



HAL
open science

Prédiction de la performance en course à pied : demi-fond et fond

Lucie Lerebourg

► **To cite this version:**

Lucie Lerebourg. Prédiction de la performance en course à pied : demi-fond et fond. Education. Normandie Université, 2021. Français. NNT : 2021NORMR097 . tel-03593527

HAL Id: tel-03593527

<https://theses.hal.science/tel-03593527v1>

Submitted on 2 Mar 2022

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.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité STAPS

Préparée au sein de l'Université de Rouen Normandie

Prédiction de la performance en course à pied : demi-fond et fond

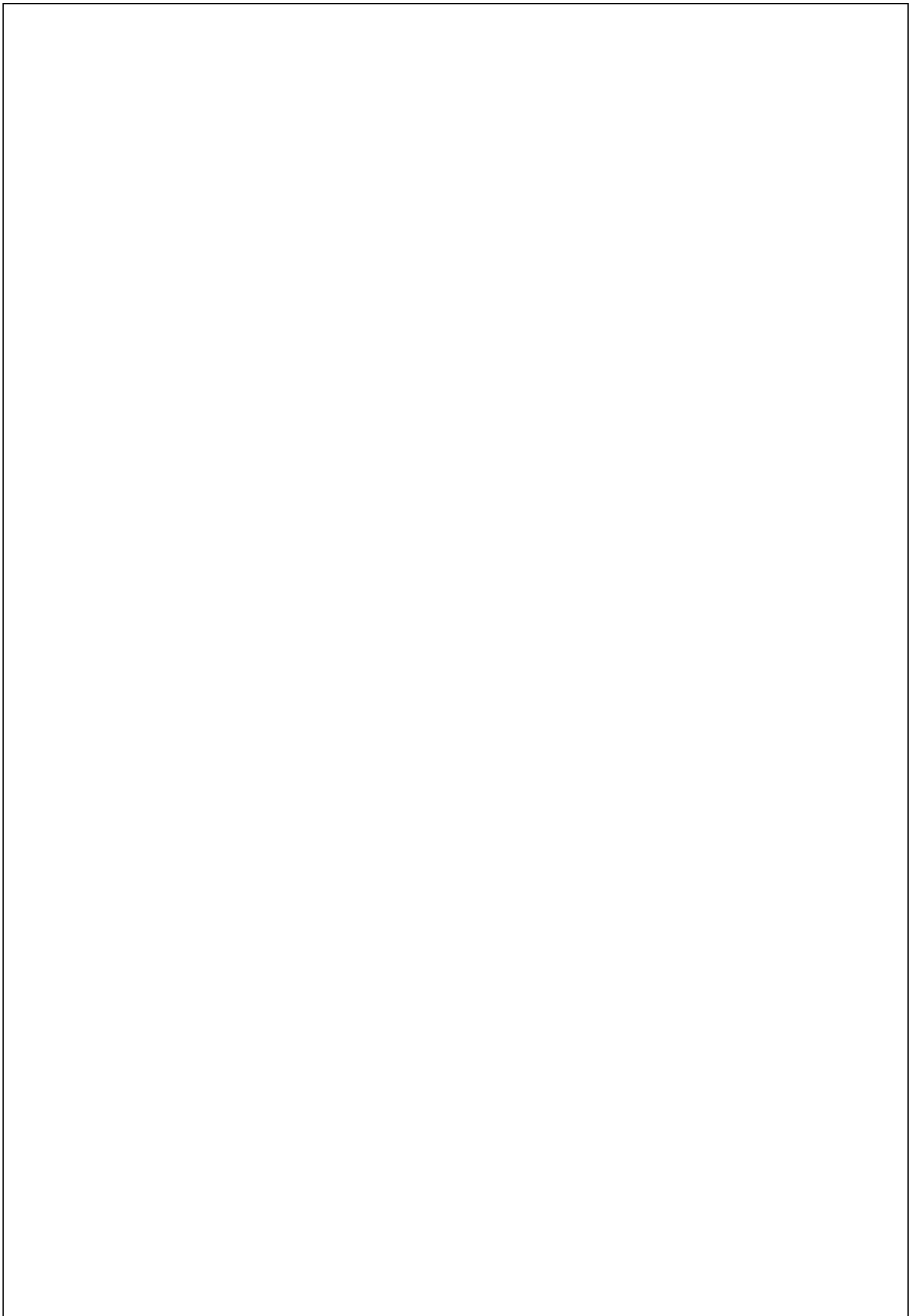
**Présentée et soutenue par
LUCIE LEREBOURG**

**Thèse soutenue le 15/12/2021
devant le jury composé de**

MME VERONIQUE BILLAT	PROFESSEUR DES UNIVERSITES, UNIVERSITE PARIS 11 PARIS-SUD	Rapporteur du jury
M. TARAK DRISS	PROFESSEUR DES UNIVERSITES, UNIVERSITE PARIS 10 PARIS-NANTERRE	Rapporteur du jury
MME NATACHA HEUTTE	PROFESSEUR DES UNIVERSITES, Université de Rouen Normandie	Président du jury
M. JEREMY COQUART	PROFESSEUR DES UNIVERSITES, Université de Rouen Normandie	Directeur de thèse

Thèse dirigée par JEREMY COQUART, CENTRE D'ETUDES DES TRANSFORMATIONS DES ACTIVITES PHYSIQUES ET SPORTIVES

**Prédiction de la performance en course à pied :
demi-fond et fond**



« *Mes pensées dorment si je les assis, mon esprit ne va si les jambes ne l'agitent* »

[Les Essais, Michel de Montaigne, XVIème siècle]

« *Sans douleur, le coureur est idiot. Apaisé, certes, mais idiot. La douleur apprend l'intelligence, la gestion du temps, de l'effort, l'attention aux muscles, la recherche permanente de moyens intérieurs pour avancer quoi qu'il arrive* »

[Petit éloge du running, Cécile Coulon, 2018]

« *Au fond, ne courons-nous pas pour ça ? Pour savoir que nous sommes capables de vaincre nos peurs et que le ruban que nous coupons en traversant la ligne d'arrivée n'est pas tenu par des hôtessees mais qu'il se trouve plutôt là où nous le voulons ?* »

[Courir ou mourir, Kilian Jornet, 2015]

« *Il faudra que je coure tous les jours
Faudra-t-il que je coure jusqu'au bout ?* »

[Je cours, Kyo, 2003]

« *J'ai appris à perdre à gagner sur les autres et le temps
À coups de revolver de course en entraînement
Les caresses étranges de la foule
Les podiums et les coups de coude
Les passions le monde et l'argent* »

[Le coureur, Jean-Jacques Goldman, 1997]

« *Si tu veux un conseil, oublie que t'as aucune chance. On sait jamais, sur un malentendu ça peut marcher* »

[Les bronzés font du ski, Patrice Leconte, 1979]

« *Cours, Forrest !* »

[Forrest Gump, Robert Zemeckis, 1994]

...

REMERCIEMENTS

Courir après le temps et soutenir dans les « délais », tels étaient les objectifs de ce double projet. Doctorante-salariée et athlète « Oné Oné » (#HOKA), l'équilibre fut rapidement trouvé. Au-delà du sujet, recentré « course à pied », le parallélisme thèse-athlé m'a permis d'avancer. Le dossard de thésard épinglé, pied sur la ligne, « à vos marques, prêts, partez » [coup de pistolet], je savais que je serais libérée le jour de la ligne dépassée. La thèse une course, certes, contre le temps, mais avant tout une course qui m'a poussée dans mes retranchements. Coureuse de « Quinze cent » voilà ce que j'avais pu imaginer comme séance de fractionné sur ces 3 années : un premier 500 m pour « se lancer » (et ça n'a pas loupé le rythme fut endiablé chez OD), un deuxième 500 m pour « temporiser » (le covid a vraiment bien aidé, #parodies) et un dernier 500 m pour « relancer » et tout « donner » (j'ai rédigé comme « jamais »). En réalité 42 mois se sont écoulés... et une fois le calcul refait (*i.e.*, 1 mois = 1 km, ce n'est pas trop compliqué...), j'ai pu réaliser sur quelle spécialité je m'étais alignée (*i.e.*, marathon ???). Qu'importe la distance, ce fut vraiment intense et de toute évidence une très belle expérience (d'endurance...XD).

MERCI...

Pour votre patience, bienveillance, confiance...

J'ai conscience de la chance d'avoir pu réaliser

avec vous cette

« Thèse-Performance » !!!

Je tiens tout d'abord à exprimer ma reconnaissance aux membres du jury de m'avoir fait l'honneur d'accepter de juger mon travail.

Je tiens également à remercier ma direction de thèse, le Professeur **Jérémy COQUART** ainsi que le Docteur **Maxime L'HERMETTE**, avec qui j'ai pu évoluer et échanger durant ces trois belles années. Jérémy (**#Jéjé**), mon directeur de thèse, je te dois un très grand MERCI pour ton suivi, ta réactivité, ta rigueur de travail et ta disponibilité, malgré la distance (**#1mdesécuritécovidobligeLOL**), qui m'auront inévitablement emmenée vers la ligne d'arrivée. Tu resteras pour moi un chasseur de coquilles hors normes !!! Et je garderai bien évidemment en souvenirs nos échanges de mails et commentaires humoristiques dissimulés dans les fichiers ! Maxime (**#Max**), mon co-encadrant, je tiens également à te remercier de m'avoir encouragé, soutenu et relu jusqu'à la fin de ce projet malgré le fait que le sujet ait « quelque peu changé ». La « prédiction de la perf » en course à pied... pas vraiment ta tasse de thé, mais tu as réussi à l'appivoiser.

Je remercie mon directeur d'entreprise, **Éric HELD**, pour la confiance et la place accordées au sein de chez OD. De beaux projets ont été menés et l'histoire n'est pas terminée... On devrait bientôt pouvoir innover !!! Je remercie au passage, tous mes collègues de chez OD, *OD Run* et *OD Cycling*, de m'avoir épaulée tout au long de ce projet. Je me permets toutefois de faire une petite dédicace particulière à ma collègue chercheuse, ex-doctorante, et surtout amie **Charlotte MENEZ (#Chacha)** pour sa confiance et son soutien sans faille durant ces trois belles années. La grosse « marrade » ne fait que continuer !

Je remercie le Centre d'Études des Transformations Physiques et Sportives (CETAPS), dans lequel j'ai eu la chance d'évoluer, ainsi que l'ensemble de mes collègues qui y sont rattachés. Comment ne pas mentionner la **#TeamdesDocs** (qui

ne cesse de s'agrandir avec les années...) et plus particulièrement la « **Roocky Family** », une famille soudée, qui m'a tant appris et tant donné... Les souvenirs resteront gravés à jamais... Et puis avec **#Amande**, ma complice... à nos « FouRires », nos « conneries » en format ILLIMITÉ !!!



PS : Je remercie au passage, les pongistes pour le nombre de balles cassées, perdues et échouées sur mon écran de PC. J'en profite également pour féliciter tout particulièrement mon petit **#Wernouz** pour ses nombreux matchs gagnés ainsi que pour sa victoire au classement « ATP CETAPS » amplement méritée.

#TeamCETAPS : **#Amande**, **#Caro**, **#Guigui**, **#LM(Elaime)**, **#MaChacha**, **#MaGio**, **#Major**, **#MonPtitBrissouille**, **#MonPtitClem(J)**, **#MonPtitDam**, **#MonPtitMimine**, **#MonPtitWernouz**, **#MonPtitWill**, **#Nana**, **#PapyRom**, **#Pierrot** ... Et sans oublier les petits derniers... **#Aya**, **#Clem**, **#Jérémy** et **#Quentin** ...

Je remercie également toutes les « *Teams* » sur qui je peux compter et qui m'ont encouragée dans mes différents projets durant ces trois dernières années... **#TeamHoka**, **#4Chagasses**, **#TeamGrard**, **#TeamLavandou<3**, **#LeCercle**, **#CetapsGirls**, **#TeamOD**, et dernièrement la **#TeamLepape** (petite pensée à **#JC** également) ... Pour les oublis...je vous prie de m'en excuser...

Enfin, je tiens à remercier mes amis, ma famille pour tout ce qu'ils m'ont apporté... même s'il m'est impossible de tous vous citer : **#PapyLampérier**, merci de m'avoir remis le pied à l'étrier et de me « culturer » au quotidien ; **#MonDridri**, je ne compte plus les années... merci d'avoir toujours été là depuis le CP ; **#Valou**, **#Nono**, **#Yas**, **#VicLaCrevette**, que ce soit en athlé ou sur les bancs de l'université...merci pour le soutien depuis... tant d'années ; **#MonPtitBrissouille**, merci d'être là... à mes côtés -77- ; **#TatyDan** & **#TontonMarc**, **#Papa** (« JohnyBigoude »), **#Maman** (« MonPtitPoussin ») & **#Pierre** (« TheBro »), mon socle familial, mon véritable pilier, merci de m'avoir aidée dans la réalisation de mes projets (appartement, thèse, athlé...).

