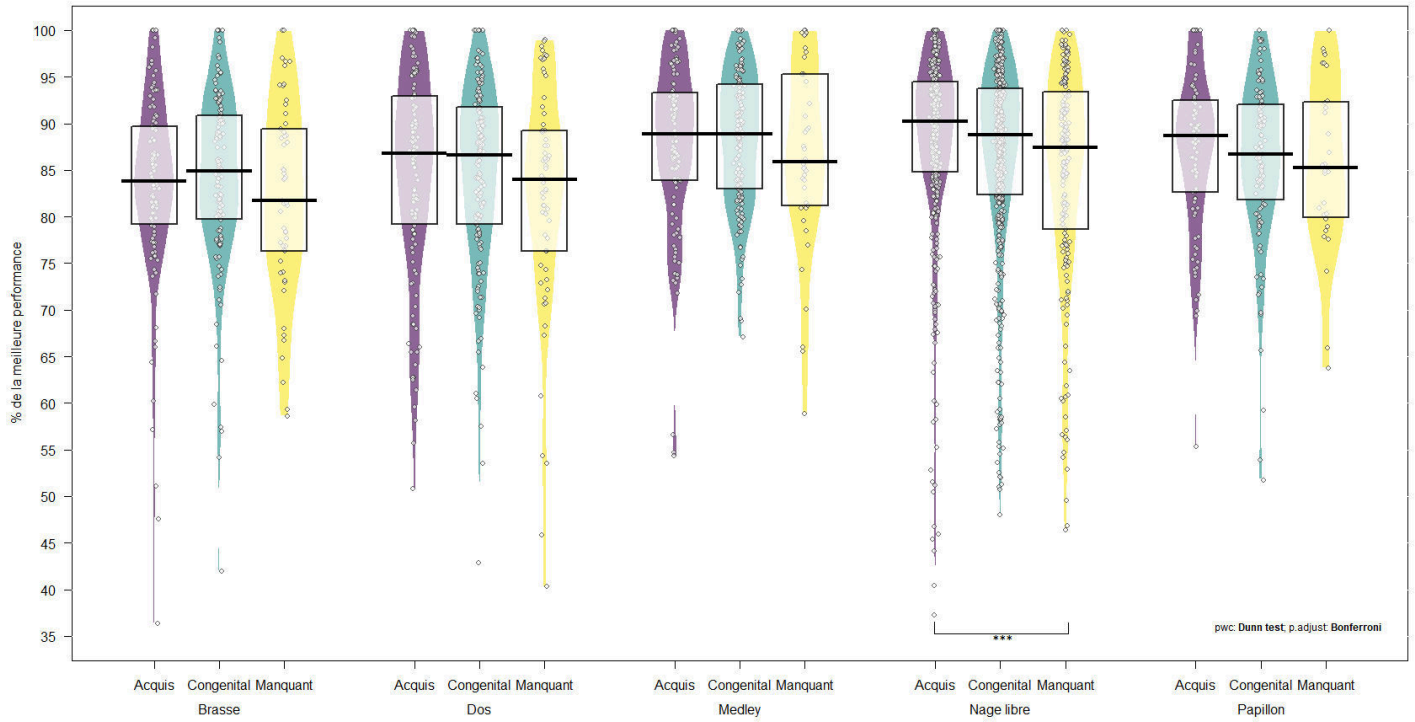


FIGURE 9 – Distribution du % de la meilleure performance en fonction de l'origine du handicap et du type d'épreuve. Toutes sévérités, femmes. Natation (n=2332).

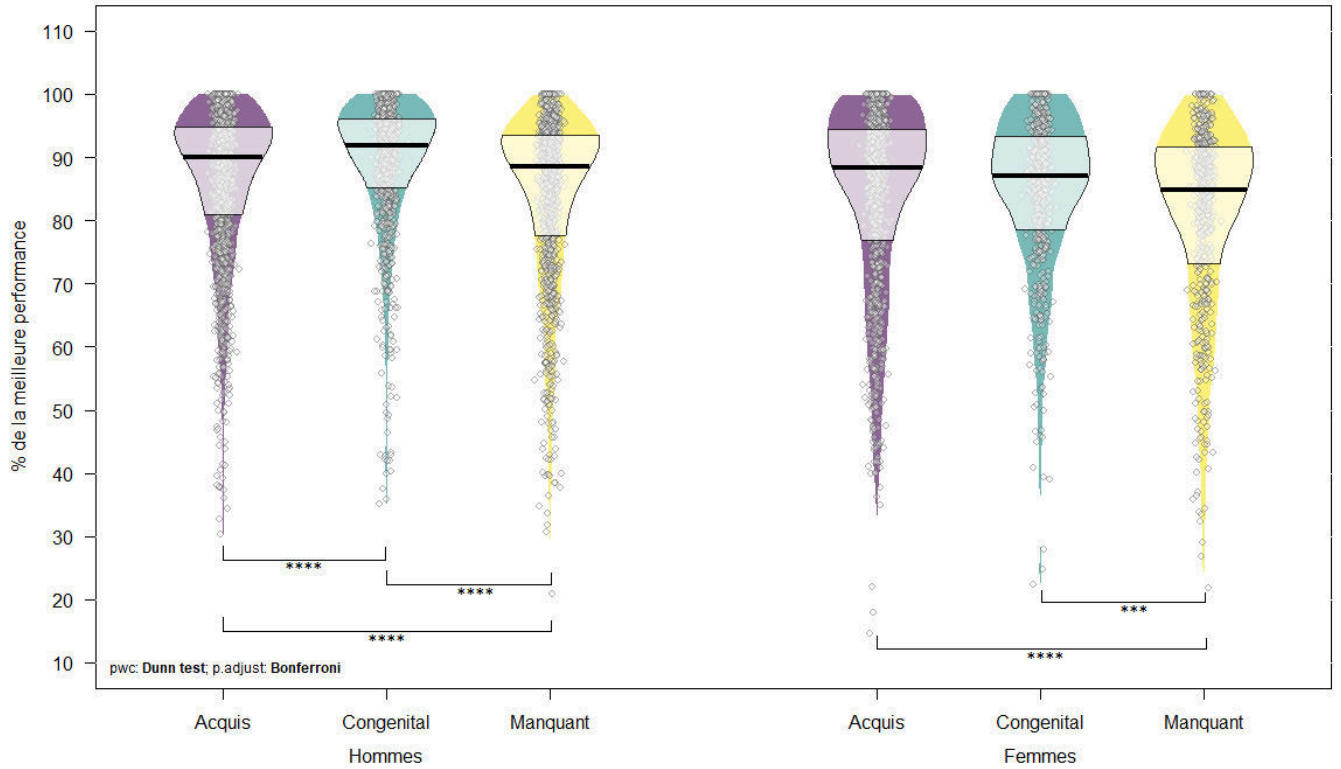


FIGURE 10 – Distribution du % de la meilleure performance en fonction de l'origine du handicap. Athlétisme (n=6097).

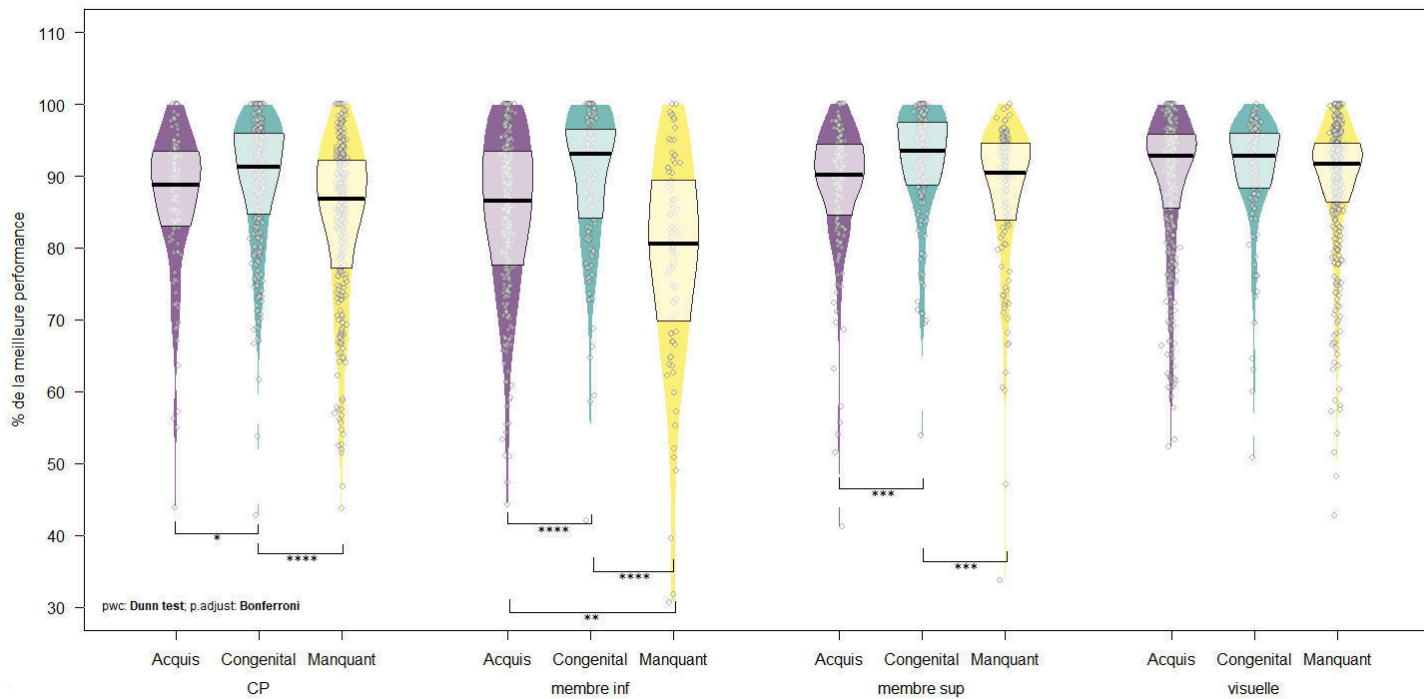


FIGURE 11 – Distribution du % de la meilleure performance en fonction de l'origine du handicap et de la sévérité. Athlétisme, épreuves debout, hommes (n=2614).

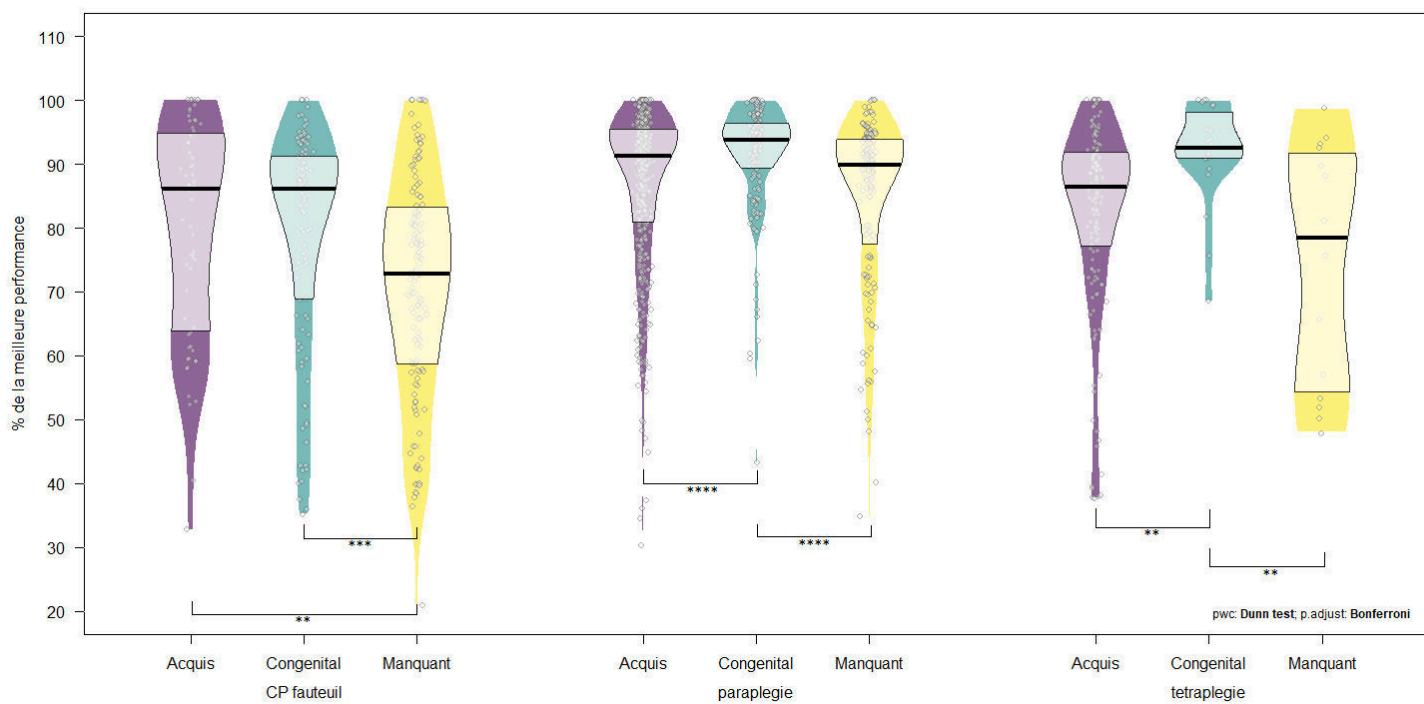


FIGURE 12 – Distribution du % de la meilleure performance en fonction de l'origine du handicap et de la sévérité. Athlétisme épreuves fauteuil, hommes (n=1304).

TABLE 5 – Caractéristiques des athlètes en natation. Données répétées.

		Athlètes	Trajectoires ¹	Performances	Âge	Nb. perf/traj.
		n	n	n	Q2[Q1-Q3] ²	Q2[Q1-Q3] ²
Hommes	Tous	538	4 954	17 809	23[19-28]	5[3-8]
	Congénital	270	2 654	9 914	21[18-25]	5[3-8]
	Acquis	268	2 300	7 895	25[21-31]	5[3-8]
Femmes	Toutes	418	4 130	14 921	20[17-25]	5[3-8]
	Congénital	231	2 306	8 964	20[17-23]	5[4-8]
	Acquis	187	1824	5 957	22[18-27]	5[3-7]
Total		956	9 084	32 730	22[18-26]	5[3-8]

1 : Une trajectoire correspond à l'ensemble des mesures répétées d'un athlète dans un évènement

2 : Q2 = médiane, Q1 = 1er quartile, Q3 = 3ème quartile

2 athlètes de la table Résultat ne figuraient pas dans la base Ranking, ce qui explique la différence entre les deux bases

TABLE 6 – Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de course.