**En réalité, je ne sais comment tous vous
remercier...**

LISTE DES ABRÉVIATIONS

95% LoA *Limites d'Agrément à 95%*

ANN *Réseaux de Neurones Artificiels, de l'anglais : Artificial Neural Networks*

ANOVA *Analyse de Variance, de l'anglais : Analysis of Variance*

ANRT *Association Nationale de Recherche et Technologie*

$Ca_{O_2}-Cv_{O_2}$ *Différence artério-veineuse en oxygène*

CDA *Capacité de Distance Anaérobie*

CE *Coût Énergétique*

CERSTAPS *Comité d'Éthique pour la Recherche en Sciences et Techniques des Activités Physiques et Sportives*

CETAPS *Centre d'Études des Transformations des Activités Physiques et Sportive*

CIFRE *Convention Industrielle de Formation par la REcherche*

CIO *Comité International Olympique*

EDSH *Structure scientifique sous le nom de "Efficiency & Dynamics in Sport & Health"*

EMS *Éducateurs Médico-Sportifs*

EQM *Erreur Quadratique Moyenne*

ES *Taille de l'Effet, de l'anglais : Effect Size*

FC *Fréquence Cardiaque*

FC_{max} *Fréquence Cardiaque Maximale*

FFA *Fédération Française d'Athlétisme*

FSFI *Fédération Sportive Féminine Internationale*

GPS *Système de Positionnement Général, de l'anglais : Global Positioning System*

HSRT *École Doctorale "Homme Société Risques et Territoires"*

IA *Intelligence Artificielle*

IAAF *Fédération Internationale d'Athlétisme, de l'anglais : International Association of Athletics Federations*

IC 95% *Intervalle de Confiance à 95%*

ICC *Coefficient de Corrélation Intra-classe*

IEAP *Ingénierie et Ergonomie de l'Activité Physique*

IMC *Indice de Masse Corporelle*

JO *Jeux Olympiques*

K-NN *K-Plus Proches Voisins, de l'anglais : K-Nearest Neighbors*

LoA *Limites d'Agrément*

MAE *Erreur Moyenne Absolue, de l'anglais : Mean Absolute Error*

MLDT *Modèle Linéaire Distance-Temps*

MP *Meilleure Performance*

MP-1 *MP de l'année précédente (-1) : MP de la saison précédant immédiatement la MP*

MP-MS *MP Multi Saisons*

Prédiction de la performance en course à pied : demi-fond et fond

OD *Orthodynamica*

Qc *Débit Cardiaque*

R&D *Recherche et Développement*

RGPD *Règlement Général sur la Protection des Données*

SD *Écart-Type, de l'anglais : Standard Deviation*

TA-1S *Taux d'Amélioration sur 1 Saison : saison précédant la MP*

TA-MS *Taux d'Amélioration Multi Saisons*

UFR STAPS *Unité de Formation de Recherche en Sciences et Techniques des Activités Physiques et Sportives*

URN *Université de Rouen-Normandie*

USFCP *Union des Sociétés Françaises de Courses à Pied*

USFSA *Union des Sociétés Françaises des Sports Athlétiques*

VDOT *V POINT (sur le V), de l'anglais : V DOT*

VES *Volume d'Éjection Systolique*

VIF *Facteur d'Inflation de la Variance, de l'anglais : Variance Inflation Factor*

VMA *Vitesse Maximale Aérobie*

$\dot{V}O_{2max}$ *Débit Maximal d'Oxygène*

LISTE DES FIGURES

Figure 1. Dispositif CIFRE associant les trois partenaires	24
Figure 2. Centre « OD Rouen ».....	26
Figure 3. Organigramme « OD FRANCE ».....	29
Figure 4. Organisation prévisionnelle des trois années de doctorat.....	31
Figure 5. Aménagement de la thèse à la suite de la crise sanitaire (covid-19).....	32
Figure 6. Planification de la thèse à la suite de la crise sanitaire (covid-19).....	33
Figure 7. Mise en place des premiers championnats nationaux	38
Figure 8. Création des premiers clubs et de la fédération (en France)	38
Figure 9. Relation entre la quantité totale d'énergie utilisée en fonction de la longueur et de la durée d'épreuve de course sur piste, d'après Howald [71].....	43
Figure 10. Évolution chronologique des dates d'entrée des courses de demi-fond et de fond dans le programme olympique, depuis la première édition de 1896 à aujourd'hui, chez les hommes et les femmes.....	46
Figure 11. Évolution chronologique des dates d'entrée aux championnats de France d'athlétisme des courses de demi-fond et de fond, depuis la première édition de 1888 à aujourd'hui, chez les hommes et les femmes.....	46
Figure 12. Disciplines de l'athlétisme sous différentes formes (<i>i.e.</i> , stade et hors-stade).....	53
Figure 13. Licences et secteurs d'activités en fonction des catégories d'âge.....	55
Figure 14. Évolution des licenciés FFA de 1999 à 2020 [96]	57
Figure 15 Performance en course à pied.....	62
Figure 16. Exemple d'épreuves du barème Femme au 1er janvier 2010, d'après la FFA [99].....	64
Figure 17. Relation vitesse-temps en course à pied et en natation [8, 101].....	65
Figure 18. Exemple de la relation vitesse-temps des records du monde hommes actuels (au premier juillet 2021) en course à pied, allant des épreuves de 100 m au 100 km.....	67
Figure 19. Illustration des différentes courbes utilisées couramment dans l'analyse des records (olympiques et du monde), d'après King et Black [102] et Billat [8]	68
Figure 20. Courbes vitesse-distance illustrant l'évolution des premiers records du monde sur piste masculins (1908-1912) et les premiers records de France masculins (1892-1895) avec les records du monde et de France actuels masculins (au 1 ^{er} juillet 2021).....	75
Figure 21. Facteurs de performance en course à pied.....	78
Figure 22. Contributions relatives des filières énergétiques en fonction des distances de courses en demi-fond et fond	80
Figure 23. Évolution de la VMA en fonction de la distance d'épreuves de demi-fond et fond, d'après Aubert et Choffin [60]	84
Figure 24. Exemples de tests de laboratoire ou de terrain	99

Prédiction de la performance en course à pied : demi-fond et fond

Figure 25. Méthodes de prédiction de performance en course à pied grâce à la compétition	103
Figure 26. Illustration des deux expressions utilisées dans le modèle VDOT	106
Figure 27. Nomogramme de Mercier et al. [26]	111
Figure 28. Exemples de prédiction de la performance par interpolation (en bleu) et extrapolation (en rouge) avec le nomogramme de Mercier et al. [26]	112
Figure 29. Premier nomogramme de Vandewalle [27].....	116
Figure 30. Second nomogramme de Vandewalle [27].....	117
Figure 31. Troisième nomogramme de Vandewalle [27].....	118
Figure 32. Caractéristiques des nomogrammes de Vandewalle [27].....	119
Figure 33. Exemples de prédiction de la performance sur 10 000 m par extrapolation avec le premier nomogramme de Vandewalle [27].....	121
Figure 34. Exemples de prédiction de la performance sur 10 000 m par extrapolation avec le second nomogramme de Vandewalle [27].....	122
Figure 35. Évolutions des vitesses de records du monde actuels en course à pied en fonction de la distance selon deux lois de puissance différentes dépendantes de la filière énergétique dominante	127
Figure 36. Exemple du processus automatique utilisé pour prédire la performance en course à pied (illustration avec le nomogramme de Mercier et al. [26]).....	158
Figure 37. Validité des mesures avec le nomogramme de Mercier et al. [26] pour prédire la performance....	162
Figure 38. Validité des mesures avec le premier nomogramme de Vandewalle [27] pour prédire la performance	163
Figure 39. Validité des mesures avec le second nomogramme de Vandewalle [27] pour prédire la performance	164
Figure 40. Validité des mesures avec le modèle linéaire distance-temps (MLDT) pour prédire la performance	179
Figure 41. Validité des mesures avec la loi de puissance pour prédire la performance.....	180
Figure 42. Validité des mesures avec l'équation simple pour prédire la performance.....	191
Figure 43. Architecture d'un réseau de neurones.....	200
Figure 44. Validité des mesures avec l'algorithme ANN pour prédire la performance	204
Figure 45. Validité des mesures avec l'algorithme KNN pour prédire la performance.....	205
Figure 46. Définitions des termes : MP, MP-1, MP-MS, TA-1S et TA-MS chez un coureur masculin de 5 000 m (sur 6 ans).....	214
Figure 47. Descriptifs de chaque variable en fonction du sexe et de la distance des dix meilleurs athlètes français de tous les temps en demi-fond (moyenne et SD).....	219
Figure 48. Index du taux d'amélioration pour les courses de 800 m, 1 500 m et 5 000 m pour les femmes et les hommes.....	223

LISTE DES TABLEAUX

Tableau 1. Exemple d'épreuves de courses de demi-fond et de fond sur piste proposées par la FFA en championnats en fonction des catégories et du sexe.....	41
Tableau 2. Classification des distances de courses du 800 m au marathon en fonction de l'intensité de la VMA, d'après Aubert et Choffin [60].....	42
Tableau 3. Exemple de contribution des systèmes énergétiques en fonction de la distance de course, d'après Dyer et Dwyer [59]	43
Tableau 4. Détails des licences et secteurs d'activités en fonction des catégories d'âge proposées par la FFA [96].....	56
Tableau 5. Records du monde actuels masculins et féminins en course à pied, du 100 m au 100 km (au 1 ^{er} juillet 2021)	69
Tableau 6. Records de France actuels masculins et féminins en course à pied, du 100 m au 100 km (au 1 ^{er} juillet 2021)	70
Tableau 7. Améliorations des records du monde masculins et féminins en demi-fond et fond lors de différentes périodes (de 1882 à 2017).....	72
Tableau 8. Évolutions des records du monde masculins et féminins en demi-fond et fond, du premier record homologué à 2001 et à aujourd'hui (2021), et évolutions sur ces vingt dernières années (2001-2021) ...	73
Tableau 9. Évolutions des records de France masculins et féminins en demi-fond et fond, du premier record homologué à aujourd'hui (2021)	74
Tableau 10. Exemple de contribution des métabolismes énergétiques en fonction de la distance de course et du sexe	79
Tableau 11. Pourcentage de VMA sollicitée en fonction de la distance de course en compétition (valeurs moyennes pour des coureurs relativement entraînés) d'après Cazorla et Leger [242].....	100
Tableau 12. Valeurs de VDOT associées aux chronos sur les distances les plus répandues, d'après Daniel's [84]	107
Tableau 13. Intensités d'entraînement basées sur la valeur VDOT actuelle, d'après Daniel's [84].....	108
Tableau 14. Barème FFA « Femme » du niveau de performance en course de demi-fond et de fond au 1 ^{er} janvier 2010	138
Tableau 15. Évolution des performances des athlètes françaises entre 2005 et 2016.....	141
Tableau 16. Évolution des pourcentages de records personnels chez les performeuses françaises entre 2005 et 2016.....	142
Tableau 17. Évolution de l'âge des athlètes françaises entre 2005 et 2016	143
Tableau 18. Évolution du nombre et du pourcentage d'athlètes françaises selon chaque niveau de performance entre 2005 et 2016.....	144
Tableau 19. Coefficient de corrélation intra-classe (ICC) entre le programme MATLAB et les méthodes manuelles.....	156

Prédiction de la performance en course à pied : demi-fond et fond

Tableau 20. Résultats des prédictions à partir du nomogramme de Mercier et al. [26] et des deux nomogrammes de Vandewalle [27] chez 158 coureuses.....	161
Tableau 21. Résultats des prédictions à partir du modèle linéaire distance-temps (MLDT), de la loi de puissance et du nomogramme de Mercier et al. [26] chez 158 coureuses.....	178
Tableau 22. Données descriptives et variables utilisées pour développer l'équation simple de prédiction (valeurs moyennes \pm SD)	189
Tableau 23. Résultats des prédictions à partir des algorithmes ANN et KNN	203
Tableau 24. Données extraites des dix meilleurs records nationaux retenus par sexe et distance (Nom, MP, date et âge de la MP).....	215
Tableau 25. Données descriptives des dix meilleures performances françaises de tous les temps en demi-fond (valeurs moyennes, SD et étendue des données)	218
Tableau 26. Exemple d'un coureur de 5 000 m homme (sur 6 ans).....	229
Tableau 27. Récapitulatif des biais et LoA à 95% pour l'ensemble des méthodes testées : nomogrammes (de Mercier et al. [26] et de Vandewalle [27]), MLDT, loi de puissance, équation simple et IA (ANN et KNN)	234