		Athlètes	Trajectoires ¹	Performances	Âge	Nb. perf/traj.
		n	n	n	Q2[Q1-Q3] ²	Q2[Q1-Q3] ²
Hommes	Tous	717	2 935	12 219	28[23-34]	6[4-9]
	Congénital	272	1 088	4 911	24[21-29]	7[4-9]
	Acquis	445	1 847	7 308	30[25-37]	6[4-8]
Femmes	Toutes	387	1 507	6 166	25[21-31]	6[4-8]
	Congénital	171	601	2 581	23[19-27]	6[4-8]
	Acquis	216	906	3 585	27[22-34]	5[3-8]
Total		1 104	4 442	18 385	27[22-33]	6[4-8]

1 : Une trajectoire correspond à l'ensemble des mesures répétées d'un athlète dans un évènement

2 : Q2 = médiane, Q1 = 1er quartile, Q3 = 3ème quartile

TABLE 7 – Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de lancer.

		Athlètes	Trajectoires ¹	Performances	Âge	Nb. perf/traj.
		n	n	n	Q2[Q1-Q3] ²	Q2[Q1-Q3] ²
Hommes	Tous	447	1 275	5 135	31[26-38]	6[4-8]
	Congénital	143	429	1 877	27[23-32]	6[4-9]
	Acquis	304	846	3 258	34[28-41]	6[3-8]
Femmes	Toutes	278	811	3 298	31[24-40]	6[4-8]
	Congénital	117	320	1 307	25[21-29]	6[4-8]
	Acquis	161	491	1 991	36[30-43]	6[4-8]
Total		725	2 086	8 433	31[25-39]	6[4-8]

1 : Une trajectoire correspond à l'ensemble des mesures répétées d'un athlète dans un évènement

2 : Q2 = médiane, Q1 = 1er quartile, Q3 = 3ème quartile

TABLE 8 – Caractéristiques des athlètes en athlétisme. Données répétées sur les épreuves de saut.

		Athlètes	Trajectoires ¹	Performances	Âge	Nb. perf/traj.
		n	n	n	Q2[Q1-Q3] ²	Q2[Q1-Q3] ²
Hommes	Tous	214	481	1 356	26[21-30]	4[2-5]
	Congénital	88	184	552	24[20-28]	4[3-6]
	Acquis	126	297	804	27[22-32]	4[2-5]
Femmes	Toutes	121	198	607	25[21-30]	4[3-5]
	Congénital	54	86	265	22[19-26]	4[3-5]
	Acquis	67	112	342	27[23-32]	4[3-5]
Total		335	679	1 963	25[21-30]	4[2-5]

1 : Une trajectoire correspond à l'ensemble des mesures répétées d'un athlète dans un évènement

2 : Q2 = médiane, Q1 = 1er quartile, Q3 = 3ème quartile

Annexe 5

TABLE 9 – Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Natation.

Sévérité handicap	Congénital		Acquis	
	R ²	Âge au pic	R ²	Âge au pic
MOORE hommes				
Toutes	0,93	20,77	0,85	23,56
Sévère	0,91	22,39	0,64	31,00
Importante	0,89	26,35	0,62	20,44
Modérée	0,93	21,20	0,69	23,09
Minime	0,90	19,82	0,64	22,78
Visuelle	0,75	21,58	0,64	22,35
IMAP hommes				
Toutes	0,77	25,61	0,80	25,53
Sévère	0,77	26,63	0,64	30,71
Importante	0,71	31,20	0,52	24,13
Modérée	0,77	23,89	0,69	23,02
Minime	0,70	21,12	0,63	22,58
Visuelle	0,71	22,36	0,61	23,31
MOORE femmes				
Toutes	0,88	19,22	0,84	21,44
Sévère	0,88	21,18	0,34	30,80
Importante	0,88	17,87	0,91	24,94
Modérée	0,89	21,24	0,20	18,51
Minime	0,96	18,78	0,91	17,87
Visuelle	0,77	18,26	0,82	17,85
IMAP femmes				
Toutes	0,85	20,95	0,52	23,94
Sévère	0,81	24,04	0,34	30,49
Importante	0,77	20,27	0,65	23,24
Modérée	0,90	21,06	0,17	20,52
Minime	0,87	19,21	0,62	21,79
Visuelle	0,65	20,49	0,80	19,94

TABLE 10 – Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de course.

Sévérité handicap	Congénital		Acquis	
	R ²	Âge au pic	R ²	Âge au pic
MOORE hommes				
Toutes	0,89	25,87	0,88	32,90
CP	0,86	21,54	0,79	23,61
CP fauteuil	0,78	19,80	0,89	34,00
membres inf.	0,12	16,81	0,81	26,82
membres sup.	0,80	22,07	0,82	22,38
paraplégie	0,90	26,60	0,82	29,85
tétraplégie	0,55	20,94	0,62	38,88
visuelle	0,62	26,36	0,90	24,09
IMAP hommes				
Toutes	0,65	24,59	0,64	28,77
CP	0,85	22,94	0,75	24,53
CP fauteuil	0,73	21,24	0,89	34,00
membres inf.	0,09	14,09	0,81	26,82
membres sup.	0,64	23,15	0,76	24,44
paraplégie	0,73	27,28	0,67	30,66
tétraplégie	0,52	22,19	0,62	38,88
visuelle	0,53	24,78	0,83	26,11
MOORE femmes				
Toutes	0.69	18.45	0.88	26.08
CP	0.94	19.75	0.60	19.34
CP fauteuil	0.83	21.35	0.48	40.00
membres inf.	0.83	20.38	0.46	24.33
membres sup.	0.74	20.38	0.33	21.73
paraplégie	0.55	24.26	0.90	27.59
tétraplégie	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>
visuelle	0.54	22.99	0.39	26.48
IMAP femmes				
Toutes	0,55	22,90	0,85	25,84
CP	0,94	20,14	0,60	18,94
CP fauteuil	0,65	21,41	0,49	40,00
membres inf.	0,52	20,52	0,47	24,34
membres sup.	0,72	21,62	0,32	22,30
paraplégie	0,56	24,34	0,83	27,68
tétraplégie	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>
visuelle	0,46	25,92	0,39	26,43

TABLE 11 – Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de lancer.

Sévérité handicap	Congénital		Acquis	
	R ²	Âge au pic	R ²	Âge au pic
MOORE hommes				
Toutes	0,85	27,45	0,71	33,67
CP	0,81	27,56	0,55	27,58
CP fauteuil	0,66	28,18	0,35	44,14
membres inf.	0,77	25,98	0,84	34,69
membres sup.	0,84	31,61	0,40	25,43
paraplégie	0,67	31,61	0,80	32,64
tétraplégie	NA	NA	NA	NA
visuelle	0,76	22,00	0,68	30,63
IMAP hommes				
Toutes	0,67	27,70	0,72	33,32
CP	0,60	28,06	0,57	27,81
CP fauteuil	0,54	27,77	0,12	29,01
membres inf.	0,49	30,14	0,67	34,35
membres sup.	0,82	30,49	0,21	28,36
paraplégie	0,40	25,16	0,79	33,38
tétraplégie	NA	NA	NA	NA
visuelle	0,76	22,00	0,57	30,03
MOORE femmes				
Toutes	0,58	21,96	0,77	40,42
CP	0,76	21,09	0,19	33,85
CP fauteuil	0,35	29,84	0,36	43,16
membres inf.	0,34	26,02	0,66	27,04
membres sup.	0,70	22,29	0,82	23,79
paraplégie	0,19	25,99	0,54	45,01
tétraplégie	NA	NA	NA	NA
visuelle	0,84	27,50	0,83	26,47
IMAP femmes				
Toutes	0,54	24,95	0,69	38,61
CP	0,70	23,35	0,07	14,00
CP fauteuil	0,35	29,81	0,28	40,14
membres inf.	0,33	25,87	0,40	25,31
membres sup.	0,41	22,96	0,45	22,09
paraplégie	0,09	32,22	0,54	44,91
tétraplégie	NA	NA	NA	NA
visuelle	0,38	32,00	0,73	29,35

TABLE 12 – Comparaison des modélisations de Moore et IMAP en fonction de l'origine et de la sévérité du handicap. Maximums de performance à chaque âge (exprimés en % de la meilleure performance). Épreuves de saut.

Sévérité handicap	Congénital		Acquis	
	R ²	Âge au pic	R ²	Âge au pic
MOORE hommes				
Toutes	0,60	23,27	0,83	23,29
CP	0,76	27,59	0,28	23,21
membres inf.	0,32	23,60	0,82	23,37
membres sup.	0,89	25,40	0,68	19,00
visuelle	0,51	26,69	0,57	24,28
IMAP hommes				
Toutes	0,60	22,41	0,82	25,26
CP	0,36	23,22	0,23	25,31
membres inf.	0,16	14,00	0,81	24,33
membres sup.	0,75	24,22	0,66	19,00
visuelle	0,51	26,69	0,57	25,26
MOORE femmes				
Toutes	0,84	20,62	0,69	26,07
CP	0,60	24,47	0,20	40,01
membres inf.	0,57	19,86	0,91	26,05
membres sup.	0,59	21,35	0,61	24,99
visuelle	0,28	25,57	0,50	29,71
IMAP femmes				
Toutes	0,68	22,83	0,71	25,97
CP	0,31	21,33	0,12	37,26
membres inf.	0,06	19,87	0,86	26,90
membres sup.	0,48	22,99	0,06	25,73
visuelle	0,27	25,10	0,32	26,78

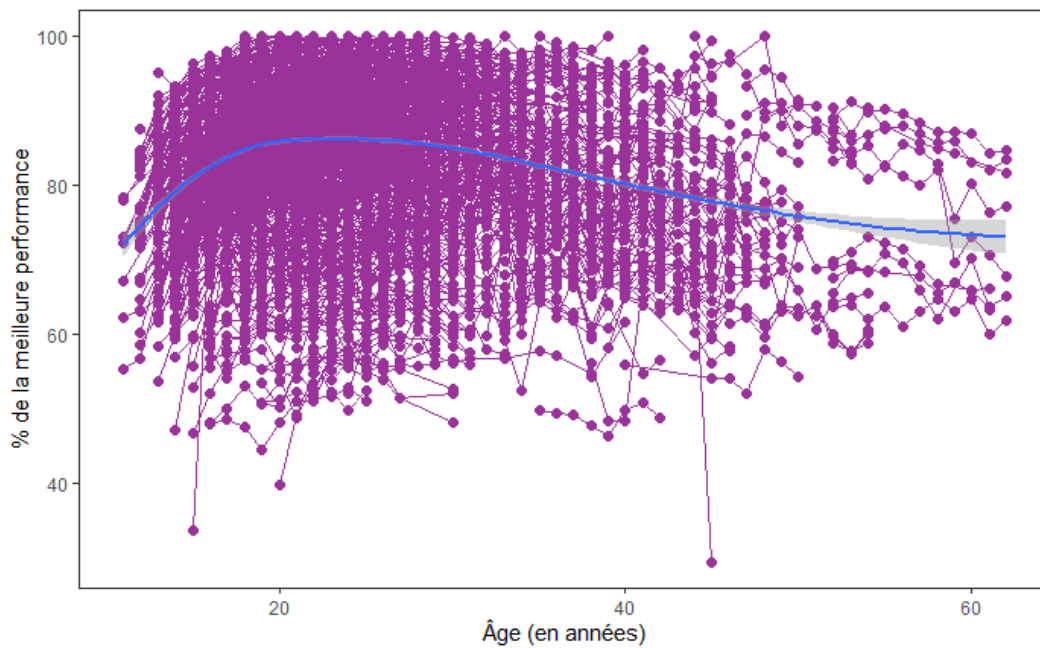


FIGURE 13 – Trajectoires de performance en natation. Échantillon de 7 ou plus performances annuelles. Courbe bleue : smoothing méthode "loess".