LISTE DES ANNEXES

Annexe 1. Article 1 : <i>The effects of shoe type on lower limb venous status during gait or exercise: A systematic review</i>	272
Annexe 2. Article 2 : <i>Changes in performances/characteristics of French female runners over the last 12 years</i>	291
Annexe 3. Article 3 : <i>Prediction of middle-distance running performances of female runners using nomograms</i>	303
Annexe 4. Article 4 : <i>Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram</i>	325
Annexe 5. Article 5 : <i>Prediction of 3000-m steeplechase performance using a simple equation</i>	346
Annexe 6. Article 6 : <i>Prediction of marathon performance using artificial intelligence</i>	361
Annexe 7. Article 7 : <i>Predicting the potential of middle- and long- distance track runners by the evaluation of performance improvement rate</i>	385
Annexe 8. Projet : Conception et développement d'une gamme de Sandales Orthopédiques « sur-mesure » et personnalisables	411

LISTE DES PUBLICATIONS ET COMMUNICATIONS

❖ Articles scientifiques publiés

- **Lerebourg L**, Coquart JB. *Changes in performances/characteristics of French female runners over the last 12 years. Research in Sports Medicine.* 2021, 29 (2), 185-195.
- **Lerebourg L**, L'Hermette M, Menez C, Coquart JB. *The effects of shoe type on lower limb venous status during gait or exercise: A systematic review. PloS One.* 2020, 15(11), e0239787.
- **Lerebourg L**, Guignard B, Racil G, Jlid MC, Coquart JB. *Prediction of middle-distance running performances of female runners using nomograms. International Journal of Sports Medicine.* 2021,19 (10). Epub ahead of print.

❖ Articles scientifiques soumis

- **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. *Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram. Journal of Strength and Conditioning Research.* Soumis.
- Barbry A, **Lerebourg L**, Racil G, Jlid MC, Coquart JB. *Prediction of the 3000-m steeplechase performance using a simple equation. Research Quarterly for Exercise and Sport.* Soumis.

- **Lerebourg L**, Saboul D, Cléménçon M, Coquart JB. *Prediction of marathon performance using artificial intelligence. Artificial Intelligence in Medicine.* Soumis.
- Menez C, L'Hermette M, **Lerebourg L**, Coquart JB. *Effects of foot orthoses on gait kinematics and pain in subjects with leg length inequality. Journal American Podiatric Medical Association.* Soumis.
- Guignard B, Cléménçon M, Proiett G, **Lerebourg L**, Iodice P, L'Hermette M, Machemehl C, Coquart JB. *Physiological and biomechanical demands of the Street performance in elite skaters. Biology of Sport.* Soumis.
- **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. *Predicting the potential of middle-distance runners by the evaluation of performance improvement rate.* À soumettre.

❖ **Communications internationales avec acte**

- **Lerebourg L**, Coquart JB. Évolution des performances de moyennes et de longues distances sur piste et des caractéristiques des coureuses françaises lors des trois derniers cycles des Jeux Olympiques. 18^{ème} Congrès de l'ACAPS. 29-31 octobre 2019. Paris, France.
- **Lerebourg L**, L'Hermette M, Menez C, Coquart JB. *Effects of shoes on lower limb venous status during exercise : A systematic review. 5th Congress of Sport Medicine & Fitness.* 14-15 novembre 2019. Paris, France.

❖ **Communications sans acte**

- **Lerebourg L**, L'Hermette M, Menez C, Coquart JB. L'impact des sandales de récupération « sur-mesure », stimulant la semelle veineuse plantaire, sur la performance, la fatigue et la récupération du sportif. Séminaire Recherche CETAPS Axe Santé. 22 novembre 2018. Saint Sébastien de Morsent, France.
- **Lerebourg L**, L'Hermette M, Menez C, Coquart JB. Impact des sandales de récupération « sur-mesure », stimulant la semelle veineuse plantaire, sur la récupération, la fatigue et la performance chez le coureur à pied. 3^{ème} Doctoriales du CETAPS. 5-6 juin 2019. Mont-Saint-Aignan, France.
- **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. Prédire le potentiel des athlètes de demi-fond français via l'évaluation du taux d'amélioration des performances. 5^{ème} Doctoriales du CETAPS. 24-25 juin 2021. Mont-Saint-Aignan, France.
- **Lerebourg L**, Guignard B, Cléménçon M, Proiett G, Iodice P, L'Hermette M, Machemehl C, Coquart JB. De la rue aux Jeux olympiques : Analyse du street et des caractéristiques de skateurs français internationaux. Colloque, Skateboard. 22 octobre 2020. Mont-Saint-Aignan, France.
- **Lerebourg L**. Modélisation et prédiction de la performance en course à pied. Webinar OD-STAPS Rouen. 26 mai 2021. Rouen, France.

TABLE DES MATIÈRES

REMERCIEMENTS.....	4
LISTE DES ABRÉVIATIONS	8
LISTE DES FIGURES	10
LISTE DES TABLEAUX.....	12
LISTE DES ANNEXES.....	14
LISTE DES PUBLICATIONS ET COMMUNICATIONS	15
INTRODUCTION GÉNÉRALE.....	21
1. Présentation du dispositif de thèse	23
2. Présentation de l'entreprise OD	25
3. Présentation du plan et des objectifs de thèse.....	30
PREMIÈRE PARTIE : CADRE THÉORIQUE	36
CHAPITRE I. COURSE À PIED	37
1. Contexte historique.....	37
2. Définition et classification des disciplines : demi-fond et fond.	39
3. Pratiques sportives masculine et féminine	44
a. Évolution vers la parité en athlétisme	44
b. Dates d'entrée des épreuves de demi-fond et de fond dans les programmes des compétitions : à l'échelle nationale et olympique.....	47
i. 800 m (demi-fond).....	47
ii. 1 500 m (demi-fond)	48
iii. 3 000 m steeple (demi-fond).....	49
iv. 5 000 m (demi-fond)	50
v. 10 000 m (fond)	50
vi. Marathon (fond).....	51
4. État des lieux de la pratique en France : Le développement de la FFA.....	52
a. Projet fédéral et plan de développement de la FFA	53
b. Évolution des pratiques : les différentes licences et catégories	54
c. Évolution du nombre de pratiquants : une croissance importante	57
d. Évolution dans une optique de performance et d'efficacité	58
5. Conclusion.....	59
CHAPITRE II. PERFORMANCE EN COURSE À PIED.....	61
1. Performance en course à pied et évolution des performances.....	61
a. Définition	61
b. Modélisation et évolution des performances	65
2. Facteurs de performance	77
a. Paramètres physiologiques.....	78
b. Paramètres biomécaniques.....	86
c. Paramètres psychologiques.....	88

Prédiction de la performance en course à pied : demi-fond et fond

d.	Paramètres environnementaux.....	91
e.	Paramètres technologiques	93
3.	Conclusion.....	96
CHAPITRE III. PRÉDICTION DE LA PERFORMANCE EN COURSE À PIED.....		97
1.	Revue de littérature : Concepts et disciplines étudiés	97
a.	Détermination du profil et du potentiel de performance des coureurs à l'aide de tests de laboratoire ou de terrain.....	98
b.	Détermination du profil et du potentiel de performance des coureurs à l'aide des performances réalisées en compétition.....	102
2.	Présentation des modèles/outils testés pour la prédiction de la performance en demi-fond.....	110
a.	Nomogrammes.....	110
i.	De Mercier et al. [26].....	110
ii.	De Vandewalle [27].....	114
b.	Concept de vitesse critique.....	123
c.	Loi puissance.....	125
d.	Intelligence artificielle (IA).....	128
3.	Conclusion.....	130
DEUXIÈME PARTIE : CADRE EXPÉRIMENTAL		132
CHAPITRE I. ÉVOLUTION DES PERFORMANCES/CARACTÉRISTIQUES DES COUREUSES FRANÇAISES SUR LES 12 DERNIÈRES ANNÉES.....		136
1.	Introduction.....	136
2.	Méthodes.....	138
3.	Résultats.....	140
4.	Discussion.....	145
5.	Conclusion.....	151
CHAPITRE II. PRÉDICTION DES PERFORMANCES EN COURSE DE DEMI-FOND ET DE FOND À L'AIDE DE NOMOGRAMMES.....		152
1.	Introduction.....	152
2.	Méthodes.....	154
3.	Résultats.....	160
4.	Discussion.....	165
5.	Conclusion.....	170
CHAPITRE III. PRÉDICTION DES PERFORMANCES EN COURSE DE DEMI-FOND ET DE FOND VIA UNE MODÉLISATION LINÉAIRE DISTANCE-TEMPS, UNE LOI PUISSANCE ET UN NOMOGRAMME.....		172
1.	Introduction.....	172
2.	Méthodes.....	174
3.	Résultats.....	176
4.	Discussion.....	181
5.	Conclusion.....	184
CHAPITRE IV. PRÉDICTION DES PERFORMANCES SUR 3000 M STEEPLE À L'AIDE D'UNE SIMPLE ÉQUATION.....		185
1.	Introduction.....	185

Prédiction de la performance en course à pied : demi-fond et fond

2. Méthodes.....	187
3. Résultats.....	189
4. Discussion.....	192
5. Conclusion.....	195
CHAPITRE V. PRÉDICTION DES PERFORMANCES SUR MARATHON À L'AIDE DE L'INTELLIGENCE ARTIFICIELLE (IA).....	196
1. Introduction.....	196
2. Méthodes.....	198
3. Résultats.....	202
4. Discussion.....	206
5. Conclusion.....	210
CHAPITRE VI. PRÉDICTION DU POTENTIEL DES COUREURS DE DEMI-FOND VIA L'ÉVALUATION DU TAUX D'AMÉLIORATION DES PERFORMANCES.....	211
1. Introduction.....	211
2. Méthodes.....	213
3. Résultats.....	217
4. Discussion.....	224
5. Conclusion.....	229
DISCUSSION GÉNÉRALE.....	231
CONCLUSION GÉNÉRALE.....	241
RÉFÉRENCES.....	242
ANNEXES.....	272

INTRODUCTION GÉNÉRALE

Les origines de la course à pied sont lointaines [1]. Cette dernière a notamment évolué vers un sport de compétition et de loisir tel que l'athlétisme moderne qui est sans conteste issu des sports pratiqués dans l'Antiquité [1]. L'athlétisme est aujourd'hui un sport universel pratiqué sur tous les continents et considéré comme le « sport-roi » des Jeux olympiques (JO) [2]. Le fondement de l'athlétisme repose sur la notion de performance, en perpétuelle évolution, où les athlètes sont constamment à la recherche de records [3, 4]. L'analyse et la prédiction des performances en course à pied peuvent présenter un véritable intérêt pour les fédérations sportives à la recherche d'exploits, mais aussi pour les athlètes et les entraîneurs en quête de victoires ou de records. Les records offrent la possibilité d'étalonner les performances à travers les âges, les générations, et suscitent des questionnements sur les conditions dans lesquelles les performances chronométrées, indispensables au cours de la formation des athlètes choisissant la pratique en compétition ou de haut niveau, sont réalisées [5-7]. Quel que soit le niveau de pratique, dès lors que l'action optimise le rapport entre les capacités physiques d'une personne et une tâche sportive, la notion de performance peut être évoquée [8, 9].

La performance sportive est multifactorielle et dépendante de nombreux facteurs physiologiques (e.g., débit maximal d'oxygène : $\dot{V}O_{2\max}$, capacité d'endurance aérobie), biomécaniques (e.g., amplitude et fréquence de la foulée), psychologiques (e.g., motivation, résistance psychologique), environnementaux (i.e., mode de vie, règlement) et technologiques (e.g., matériel, substances ergogéniques) [3, 5, 6, 9-20]. D'autres facteurs hors de contrôle de l'athlète (e.g., facteurs socioculturels et politiques, conditions climatiques) contribuent également à la réalisation de la performance sportive [3, 5, 6, 9, 10, 12, 18]. L'analyse de ces

multiples facteurs et la compréhension des circonstances dans lesquelles les performances sportives sont exécutées et évoluent peuvent permettre de mieux cerner la pratique de course à pied, d'expliquer l'hétérogénéité des performances des athlètes ou encore de discuter des éventuelles améliorations [3, 5, 9, 13, 21].