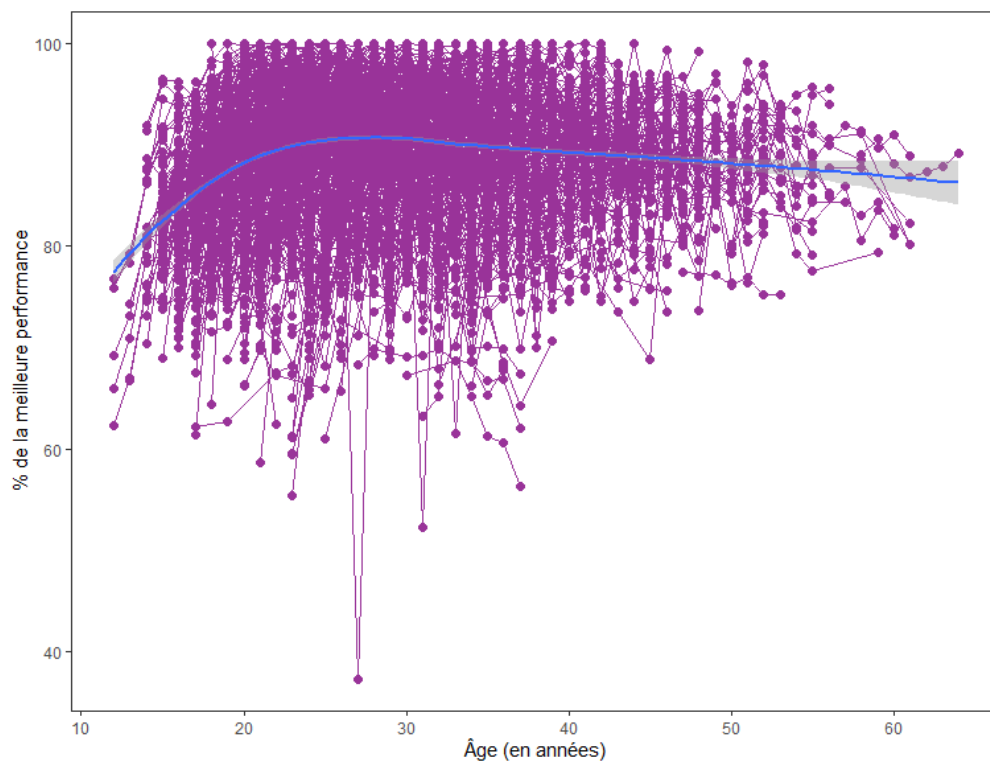


FIGURE 14 – Trajectoires de performance en athlétisme. Échantillon de 7 ou plus performances annuelles. Courbe bleue : smoothing méthode "loess".

```

### Définition de la fonction IMAP
# psi : matrice des paramètres IMAP (5 colonnes:a,b,c,d,e)
# id : vecteur des indices
# xidep : variable dépendante (nb lignes = length id)
# resp : vecteur des prédictions de longueur length id
saemixFunction2 <- fonction(psi, id, xidep) {
  t <- xidep[, 1]
  a <- psi[, 1]
  b <- psi[, 2]
  c <- psi[, 3]
  d <- psi[, 4]
  e <- psi[, 5]
  res <- a*exp((b/c)*(1-exp(-c*t)))*(1-exp(d*(t-e)))
  return(resp)
}

### Données à utiliser
saemix_data <- saemixData(name.data= cases_ALL2, # table données
name.group = "ID_athD", # variable de grouping
name.predictors = "Age", # temps
name.response = "Perf", # variable réponse
units = list(x = "Années", y = "% Performance"), # labels
name.covariates = c("Impairment_bin","Gender_bin","age_init","Rank"),
name.X = "Age")

### Définition du modèle
test_IMAP <- saemixModel(model = saemixFunction2, # fonction IMAP
description = "Saemix IMAP Natation", # titre (facultatif)
psi0 = c(a=parametre[1],b=parametre[2],c=parametre[3],
d=parametre[4],e=parametre[5]), # initialisation paramètres IMAP (moindres carrés)
fixed.estim = c(1,1,1,1,0), # estimation des effets fixes
transform.par = c(1,1,1,1,1)) # distribution log-normale paramètres
# matrice de covariance : argument covariance.model (défaut identité)
# matrice covariables : argument covariate.model (nb covar x nb paramètres)

### Run
RUN_saemix <- saemix(test_IMAP, saemix_data) # possible + options : seed etc

```

FIGURE 15 – Code modèle SAEMIX.

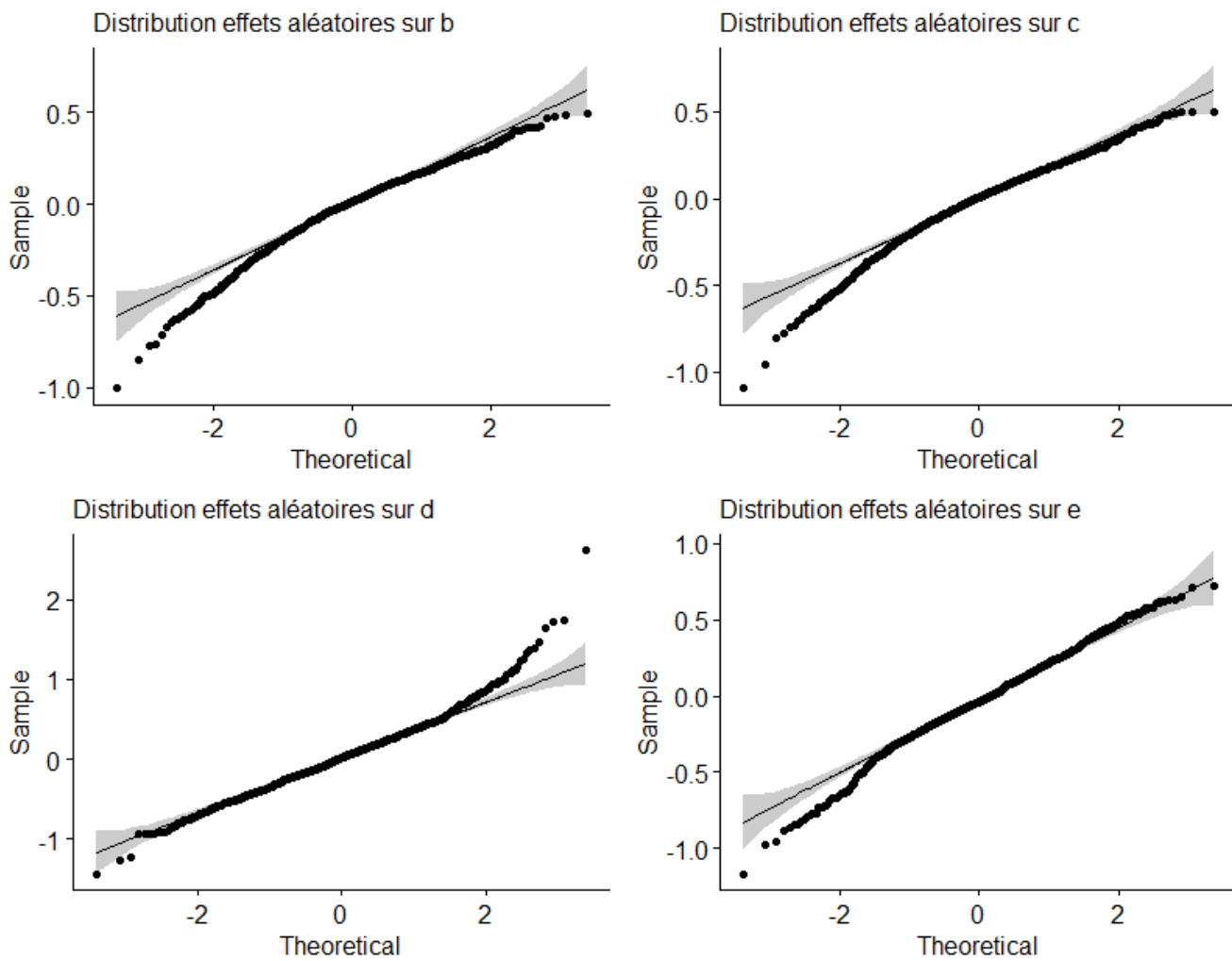


FIGURE 16 – Distributions des effets aléatoires. Modèle mixte multivariable. Natation.

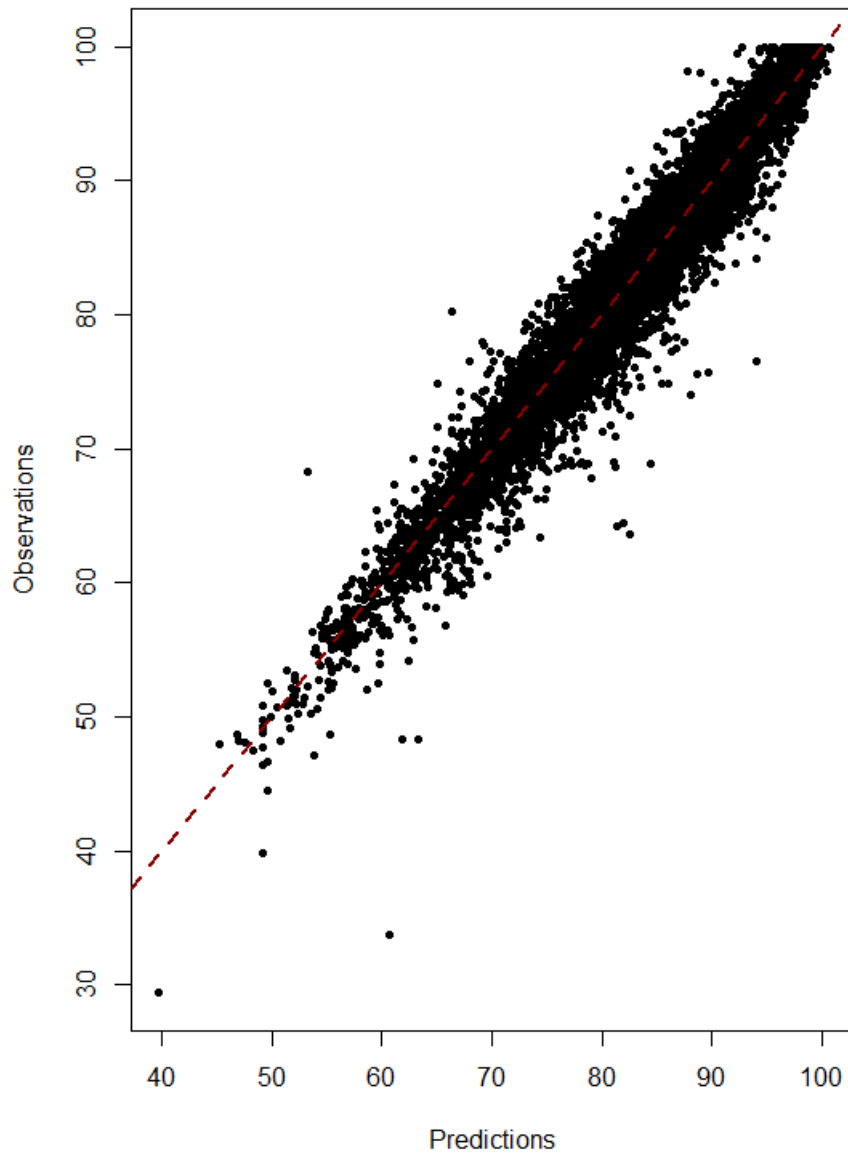


FIGURE 17 – Predictions versus observations. Modèle mixte multivariable. Natation.

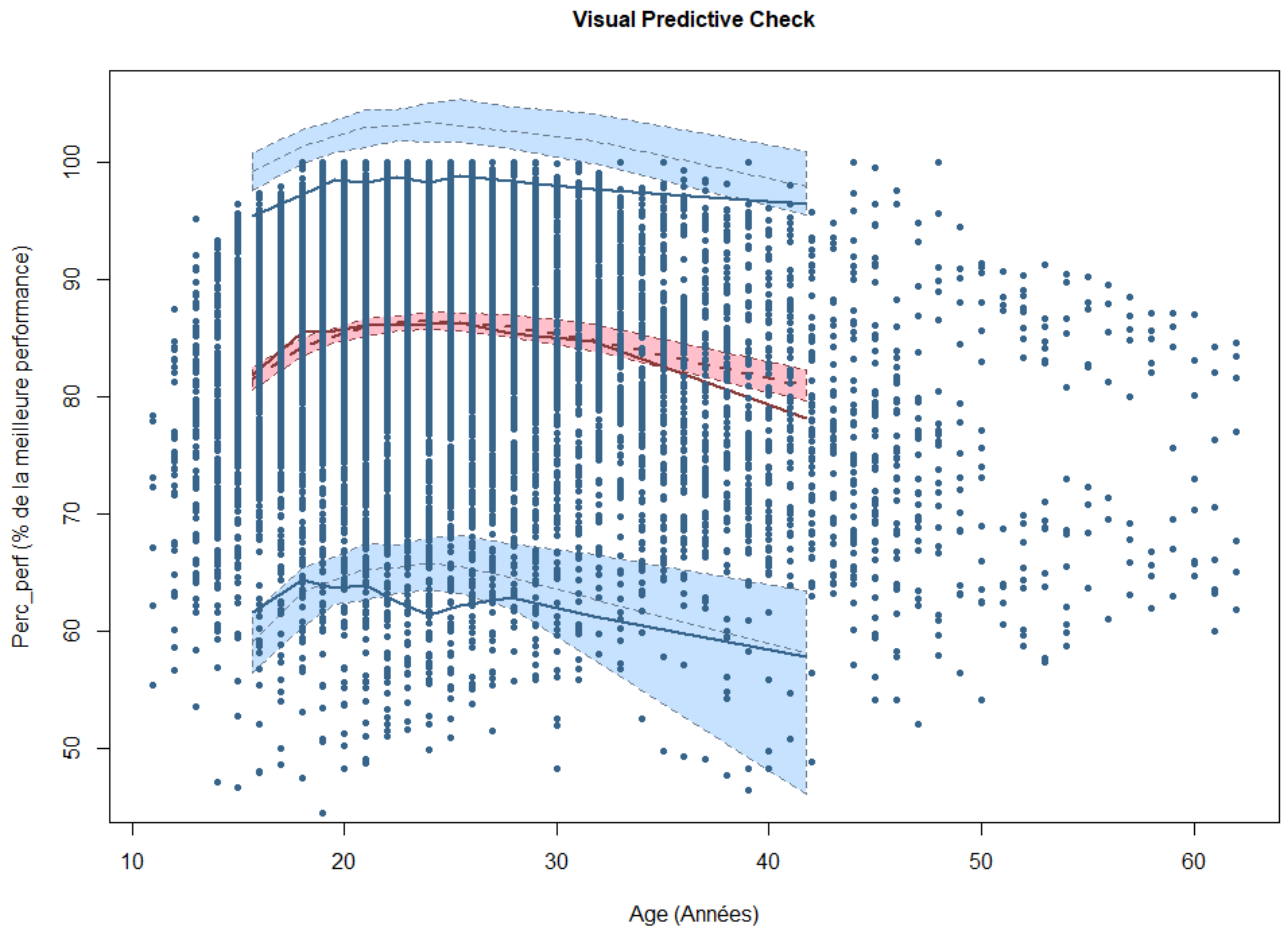


FIGURE 18 – Visual predictive check (vpc). Modèle mixte multivariable. Natation.

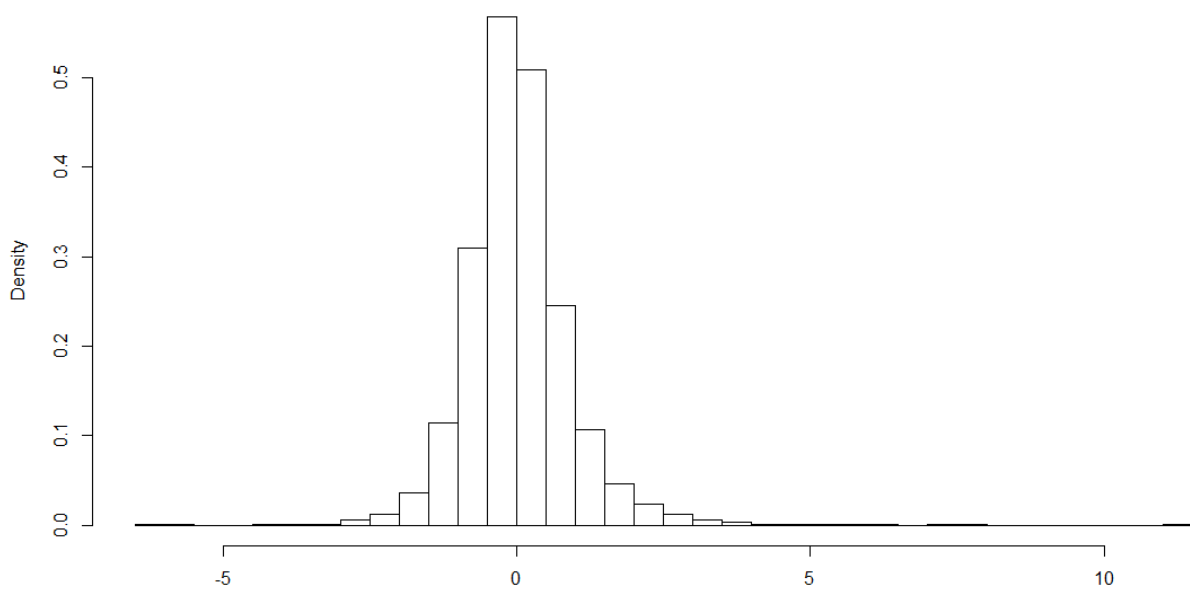


FIGURE 19 – Distribution des erreurs. Modèle mixte multivariable. Natation.