Au fil des décennies, de nombreux auteurs ont notamment cherché à examiner la progression des performances humaines, en comparant les sexes [10, 12, 22] et/ou les distances des épreuves de course à pied [3, 13], afin parfois de les modéliser et/ou de tenter de les prédire précisément. La prédiction de la performance en course à pied, et donc de la performance potentielle, peut-être utile pour prescrire des allures d'entraînement afin d'optimiser la performance en compétition, mais aussi détecter, par exemple, les futurs athlètes au fort potentiel [23-25]. Différentes théories et plusieurs outils modélisant la relation entre le temps et la distance (ou la vitesse) ont été développés et étudiés pour prédire avec précision les performances en course à pied, sur route ou sur piste. Les nomogrammes de Mercier et al. [26] ou de Vandewalle [27], le concept de vitesse critique [28, 29] et la loi puissance [30-32] en sont des exemples notables.

Néanmoins, les travaux menés et les outils développés pour la prédiction de la performance en course à pied ont, dans la plupart des cas, ciblé les athlètes de sexe masculin et de haut niveau (niveau élite, et souvent de classe mondiale) [23, 27, 31, 32]. En l'occurrence la compréhension et la prédiction des performances chez des athlètes, quel que soit le sexe, de niveau plus modeste en course à pied demeurent plus limitées. Ce champ de recherche apparaît alors riche d'opportunités visant à étendre la compréhension des performances en course à pied, sur piste ou sur route, pour tout athlète à la recherche d'une optimisation de performance. Les athlètes élités possèdent, dans un grand nombre de cas, des capacités physiques, physiologiques, psychologiques et/ou tactiques au-delà des normes [33-35], et

bénéficient souvent des techniques d'entraînement et technologies les plus avancées. L'analyse des performances, indépendamment du niveau et du sexe, pourrait permettre potentiellement de valider et de comparer les outils de prédiction existants afin d'aider les athlètes à optimiser leur potentiel de performance à l'entraînement et en compétition.

1. Présentation du dispositif de thèse

Ce travail de thèse a été mené dans le cadre d'une Convention Industrielle de Formation par la REcherche (CIFRE) dont le dispositif a permis à l'entreprise, Orthodynamica (OD), de bénéficier d'une aide financière afin de me recruter en tant que « doctorante-salariée ». L'objectif était de donner, à cette entreprise, un accès à des méthodes de recherche afin d'en assurer son bon développement à travers mes travaux de thèse. Le projet de recherche a été mené avec un laboratoire de recherche publique, le Centre d'Études des Transformations des Activités Physiques et Sportives (CETAPS, EA 3832), de l'Université de Rouen-Normandie (URN). Les thématiques de recherche du CETAPS s'articulent avec les domaines de spécialisations proposés au niveau des formations qualifiantes de l'Unité de Formation de Recherche en Sciences et Techniques des Activités Physiques et Sportives (*i.e.*, management du sport, éducation et motricité, entraînement et optimisation de la performance sportive, et activité physique adaptée et santé). Le développement des travaux et activités de ce laboratoire interagissent avec la formation des doctorants qui se réalise au sein de l'École Doctorale « Homme Société Risques et Territoires » (HSRT) (ED 556) de l'URN.

Cette CIFRE, financée par l'Association Nationale de Recherche et Technologie (ANRT), associe, par conséquent, trois partenaires : un doctorant, une entreprise et

un laboratoire de recherche qui assure l'encadrement de la thèse durant trois années (Figure 1).

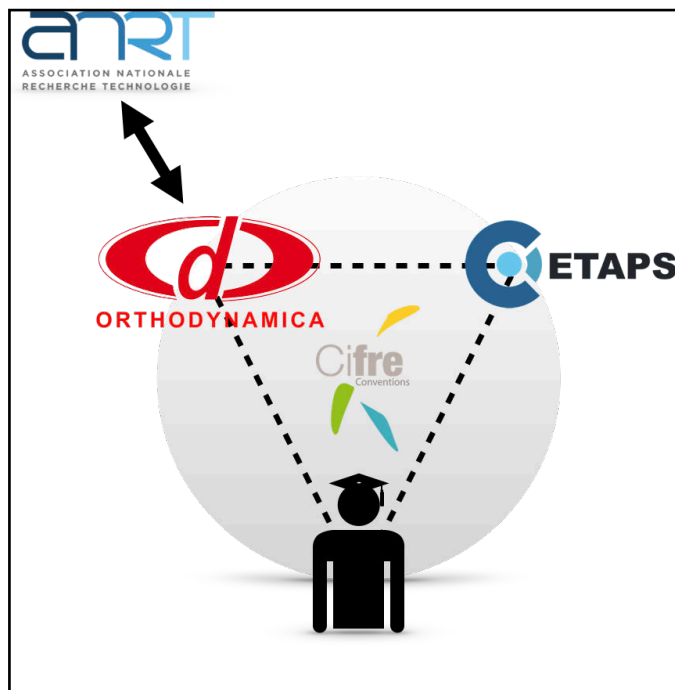


Figure 1. Dispositif CIFRE associant les trois partenaires

L'organisation de ce doctorat au sein de l'entreprise OD était définie de la façon suivante :

- 50% du temps destiné aux travaux de recherche.
- 50% du temps restant dédié aux missions propres de l'entreprise, dont la conception de produits « sur-mesure » (sandales orthopédiques) et la réalisation d'analyses du mouvement (marche et course principalement) dans le cadre des consultations podologiques des patients ou dans un cadre sportif.

Au-delà de ces principales missions, et en tant que membre du pôle Recherche et Développement (R&D) d'OD, j'ai également été assignée avec mes collègues chercheurs, à d'autres travaux de recherche et de communication sur des sujets divers en lien avec la podologie, les troubles musculo-squelettiques, la posturologie

et l'optimisation de la performance sportive (événements sportifs, colloques, séminaires, webinaires).

Ce travail de thèse a été réalisé sous la direction scientifique du Professeur Jérémie COQUART et co-encadré par le Docteur Maxime L'HERMETTE, tous deux rattachés initialement à l'URN et plus particulièrement à l'UFR STAPS, ainsi qu'au laboratoire CETAPS. Cependant, depuis septembre 2021, le Professeur Jérémie COQUART est rattaché à la Faculté des Sciences du Sport et de l'Éducation Physique de l'Université de Lille et à l'Unité de Recherche Pluridisciplinaire Sport, Santé, Société. Enfin, ce travail de thèse a également été tutoré par le Podologue Éric HELD au sein de l'entreprise OD.

2. Présentation de l'entreprise OD

L'entreprise OD est une société à exercice libéral à responsabilité limitée, qui a ouvert ses portes en novembre 2015, sous la direction et l'initiative de Monsieur HELD et de son équipe. Le centre mère de l'entreprise est installé au deuxième étage de la Clinique Mathilde 2 située à Rouen. C'est un centre pluridisciplinaire, alliant à la fois l'analyse du mouvement, la prise en charge orthopédique et podologique sur mesure au service de la santé et du sport. Il offre une continuité dans la prise en charge avec un matériel de haute qualité et une association de multiples pôles, dont un pôle de podologie et d'orthopédie générale (orthopédie sur-mesure, orthopédie en série, contention, prothèse mammaire), un pôle de communication, un pôle d'éducateurs médico-sportifs (EMS) et d'un pôle R&D (Figure 2). La chiropraxie, la naturopathie et le massage sportif sont également des disciplines complémentaires rattachées au centre OD (Figure 2).

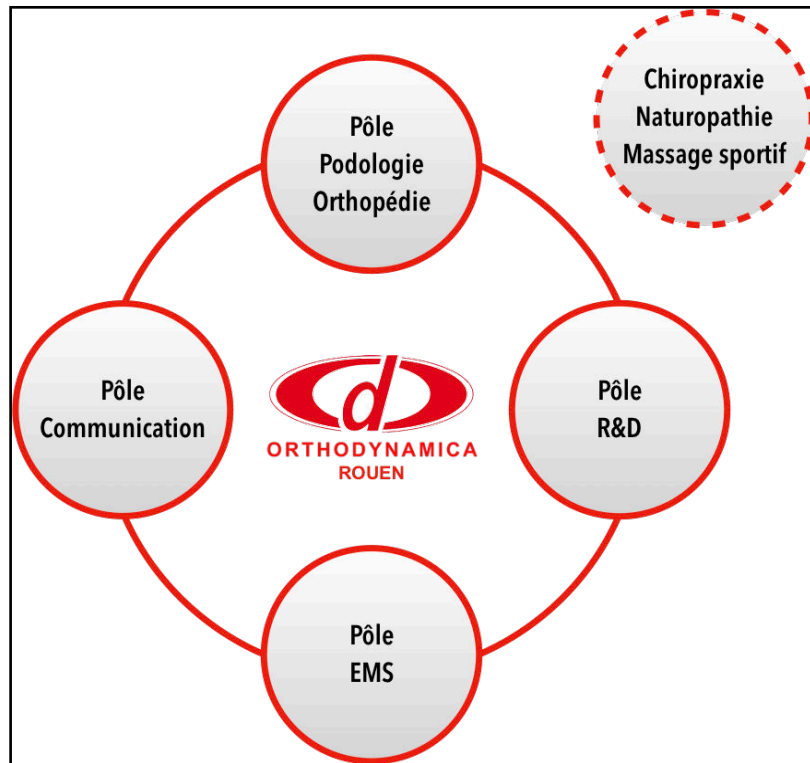


Figure 2. Centre « OD Rouen »

Les praticiens exercent leur activité dans le même bâtiment afin de pouvoir subvenir et répondre aux besoins des patients pouvant présenter des troubles de l'appareil musculo-squelettique ou podologiques en utilisant des moyens d'analyse performants. En effet, au-delà de la podologie, considérée comme une science médicale qui étudie le pied normal et pathologique et qui a pour visée de diagnostiquer et de traiter les différentes maladies et atteintes du pied en proposant notamment des soins ou des dispositifs médicaux podologiques externes [36] ; il semble essentiel de comprendre ces dernières et leurs répercussions sur la locomotion et sur le corps humain.

Si l'on prend quelques notions basiques de la podologie, le pied est le premier segment du corps qui entre en contact avec le sol et sa fonction de portance est essentielle [37]. De plus, le pied dirige la marche, supporte et propulse le corps [38]. Cependant, ce dernier peut être à l'origine de troubles morpho-statiques et dynamiques et peut ainsi engendrer des problèmes de santé [39]. En effet, la présence d'anomalies plantaires peut avoir de lourdes répercussions sur l'exécution

ou la répétition de mouvements, et par conséquent entraîner des douleurs ou des blessures, pouvant altérer directement la performance sportive par exemple [40-42]. Pour soigner une pathologie, le pied ne doit donc pas être considéré comme le segment le plus distal, mais bien le segment proximal [43, 44]. Pour cela, le podologue se doit alors de tenir compte de la statique et de la dynamique du pied tout en prenant en considération les interactions avec l'appareil locomoteur afin de proposer des solutions adaptées à chaque patient [43-45]. L'utilisation de topiques, de pansements ou encore d'orthèses plantaires [46] est un ensemble de solutions, non invasives et modulables, qui permet au podologue d'accomplir un acte thérapeutique sur diverses pathologies du pied, mais aussi du membre inférieur [36]. Si le podologue intervient généralement sur des pathologies d'origine mécanique comme l'hallux valgus [47] ou la talalgie [46], il peut également agir sur certaines pathologies inflammatoires telles que des œdèmes veineux ou lymphatiques [48]. Ces pathologies inflammatoires peuvent être enclenchées et accentuées, suite à un effort physique tel que la course à pied, et être responsables de l'apparition de troubles biomécaniques [49] ou hémodynamiques [50-52] au niveau du membre inférieur. La qualité et la vitesse de récupération du sportif peuvent, alors, se retrouver altérées [53].

Ainsi, selon le diagnostic thérapeutique établi [43-45], le podologue peut aboutir à la conception et à la réalisation de dispositifs spécifiques, comme la confection d'orthèses plantaires [46] ayant pour objectif de favoriser la récupération du sportif par exemple.

Par conséquent, la podologie et l'orthopédie générale sont, à l'heure actuelle, des disciplines paramédicales qui analysent le corps humain dans son ensemble, notamment *via* l'interface « Homme-Matériel », et qui offrent des perspectives et des objectifs de travail variés au centre OD au regard de leurs rôles préventif et curatif.