Annexe 9

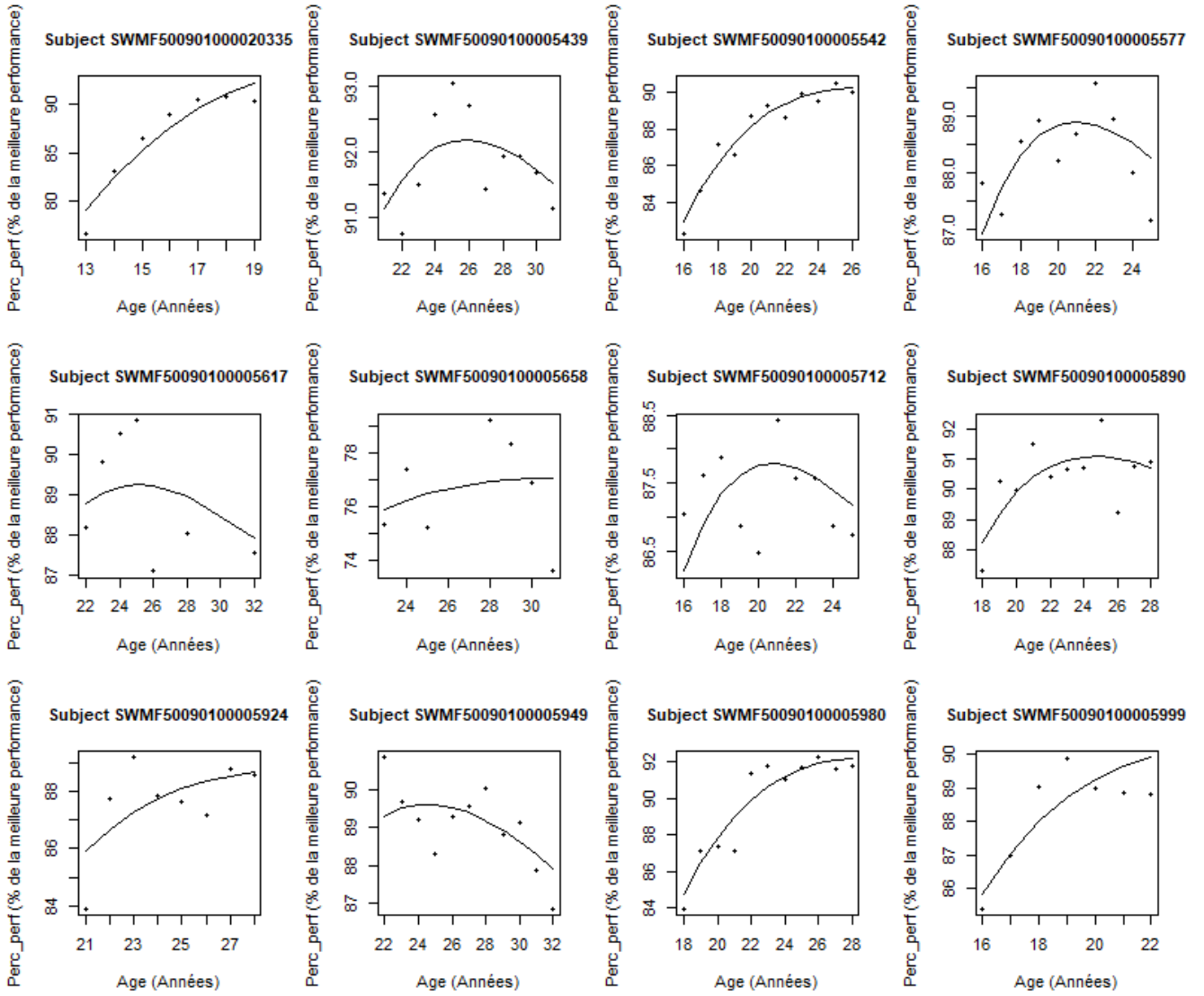


FIGURE 20 – Exemples de fit individuels. Modèle mixte multivariable. Natation.

Annexe 10

TABLE 13 – Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de course.

Paramètres	Effets fixes (ET% ¹)	Variance effets aléatoires (ET% ¹)	<i>p</i>
<i>a</i>	0,63(14,5%)	-	-
<i>b</i>	1,86(5,7%)	0,06(7,62%)	-
<i>c</i>	0,37(4,5%)	0,07(7,7%)	-
<i>d</i>	0,03(10,9%)	0,36(84,6%)	-
<i>e</i>	122(-)	0,27(42,4%)	-
<i>a</i> . ²	2,9(0,8%)	-	-
$\beta_{handicap,b}$	0,06(36,5%)	-	<0,01
$\beta_{sexe,b}$	0,06(31,6%)	-	<0,001
$\beta_{age-initial,b}$	-0,02(6,6%)	-	<0,001
$\beta_{position,b}$	-0,18(11,7%)	-	<0,001
$\beta_{handicap,c}$	0,06(35,2%)	-	<0,01
$\beta_{sexe,c}$	0,06(33,4%)	-	<0,01
$\beta_{age-initial,c}$	-0,02(7,0%)	-	<0,001
$\beta_{age-initial,e}$	0,01(26,9%)	-	<0,001

1 : Coefficient de variation $\Rightarrow ET\% = (\text{Écart-type}/\text{Estimation}) * 100$.

2 : Constante d'erreur.

TABLE 14 – Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de lancer.

Paramètres	Effets fixes (ET% ¹)	Variance effets aléatoires (ET% ¹)	<i>p</i>
<i>a</i>	16,6(12,6%)	0,38(10,8%)	-
<i>b</i>	0,16(9,0%)	-	-
<i>c</i>	0,05(14,8%)	0,21(11,6%)	-
<i>d</i>	0,006(22,0%)	0,6(38,7%)	-
<i>e</i>	122(-)	0,15(26,3%)	-
<i>a</i> . ²	4,9(1,3%)	-	-
$\beta_{age-initial,a}$	0,01(38,5%)	-	<0,01
$\beta_{handicap,b}$	0,12(40,61%)	-	<0,01
$\beta_{age-initial,b}$	-0,01(37,0%)	-	<0,01
$\beta_{handicap,c}$	0,26(26,7%)	-	<0,001
$\beta_{age-initial,c}$	0,01(34,6%)	-	<0,01
$\beta_{position,c}$	0,06(23,3%)	-	<0,001
$\beta_{handicap,d}$	1,00(17,5%)	-	<0,001
$\beta_{sexe,d}$	0,14(30,0%)	-	<0,001
$\beta_{age-initial,d}$	0,04(16,6%)	-	<0,001
$\beta_{handicap,e}$	-0,44(17,3%)	-	<0,001

1 : Coefficient de variation $\Rightarrow ET\% = (\text{Écart-type}/\text{Estimation}) * 100$.

2 : Constante d'erreur.

TABLE 15 – Résultats modèle mixte IMAP après sélection de modèle. Analyse multivariable. Athlétisme, épreuves de saut.

Paramètres	Effets fixes (ET% ¹)	Variance effets aléatoires (ET% ¹)	<i>p</i>
<i>a</i>	0,06(90,3%)	-	-
<i>b</i>	1,80(17,1%)	0,08(27,0%)	-
<i>c</i>	0,24(7,8%)	0,07(28,0%)	-
<i>d</i>	0,04(93,0%)	-	-
<i>e</i>	122(-)	0,28(80,0%)	-
a. ²	3,2(3,0%)	-	-
$\beta_{handicap,b}$	0,04(237,5%)	-	0,34
$\beta_{handicap,c}$	0,04(236,9%)	-	0,34
$\beta_{age-initial,c}$	7,6e-4(36,5%)	-	<0,01
$\beta_{sexe,d}$	0,18(200,6%)	-	0,31
$\beta_{age-initial,d}$	-0,032(192,5%)	-	0,30
$\beta_{age-initial,e}$	0,04(94,7%)	-	0,15

1 : Coefficient de variation $\Rightarrow ET\% = (\acute{E}cart\text{-}type/Estimation)*100$.

2 : Constante d'erreur.

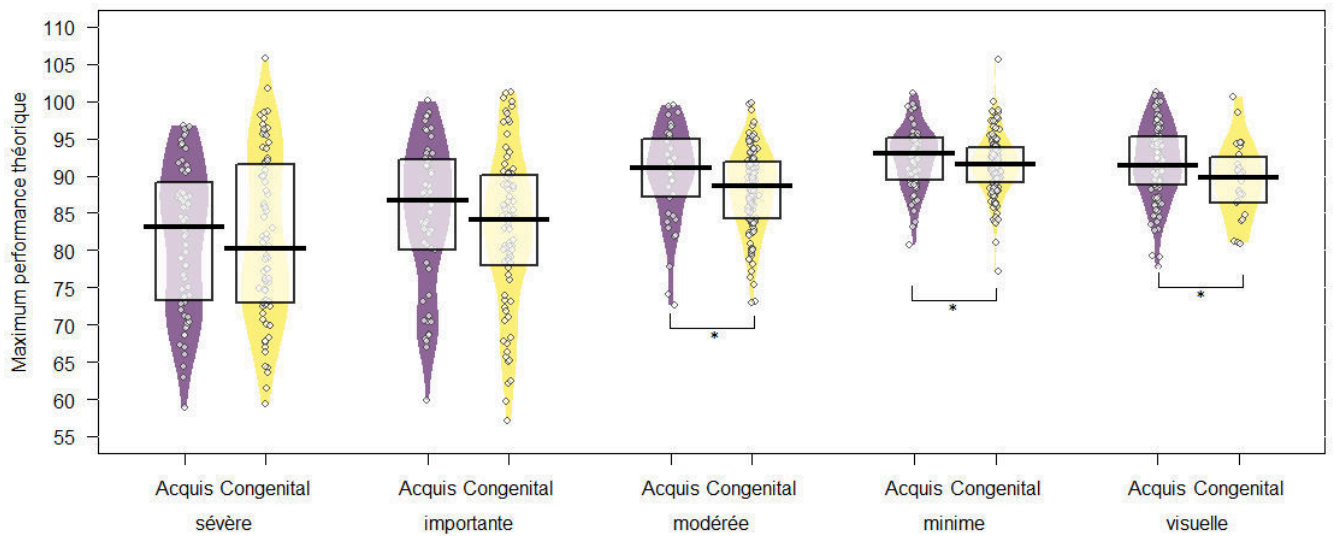


FIGURE 21 – Distribution des maximums de performance théorique en fonction de l’origine du handicap et de sa sévérité. Natation, hommes (n=779).

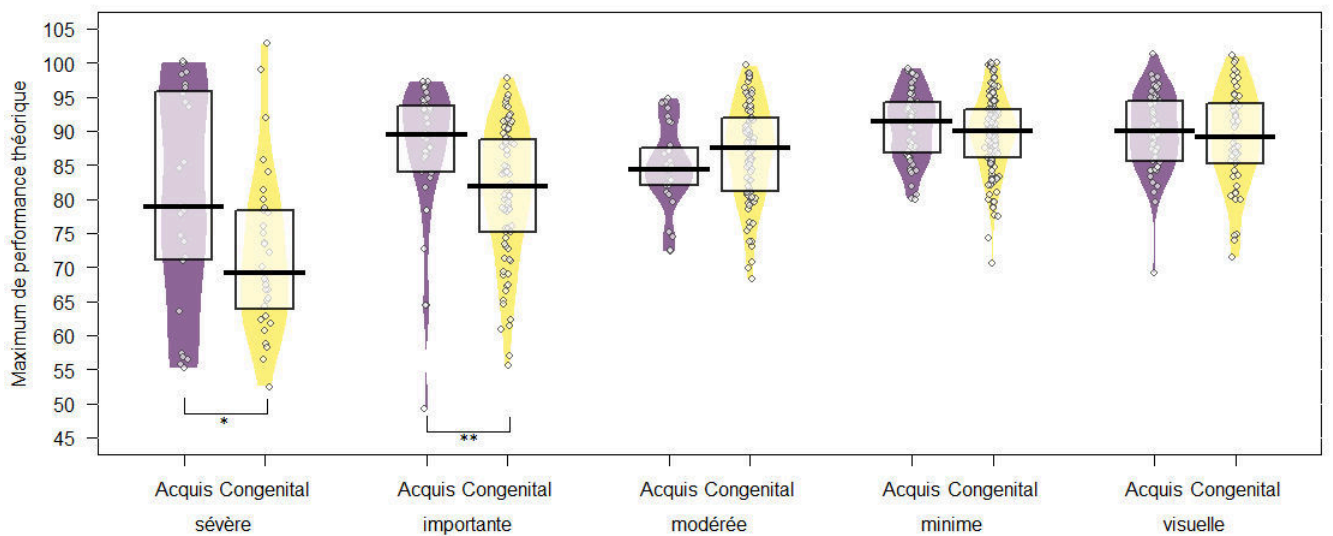


FIGURE 22 – Distribution des maximums de performance théorique en fonction de l’origine du handicap et de sa sévérité. Natation, femmes (n=621).

Annexe 12

On part de l'équation d'IMAP :

$$P(t) = \beta_0 N_0 \cdot e^{\frac{\alpha_0}{\alpha_r}(1-e^{-\alpha_r t})} \cdot (1 - e^{\beta_r(t-t_d)})$$

- En développant et en utilisant la formulation $e^{a+b} = e^a \times e^b$ on a :

$$P(t) = \beta_0 N_\infty \cdot e^{-\frac{\alpha_0}{\alpha_r} e^{-\alpha_r t}} \cdot (1 - e^{\beta_r(t-t_d)})$$

Et $N_\infty = N_0 \cdot e^{\frac{\alpha_0}{\alpha_r}}$ correspond à la limite de densité. Pour plus de lisibilité on utilisera ensuite les notations suivantes : $\beta_0 N_0 = a$, $\alpha_0 = b$, $\alpha_r = c$, $\beta_r = d$ et $t_d = e$, donnant ainsi :

$$P(t) = a \cdot e^{-\frac{b}{c} e^{-ct}} \cdot (1 - e^{d(t-e)})$$

- La dérivée de cette équation est :

$$\frac{\partial P(t)}{\partial t} = a e^{-\frac{b}{c} e^{-ct}} (b e^{-ct} - d e^{d(t-e)}) (1 - e^{d(t-e)})$$

- L'âge où la performance est maximale t_{max} correspond au point où la dérivée s'annule :

$$t_{max} := \arg \max_{t \in]0, e[} P(t) = \frac{\log(b/d) + de}{d + c}$$

car $a e^{-\frac{b}{c} e^{-ct}} > 0$ et $1 - e^{d(t-e)} > 0$ car e correspond à l'âge maximal
Cela revient donc à résoudre $b e^{-ct} - d e^{d(t-e)} = 0$

- Et la performance maximale est :

$$\max P(t) = P(t_{max}) = a \cdot e^{-\frac{b}{c} e^{-ct_{max}}} \cdot (1 - e^{d(t_{max}-e)})$$