Depuis sa création, la société OD ne cesse d'accroître son activité avec l'ouverture de plusieurs centres en France, dont Évreux et Caen, mais également à l'étranger en s'installant à l'Île Maurice (Figure 3). La société OD a également pour projet de créer dans les années à venir de nouvelles antennes dans d'autres zones du territoire français, car elle a pour principal objectif de développer son expérience dans l'expertise de la performance motrice et l'efficience dans l'analyse thérapeutique. Le développement de la recherche dans le centre mère (*i.e.*, Rouen) a pour but de valider, ou non, les différentes hypothèses (*e.g.*, solutions thérapeutiques) et outils d'analyses (*e.g.*, système de capture de mouvements *Qualisys*, système de mesure des pressions plantaires *Zebris*, système d'évaluation posturale *DIERS...*) mis en place dans les domaines de la podologie et de l'orthopédie, afin d'optimiser la performance des sportifs et d'améliorer la prise en charge des patients. Au-delà de la podologie et au sein du centre d'analyse, OD propose d'autres services tels que « OD Run » et « OD Cycling » (Figure 3). L'espace « OD Cycling », dédié à l'amélioration de la position des cyclistes, permet de coupler le système de caméras trois dimensions *Qualisys* avec le système *bikefitting Shimano* afin de proposer des réglages optimaux et adaptés à la morphologie du sportif et à son niveau de pratique. L'espace « OD Run » est, quant à lui, dédié aux chaussures de course à pied ou de marche afin d'optimiser la pratique de ces activités en orientant les sportifs vers des chaussures adaptées à leur profil.

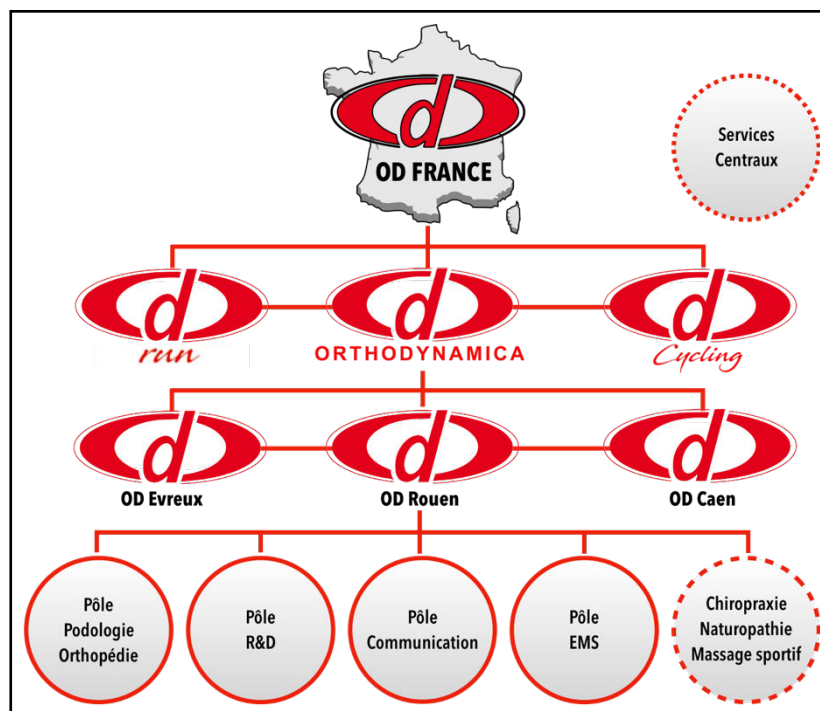


Figure 3. Organigramme « OD FRANCE »

La mise en avant de la recherche en collaborant avec l'URN est également l'une des priorités du centre en embauchant de jeunes chercheurs ou en proposant des thématiques de stages à des étudiants dans le cadre de leur formation universitaire. Désireux de développer de nouveaux projets de recherche par l'intermédiaire de plusieurs structures (e.g., universités, entreprises, cliniques), le Centre OD est à l'origine, avec le CETAPS et la société TRINOMA, de la création d'une entité scientifique « *Efficiency & Dynamics in Sport & Health* » (EDSH) composée de plusieurs jeunes chercheurs, d'enseignants-chercheurs, de chirurgiens et de professionnels médicaux et paramédicaux. À terme, l'objectif de cette entité est de mettre en place des recherches innovantes afin d'améliorer les connaissances dans les domaines du sport et de la santé.

La société OD a pour ambition d'apporter au domaine de la podologie et de l'orthopédie, mais aussi au domaine sportif, un regard scientifique par la mise en place d'expérimentations rigoureuses. L'analyse dynamique, *via* des systèmes

d'analyse quantifiée du mouvement et plus particulièrement des systèmes d'analyse tridimensionnelle [54-56], est le parti pris du centre OD afin de valoriser les solutions proposées aux sportifs et patients. Chez OD, la podologie moderne s'appuie notamment sur de nouvelles conceptions d'élaboration pour créer des produits innovants et sur-mesure tels que des semelles ou des sandales orthopédiques. Une semelle connectée est également à l'étude pour toujours plus de suivi du patient. La recherche et le développement des connaissances font donc partie des priorités de la structure afin d'optimiser l'efficacité et le savoir-faire de cette équipe pluridisciplinaire.

3. Présentation du plan et des objectifs de thèse

Comme évoqué précédemment, la société OD a pour objectif de développer de nouveaux projets afin de devenir de plus en plus performante dans sa proposition de services. De ce fait, l'entreprise n'hésite pas à faire intervenir et embaucher des doctorants dans sa structure. Après m'être spécialisée dans le domaine de l'interface « Homme-Matériel », et plus particulièrement dans la conception de produits sportifs, lors de mes années de *Master STAPS* « Ingénierie et Ergonomie de l'Activité Physique » (IEAP) à l'Université Savoie Mont-Blanc, il me semblait intéressant de poursuivre dans cette dynamique de R&D par l'intermédiaire d'une thèse CIFRE au sein de la structure OD et du laboratoire CETAPS. Ainsi, la réflexion et la recherche sur la motricité humaine dans un objectif de performance et de conception de produits ont été au cœur de ce projet de thèse CIFRE et ont notamment animé mes premières années de doctorat. En effet, les études menées lors de ce projet avaient pour objectifs initiaux de concevoir des sandales orthopédiques sur-mesure et de mesurer l'évolution de paramètres physiologiques, biomécaniques et perceptifs chez des sportifs afin de valider, ou non, les différents procédés de fabrication développés chez OD.

De ce fait, la revue systématique de littérature, réalisée lors des premiers mois de la thèse (*i.e.*, Juin-Octobre 2018) (Figure 4), était dédiée à faire l'état des connaissances actuelles sur l'impact du type de chaussage (*e.g.*, port de semelles, de tongs ou de chaussures) dit de « récupération » sur des paramètres physiologiques, biomécaniques et perceptifs en lien avec la récupération, la fatigue et la performance du sportif. Ce travail a notamment fait l'objet d'une publication scientifique internationale, portant sur les effets du type de chaussage sur le retour veineux au niveau des membres inférieurs durant l'exercice [57] (*cf.* Annexe 1), afin de permettre la mise en place de deux protocoles expérimentaux :

- Effet du port de semelles de récupération sur la fatigue perçue et la performance chez des golfeurs.
- Effet du port de sandales de récupération « sur-mesure » sur la fatigue et la performance chez des coureurs à pied.

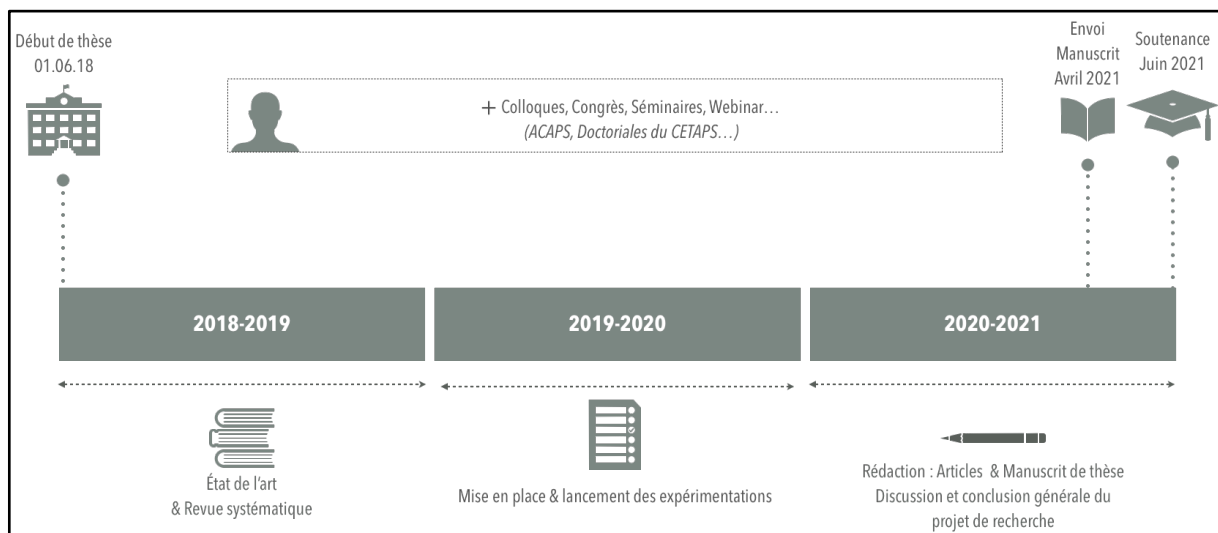


Figure 4. Organisation prévisionnelle des trois années de doctorat

Ces protocoles avaient pour enjeux de répondre aux objectifs suivants, dont examiner l'intérêt potentiel de l'utilisation de semelles (pendant l'effort) ou de

sandalettes de récupération « sur-mesure » (en post-effort) sur la fatigue et la performance chez le golfeur ou chez le coureur à pied, respectivement.

Toutefois, malgré l'accord favorable du Comité d'Éthique pour la Recherche en STAPS (CERSTAPS) pour le lancement des protocoles et le bon déroulement des pré-expérimentations avec les participants, ces deux études de terrain ont été reportées à deux reprises en raison de la crise sanitaire (*i.e.*, covid-19), avant d'être interrompues temporairement. Ces études seront donc menées prochainement, lorsque l'ensemble des conditions nécessaires à leur réalisation seront réunies.

Au regard des conditions sanitaires des dix-huit derniers mois de mon doctorat, afin de pallier l'annulation des études de terrain initialement prévues, et dans l'attente des conditions permettant la réalisation des protocoles, j'ai, par conséquent, été amenée à revoir l'organisation de ma thèse (Figure 5).



Figure 5. Aménagement de la thèse à la suite de la crise sanitaire (covid-19)

À gauche : Thématique de recherche développée lors de la première partie du doctorat (Juin 2018-Octobre 2020) et provisoirement suspendue (à la suite du deuxième confinement).

À droite : Thématique de recherche travaillée lors de la seconde partie de doctorat (Mars 2020-Novembre 2021).

Les multiples confinements m'ont notamment permis de travailler sur d'autres thématiques de recherche à partir d'études rétrospectives, non interventionnelles. Ces thématiques de recherche additionnelles ont donc été incluses, avec l'accord de l'entreprise OD et de mon directeur, dans mon travail doctoral. Ces travaux portent sur la prédiction de la performance en course à pied, et plus particulièrement en demi-fond et fond, et vous seront donc présentés dans cette thèse (Figure 6).

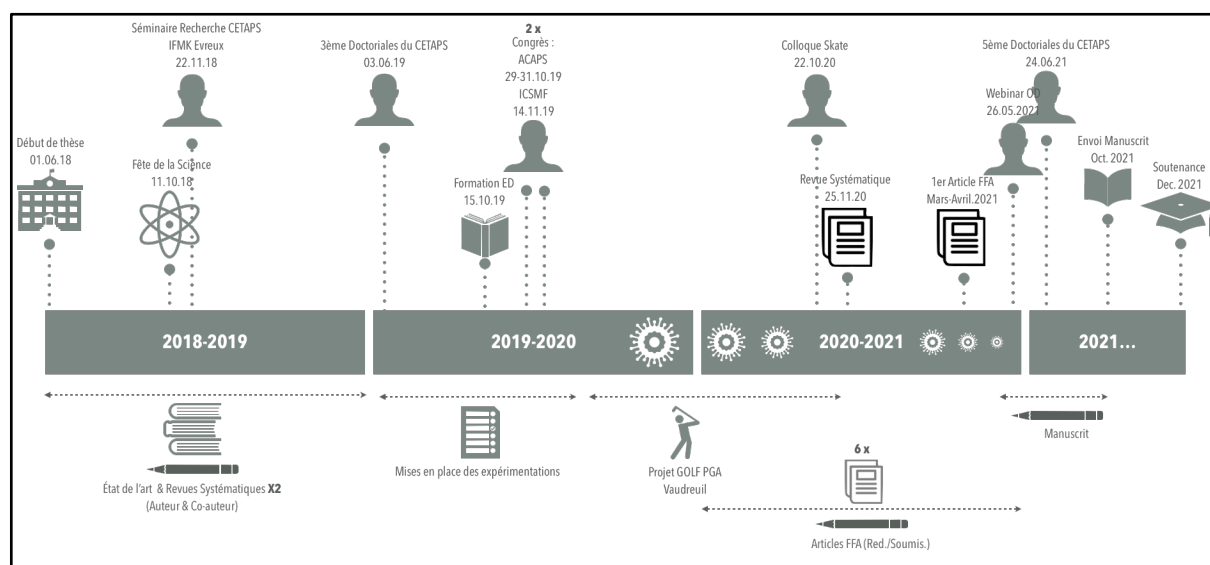


Figure 6. Planification de la thèse à la suite de la crise sanitaire (covid-19)

Suite à cette réorganisation, les différentes études menées ont eu pour objectifs d'analyser et de prédire les performances, ou le potentiel de performance, d'athlètes pratiquant la course à pied. Tout au long de cette deuxième partie de doctorat, des performances ont été analysées, différents outils tels que les nomogrammes, les concepts de vitesse critique (modèle linéaire) ou de loi

puissance, et l'intelligence artificielle (IA) ont été testés et comparés afin de vérifier leur validité et leur précision de prédiction. Un index de performance, construit sur l'analyse de multiples taux de progression d'athlètes figurant dans le top 10 national français, proposé et discuté en fin de manuscrit en guise d'application pratique, a également fait l'objet de l'une des études de ce travail de recherche. L'enjeu global de ces travaux était notamment d'identifier les outils pouvant contribuer théoriquement à optimiser les programmes d'entraînement des athlètes.

Le plan de thèse a, par conséquent, été revu de la façon suivante :

- Une première partie composée de trois chapitres ayant pour objectifs de présenter la pratique de la course à pied ; de définir la notion de performance en course à pied et les facteurs limitants ou propices à la réalisation de cette dernière ; et enfin un dernier chapitre axé sur la prédiction de la performance en course à pied et les principaux outils existants.
- Une seconde partie, organisée en six chapitres, destinée à présenter les résultats des recherches réalisées et contribuant à documenter et alimenter les connaissances des athlètes et entraîneurs de course à pied. Le premier chapitre de cette deuxième partie expose notamment l'évolution des performances en courses de demi-fond et de fond chez les femmes. Les chapitres II à V, portent sur la validation et la comparaison de méthodes de prédiction des performances des courses de demi-fond et de fond, sur piste et sur route. Enfin le chapitre VI met en avant l'analyse de taux d'amélioration des performances de demi-fond et de fond avant l'atteinte de records personnels. Un index de performance (en application pratique), est également présenté, dans le but d'évaluer la progression des athlètes, de prédire le potentiel et éventuellement de détecter les athlètes au fort potentiel.

- Une discussion et une conclusion générales permettent d'exposer les apports des différentes études présentées dans les chapitres I à VI (cf. deuxième partie), mais également de discuter des objectifs annoncés en lien avec cette thématique de recherche. Enfin, les limites et les perspectives de ce travail sont abordées en fin de ce manuscrit.

PREMIÈRE PARTIE : CADRE THÉORIQUE

La première partie de ce manuscrit a pour enjeu de définir le cadre théorique, dans lequel se sont déroulés les travaux de recherche de ce doctorat, recentré vers la prédiction de performance en demi-fond et fond. Il s'agit dans un premier temps de présenter l'activité « course à pied » (*i.e.*, contexte historique, classification des disciplines, pratiques masculines et féminines et état des lieux de la pratique en France) en ciblant notamment les disciplines constituant l'objet des études présentées dans la seconde partie de ce manuscrit. Ensuite, il sera question d'aborder la notion de performance, de records personnels (références chronométriques) et de définir les différents facteurs propices à la performance en course à pied tout en ciblant encore une fois les disciplines étudiées. Enfin, dans un dernier temps, et en lien avec les évolutions chronométriques évoquées (*e.g.*, records réalisés) et la notion de performance, seront présentés différents outils et méthodes développés pour prédire et optimiser la performance en course à pied.

CHAPITRE I. COURSE À PIED

Jadis, les hommes durent courir pour se nourrir ou se sauver des prédateurs qui se présentaient sur leur chemin. La course à pied est une locomotion relativement « simple » qui consiste à mettre un pied devant l'autre, sans que les deux pieds ne soient toujours en contact avec le sol (à la différence de la marche), pour générer un déplacement vers l'avant [58].

1. Contexte historique

Si l'origine de la course à pied est préhistorique, les premières disciplines athlétiques, essentiellement des épreuves de course, remontent quant à elles à l'Antiquité [1]. La course à pied a connu un grand succès chez les Grecs, avec les courses aux flambeaux (*i.e.*, lampadophories), et en demeurant l'une des disciplines reines lors des grands jeux helléniques [1]. Toutefois, la course à pied a évolué, d'un simple moyen de locomotion à une activité sportive de compétition et de loisir, et les épreuves antiques ont laissé place à l'athlétisme moderne, dont la naissance remonte à la fin du XIX^{ème} siècle, sous l'impulsion de la Grande-Bretagne [59, 2, 1]. Parallèlement à la rénovation des JO par Pierre de Coubertin, l'athlétisme moderne a pris véritablement racine au XIX^{ème} siècle, avec la création de clubs dans les collèges, la fondation des premiers clubs d'athlétisme amateurs dans différents pays (*e.g.*, *Amateur Athletic Club* en Angleterre en 1866, *New-York Athletic Club* aux États-Unis en 1868) et la mise en place des compétitions à différentes échelles (*e.g.*, nationales, internationales) (Figure 7) [59, 1, 60, 2]. La fédération internationale d'Athlétisme « *World Athletics* », anciennement « *International Association of Athletics Federations* » (IAAF), est fondée à la suite des JO de Stockholm en 1912, afin d'organiser les compétitions internationales mondiales.

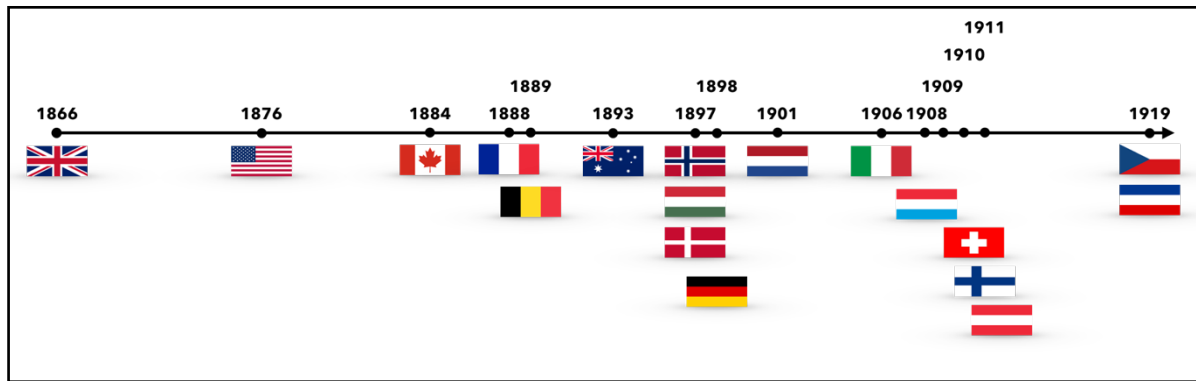


Figure 7. Mise en place des premiers championnats nationaux

En France, les premiers clubs ou expériences de groupes tels que le « Club des coureurs » fondé en 1875 ou le « Cercle de Madrid » créé en 1876, furent éphémères [1, 60]. Toutefois, deux clubs « Le Racing-Club de France » et « Le Stade Français », fondés quelques années plus tard (1882 et 1884, respectivement) réussirent à coexister et participeront à la création de l'Union des Sociétés Françaises de Courses à Pied (USFCP) en 1887, qui deviendra en 1889 l'Union des Sociétés Françaises des Sports Athlétiques (USFSA), et qui sera précurseur de la Fédération actuelle, la FFA, créée en 1920 (Figure 8) [1, 60, 2].

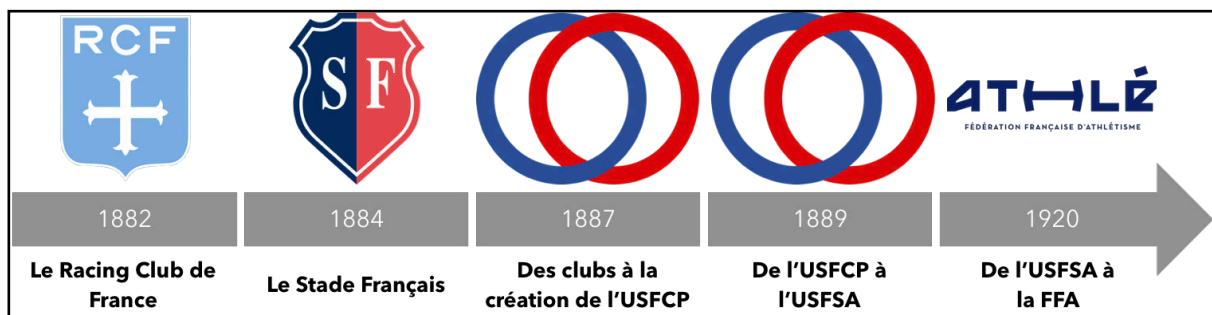


Figure 8. Création des premiers clubs et de la fédération (en France)

À l'échelle nationale, les premiers championnats de France d'athlétisme ont eu lieu en 1888 [60]. Toutefois, le nombre d'épreuves était limité et seules quelques courses masculines étaient au programme dont le 100 m, 400 m, 800 m, 1 500 m et 110 m haies [60].

À l'échelle internationale, il faudra attendre 1928 aux JO d'Amsterdam pour que le programme des courses athlétiques hommes, du 100 m au marathon, soit stabilisé [1].

Le programme des courses féminines s'est peu à peu aligné au fil des années malgré les craintes masculines concernant la capacité des femmes à pouvoir s'exprimer dans les différents registres d'efforts athlétiques [59, 1, 60]. Pierre de Coubertin, n'était absolument pas partisan des courses pour les femmes et confortait le fait que leur place était à la maison : « Elle est avant tout la compagne de l'homme, la future mère de famille, et doit être élevée en vue de cet avenir immuable ». Selon lui, les JO constituaient « l'exaltation solennelle et périodique de l'athlétisme mâle avec [...] l'applaudissement féminin pour récompense », jugeant que des « Olympiades avec des « femelles » seraient inintéressantes et inesthétiques » [1]. Sous l'impulsion d'Alice Milliat, grande militante du sport féminin pour la lutte de l'égalité, la Fédération Sportive Féminine Internationale (FSFI) est créée en 1921 [1, 2]. Toutefois, il aura fallu attendre plusieurs années avant d'atteindre l'alignement du programme féminin sur celui de leurs homologues masculins dans les compétitions nationales et internationales d'athlétisme [59, 1, 60, 61].

2. Définition et classification des disciplines : demi-fond et fond.

Les courses de demi-fond et de fond se courent majoritairement en peloton où les athlètes peuvent se protéger du vent et ainsi limiter la résistance de l'air, en courant à l'abri derrière l'athlète faisant la course en tête ou en se cachant dans la masse du peloton [60]. Les finalités de ces courses sont diverses et variées (e.g., battre un record, chercher la victoire) et requièrent l'acquisition d'automatismes tactiques et de capacités d'adaptation (*i.e.*, savoir se placer au départ, rester positionné au contact de la tête de course, imposer un train de course, connaître le meilleur moment pour relancer ou attaquer, se protéger du vent) [60]. Pour chaque course,

la confrontation entre les athlètes est directe (*i.e.*, lutte d' « homme à homme ») et le résultat est immédiat avec un départ commun, une même distance et une référence chronométrique comme juge de paix qui s'est singularisée au fil du temps sur chaque distance.

Le panel des distances de course sur piste a pour origine les fractions et les multiples du mile (1 609,32 m) soit la distance standard anglaise (encore disputée à l'heure actuelle) [59, 60, 1]. En compétition, et plus précisément dans le programme olympique, les épreuves du mile, 3 miles (*i.e.*, 4 828,03 m) et 6 miles (*i.e.*, 9 656,06 m) ont laissé place aux distances du 1 500 m, 5 000 m et 10 000 m, respectivement. La piste d'athlétisme de 400 m est par conséquent l'héritière du quart de mile (*i.e.*, 402,34 m). Toutefois, l'adoption du système métrique n'a pas empêché la coexistence des distances anglaises sur les tablettes des records [59, 60]. Aux distances olympiques, s'ajoutent des épreuves déclinées en fonction des catégories d'âge et de sexe pour les rendre accessibles aux jeunes (Tableau 1) [60, 62] ; ainsi que des épreuves faisant l'objet d'un record du monde, mais qui ne figurent pas au programme des championnats comme le 1 000 m, le mile, le 2 000 m, le 3 000 m, le 20 000 m, l'heure et les relais (*e.g.*, 4 × 800 m, 4 × 1 500 m) [2, 60]. Ces épreuves non olympiques sur piste font généralement l'objet de tentatives de records en *meetings*. En effet, il est possible de distinguer plusieurs types de compétitions sur piste dont les championnats (*e.g.*, régionaux, interrégionaux, nationaux, internationaux) où les athlètes tentent d'obtenir la meilleure place ou une éventuelle qualification au temps (*i.e.*, minima) pour accéder au tour ultérieur (*e.g.*, des championnats départementaux aux championnats régionaux) ; les compétitions collectives (*e.g.*, interclubs) où les athlètes tentent de réaliser la meilleure performance possible pour l'équipe dans un objectif de points ; et les *meetings* où les athlètes tentent de réaliser la meilleure performance chronométrique possible [60]. À la différence des courses de championnats, par exemple, organisées dans

un cadre fédéral, où la place est décisive (*i.e.*, classement), les courses de *meeting* sont des initiatives « privées » qui s'appuient sur des « partenaires » publics ou « sponsors » privés et qui prévoient généralement un programme réduit, pour des athlètes invités avec des primes aux participants et aux performances réalisées pour assurer le spectacle [2, 60]. Ces courses ont pour avantage d'être rapides. Le rythme de course est généralement emmené par des athlètes (*i.e.*, meneur d'allure ou « *pacers* ») chargés d'assurer une allure élevée ou conforme aux objectifs demandés par les meilleurs compétiteurs ou l'organisation [2, 60]. Par conséquent, bien que la nature des courses accorde une importance à la place et au classement, l'athlétisme moderne accorde une place considérable à la chasse aux records qui s'organise régulièrement dans le cadre de *meetings* ou de rencontres spécifiques.

Tableau 1. Exemple d'épreuves de courses de demi-fond et de fond sur piste proposées par la FFA en championnats en fonction des catégories et du sexe

Catégorie	Âge	Épreuve	
		Homme	Femme
Éveil athlétique	7-9 ans	Par équipes : Endurance 8', course d'endurance progressive (9-10)	
Poussin	10-11 ans	4' course, 1 000 m	
Benjamin	12-13 ans	1 000 m	
Minime	14-15 ans	1 000 m, 3 000 m	1 000 m, 2 000 m
Cadet	16-17 ans	800 m, 1 500 m, 3 000 m, 2 000 m <i>steeple</i> , 30 min	
<i>Junior</i>	18-19 ans	800 m, 1 500 m, 5 000 m, 10 000 m, 3 000 m <i>steeple</i>	800 m, 1 500 m, 3 000 m, 10 000 m, 2 000 m <i>steeple</i>
Espoir, <i>Senior</i> <i>Master</i>	20-22 ans	800 m, 1 500 m, 5 000 m, 10 000 m, 3 000 m <i>steeple</i>	
	23-34 ans		
	35 ans et +		

La classification des distances de course à pied, en demi-fond et fond, est une question récurrente qui semble purement arbitraire. Il est courant de classer les épreuves en fonction de la distance parcourue et du temps réalisé (ou puissance développée), ou en fonction de l'aspect physiologique [8, 59, 60].

Aubert et Choffin [60] proposent par exemple, une classification des distances de course en fonction du pourcentage de la vitesse maximale aérobie (VMA) sollicitée (Tableau 2). Ils distinguent les épreuves de :

- « Demi-fond court » telles que le 800 m ou le 1 000 m, se courant à une vitesse supérieure à la VMA et dans lesquelles la contribution de la voie glycolytique (*i.e.*, système anaérobie lactique) est prépondérante.
- « Demi-fond » entre le 1 500 m et le 3 000 m (dont le 3 000 m *steeple*), sollicitant un pourcentage de VMA légèrement supérieur ou inférieur à 100% (Tableau 2) et au cours desquelles le processus de phosphorylation oxydative à puissance maximale est le facteur essentiel de performance. Le caractère technique est également un déterminant de performance à ne pas négliger dans les courses de *steeple* (*i.e.*, courses avec franchissement de barrières et de rivières).
- « Demi-fond long » telles que le 5 000 m et le 10 000 m, nécessitant une puissance légèrement supérieure ou à hauteur du seuil anaérobie, et donc inférieure à la VMA.
- « Fond » telles que le semi-marathon (*i.e.*, 21,1 km) et le marathon (*i.e.*, 42,195 km), se courant à des allures légèrement inférieures au seuil anaérobie.

Tableau 2. Classification des distances de courses du 800 m au marathon en fonction de l'intensité de la VMA, d'après Aubert et Choffin [60]

Distance (m)	Intensité (%VMA)	Classification
800	115-120	Demi-fond court
1 500	105	Demi-fond
3 000	98-100	Demi-fond
5 000	95	Demi-fond long
10 000	88	Demi-fond long
Semi-marathon	83	Fond
Marathon	80	Fond

En d'autres termes, cette classification s'établit principalement en fonction du pourcentage de VMA sollicité lors de l'épreuve [59, 63-66, 60, 8, 67] (Tableau 3, Figure 9). Toutefois, il n'est pas toujours évident de classer les disciplines de course à pied, notamment celles, où la prépondérance d'une voie par rapport à l'autre n'est pas si évidente, et peut donc faire basculer la distance de course en demi-fond ou

en fond. Si certains auteurs ou fédérations définissent les épreuves de demi-fond (court ou long) comme étant les distances allant du 800 m au 5 000 m [68, 8, 69], ou les distances allant du 800 m au 10 000 m [60], d'autres distinguent le demi-fond du fond dès le 3 000 m en incluant cette distance parmi les courses de fond [59, 70].

Tableau 3. Exemple de contribution des systèmes énergétiques en fonction de la distance de course, d'après Dyer et Dwyer [59]

Distance (m)	Temps de course approximatif	Part des métabolismes d'énergie requis (%)	
		Anaérobie	Aérobie
100	11 s	95	5
400	50 s	75	25
800	2 min	50	50
10 000	30 min	15	85
Marathon	150 min	5	95

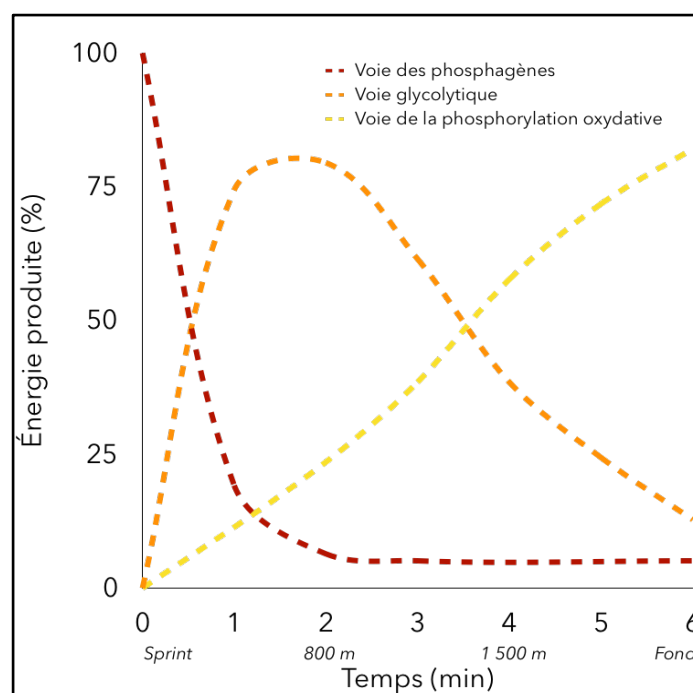


Figure 9. Relation entre la quantité totale d'énergie utilisée en fonction de la longueur et de la durée d'épreuve de course sur piste, d'après Howald [71]

Enfin, une classification des distances en fonction du type de terrain peut également être envisagée. En France, par exemple, les courses « hors-stade » (*i.e.*, sur route) allant du 10 km au 24 h sont associées aux courses de fond, alors que les courses

de plus courtes distances, du 800 m au 5 000 m, exécutées sur piste sont associées aux courses de demi-fond par la fédération [69].

Pour la suite de ce manuscrit, et en lien avec les travaux de recherche menés en collaboration avec la FFA, nous adopterons la classification de la FFA classant les épreuves allant du 800 m au 5 000 m et les épreuves d'une distance égale ou supérieure à 10 000 m, en demi-fond et en fond, respectivement.

3. Pratiques sportives masculine et féminine

a. Évolution vers la parité en athlétisme

Si de nos jours, le sport est reconnu comme jouant un rôle sociétal important pour promouvoir l'égalité des sexes [72], ce dernier n'a pas toujours été accessible aux athlètes féminines. À l'heure actuelle, le Comité International Olympique (CIO), encourage et soutient la promotion des femmes dans le sport, en développant le *leadership*, en menant des campagnes de sensibilisation, et en nommant davantage de femmes à des postes de direction au sein de l'administration ou à des postes de gouvernance [72]. Longtemps cantonnées à un rôle de mère par la société et écartées des pratiques sportives, les femmes ont pu progressivement accéder aux événements sportifs et s'exprimer à l'échelle nationale et internationale (e.g., championnats nationaux, JO...) sous l'impulsion d'Alice Milliat [1, 59, 60].

L'athlétisme est une discipline qui figure au programme des JO d'Athènes depuis la première édition en 1896 pour les hommes exclusivement et qui a été admise officiellement au programme féminin en 1928 lors des JO d'Amsterdam avec seulement cinq épreuves (100 m, 800 m, 4 × 100 m, hauteur et disque) contre 22 épreuves pour les hommes lors de cette édition [1]. À noter qu'auparavant eurent lieu des JO féminins « non officiels » en 1922 à Paris à l'initiative d'Alice Milliat et en 1926 à Göteborg [1]. Si l'athlétisme a donc toujours été représenté aux JO, ces

derniers organisés tous les quatre ans par le CIO avec l'appui de la Fédération Internationale ont cependant connu quelques interruptions (*i.e.*, 1916, 1940 et 1944) lors des périodes de guerres mondiales (1914-1918 et 1939-1945) [1].

Au fil du temps, le programme olympique d'athlétisme a évolué avec l'ajout, la modification ou la suppression d'épreuves de course, de saut et de lancer. Le programme féminin a vu peu à peu la parité l'emporter sur le programme masculin [1, 59, 60]. Actuellement, les programmes comprennent 24 épreuves, pour les hommes comme pour les femmes, dont :

- des courses : 100 m, 200 m, 400 m, 800 m, 1 500 m, 10 000 m, relais 4 × 100 m, relais 4 × 400 m, marathon, 100 m haies (pour les femmes) ou 110 m haies (pour les hommes), 400 m haies, 3 000 m *steeple*, 20 km marche et 50 km marche.
- des sauts : longueur, hauteur, perche, triple-saut.
- des lancers : poids, javelot, disque, marteau.
- et des épreuves combinées avec l'heptathlon ou le décathlon, pour les femmes ou les hommes, respectivement.

Les hommes et les femmes peuvent désormais s'exprimer dans tous les registres d'efforts athlétiques.

Toutefois, au regard des dates d'entrée des différentes épreuves qui définissent ces programmes d'athlétisme [1, 60], il demeure intéressant de prendre en considération les dates d'ajout de ces dernières à l'échelle internationale (*e.g.*, olympique, Figure 10) ou nationale (*e.g.*, en France, Figure 11) en fonction des sexes. Les disciplines de course à pied ayant fait l'objet de ce travail de recherche ont particulièrement retenu notre attention dans le but de pouvoir témoigner de l'évolution des représentations des capacités athlétiques (*i.e.*, performances ou records) des coureurs, hommes et femmes [10, 73], en reflet de changements de société. En effet, la pratique sportive compétitive sous sa forme actuelle s'est notamment développée parallèlement à la progression des nations durant ces deux

derniers siècles (e.g., influences géopolitique, culturelle et sociale, et progression économique) [74]. Le sport, initialement amateur, a progressivement laissé place au sport professionnel lui donnant ainsi une nouvelle dimension (i.e., rémunération, médiatisation...) [1, 59, 75, 76].

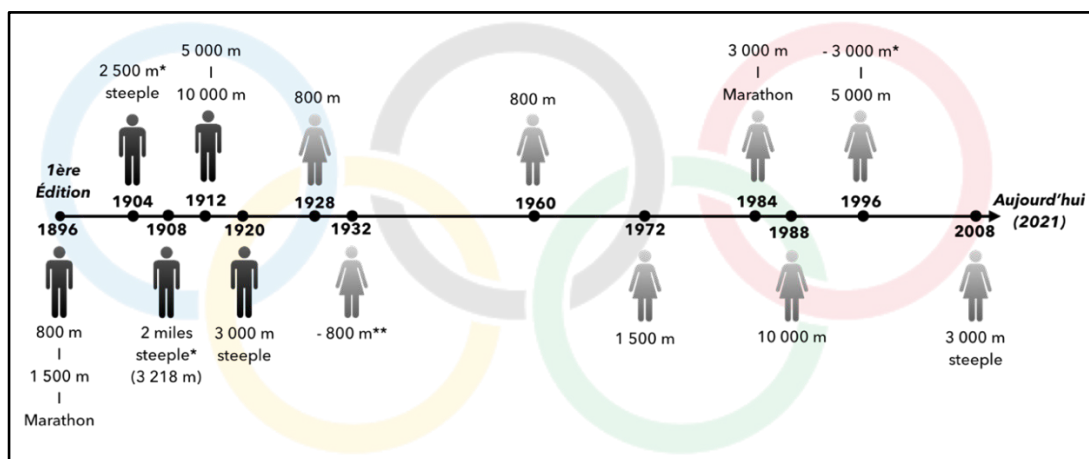


Figure 10. Évolution chronologique des dates d'entrée des courses de demi-fond et de fond dans le programme olympique, depuis la première édition de 1896 à aujourd'hui, chez les hommes et les femmes

Note : *Distance ayant été modifiée ou remplacée. **Distance ayant été supprimée.

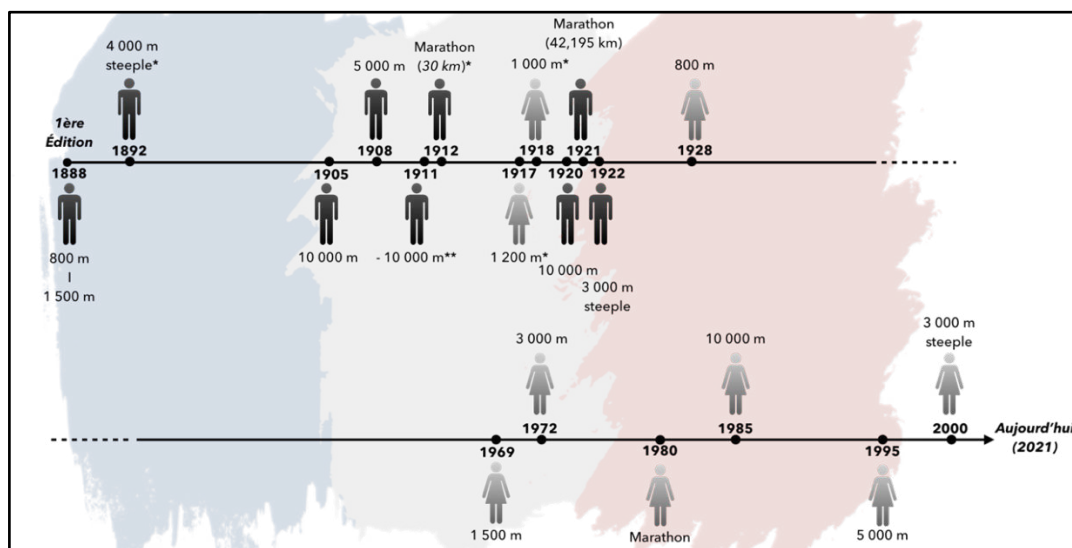


Figure 11. Évolution chronologique des dates d'entrée aux championnats de France d'athlétisme des courses de demi-fond et de fond, depuis la première édition de 1888 à aujourd'hui, chez les hommes et les femmes

Note : *Distance ayant été modifiée ou remplacée. **Distance ayant été supprimée.

b. Dates d'entrée des épreuves de demi-fond et de fond dans les programmes des compétitions : à l'échelle nationale et olympique.

Comme évoqué précédemment, l'inscription des différentes épreuves officielles de demi-fond et de fond dans les programmes nationaux français et olympiques ont eu lieu progressivement dès 1888 et 1896, respectivement, pour les hommes et à partir de 1928 pour les femmes (Figures 10 et 11) [1, 59, 60]. Avant d'être retenue au programme des JO, chacune de ces épreuves est apparue lors de championnats nationaux, continentaux ou mondiaux. Certaines épreuves préalablement inscrites dans les programmes se sont vues, au fil des années, retirées puis réintroduites (e.g., 800 m féminin aux JO), ou modifiées (e.g., 4 000 m *steeple* en France) afin d'uniformiser le règlement [77], d'homogénéiser les distances de course à pied entre les nations et tendre vers une égalité entre les sexes (*i.e.*, parité des distances disputées) [1, 60]. Les dates d'entrée des différentes disciplines de course à pied, faisant l'objet de ce travail de recherche (*cf.* cadre expérimental), ont été détaillées ci-après. À noter que les caractéristiques physiologiques de chaque épreuve, présentée succinctement dans ce premier chapitre, feront l'objet de plus amples détails dans le chapitre suivant de ce manuscrit (*cf.* première partie, chapitre II).

i. 800 m (demi-fond)

À l'échelle olympique, l'épreuve de 800 m figurait au programme masculin de la première édition des jeux en 1896 à Athènes et au programme féminin en 1928 à Amsterdam, lors de la première édition des jeux autorisés pour les femmes [1] (Figure 10). Le 800 m, caractérisé par les 2 tours de piste et qui requiert à la fois une grande puissance aérobie et anaérobie, est une discipline tactique où la voie glycolytique prédomine et où les fautes de placement dans le peloton ou la mauvaise gestion d'allures peuvent être coûteuses [60, 63, 65]. Cette épreuve de

demi-fond a été retirée du programme olympique féminin dès 1929 par le CIO, car jugée trop difficile pour les femmes qui se sont allongées sur la piste lors de leur arrivée pour récupérer [1] : « Néophytes, elles répartissent mal leur effort, et donc s'époumonent. Après leur arrivée, elles ont le mauvais goût de se laisser tomber sur le gazon, ce qui alors ne se faisait pas. Le 800 m ne leur sera rendu qu'en 1960 » [78]. En conséquence, aucune course féminine supérieure à 200 m ne sera disputée dans le cadre des JO avant 1960 [1].

À l'échelle nationale, le 800 m a été inscrit au programme des premiers championnats de France d'athlétisme en 1888 pour les hommes et en 1928 pour les femmes [1] (Figure 11). Malgré la suppression de cette distance dans le programme olympique, le 800 m reste disputé lors des championnats nationaux, notamment en France.

ii. 1 500 m (demi-fond)

L'épreuve de 1 500 m figurait également au programme masculin de la première édition des JO en 1896 et il a été ajouté au programme féminin en 1972 aux JO de Munich [1] (Figure 10). Le 1 500 m, caractérisé par les 3 tours $\frac{3}{4}$ de piste, est une discipline de demi-fond exigeante sollicitant principalement la voie de phosphorylation oxydative, mais également fortement la voie glycolytique [60, 63, 66].

En France, le 1 500 m a été disputé dès les premiers championnats nationaux de 1888 pour les hommes, et est apparu 81 ans plus tard, en 1969 dans le programme féminin (Figure 11).

iii. 3 000 m *steeple* (demi-fond)

Les origines du 3 000 m *steeple* sont britanniques, comme pour toutes les courses d'obstacles en athlétisme, et issues des courses hippiques [1, 60]. Le *steeple* a fait son entrée dans le programme olympique masculin en 1900 à Paris, sur deux distances (*i.e.*, 2 500 m et 4 000 m). Toutefois, en 1904, seule l'épreuve de 2 500 m *steeple* était au programme, avant qu'elle ne soit modifiée en 1908 par le 3 218 m *steeple* (2 miles) aux JO de Londres, puis retirée en 1912 des JO de Stockholm et définitivement transposée en 3 000 m *steeple* en 1920 aux JO d'Anvers [1, 60] (Figure 10). Cette discipline n'est codifiée qu'en 1954 par l'IAAF [1]. Chez les femmes, cette discipline est la dernière des courses à avoir été inscrite au programme de l'athlétisme et a fait son entrée très tardivement en 2005 aux championnats du monde et en 2008 aux JO, établissant ainsi la parité entre les programmes homme-femme [60]. Le *steeple* est une discipline qui a pour particularité d'avoir cinq obstacles à franchir par tour de piste, dont l'un est une rivière (*i.e.*, fosse remplie d'eau) [79]. La réglementation de cette course est identique pour les deux sexes, hormis la hauteur des obstacles qui est adaptée et abaissée à 76 cm pour les femmes contre 91 cm pour les hommes [1, 60, 79]. Comme pour le 1 500 m et le 5 000 m, le *steeple* requiert un fort potentiel aérobie et des qualités techniques et physiques (*e.g.*, coordination, puissance musculaire, vitesse et équilibre) [14, 80]. L'économie de course est un paramètre essentiel afin de limiter les ruptures de rythme lors du franchissement des obstacles, par exemple [60, 81, 82].

À l'échelle nationale, le *steeple* a fait son apparition aux championnats de France dès 1892. La distance courue était alors de 4 000 m avant d'être remplacée officiellement par le 3 000 m *steeple* en 1922 dans le programme masculin (Figure

11). Chez les femmes, le 3 000 m *steeple* figure au programme national depuis l'an 2000 (Figure 11).

iv. 5 000 m (demi-fond)

L'épreuve de 5 000 m est inscrite aux programmes masculin et féminin des JO depuis 1912 et 1996, respectivement [1] (Figure 10). Cette course remplace alors le 3 000 m qui figurait au programme féminin des JO de 1984 à 1992 et qui était apparu pour la première fois sur la scène internationale en 1974 lors des championnats d'Europe. Chez les hommes, le 3 000 m n'a jamais été une discipline olympique, hormis lorsque ce dernier était couru en équipe lors des JO de 1912, 1920 et 1924. Le 5 000 m est une course de demi-fond [60, 69], et plus précisément demi-fond long, nécessitant 12 tours et ½ de piste. Le 5 000 m est une épreuve à dominante aérobie sollicitant principalement la voie de la phosphorylation oxydative et où la puissance aérobie et l'économie de course sont importantes [14, 60, 68, 83].

En France, le 5 000 m figure au programme masculin des championnats de France depuis 1908 et est apparu au programme féminin en 1995 (Figure 11).

v. 10 000 m (fond)

Comme pour le 5 000 m, le 10 000 m masculin a fait son entrée en 1912 aux JO de Stockholm et il a fallu attendre 1988 pour le voir apparaître dans le programme féminin lors des JO de Séoul [1] (Figure 10). Le 10 000 m est une épreuve de fond, de 25 tours de piste, à dominante aérobie et physiologiquement proche du 5 000 m. À la différence du 5 000 m, la sollicitation du $\dot{V}O_{2max}$ (i.e., $\% \dot{V}O_{2max}$) sera

légèrement moindre [60, 83, 84]. L'économie de course et la capacité d'endurance aérobie sont des facteurs à prendre en considération pour ce type d'effort [85].

Sur le plan national, chez les hommes, le 10 000 m a été disputé en 1905 aux championnats de France, puis supprimé en 1911 et remis au programme en 1920 (Figure 11). Chez les femmes, le 10 000 m n'apparaît qu'en 1985 au programme [60] (Figure 11).

vi. Marathon (fond)

À la différence des épreuves présentées précédemment, le marathon est une épreuve de 42,195 km qui se dispute sur route. La voie de la phosphorylation oxydative et la capacité d'endurance aérobie sont fondamentales lors de cet effort [60, 83, 86]. Le marathon est l'une des cinq épreuves qui figuraient au programme de la première édition des JO de 1896 (Figure 10). Toutefois, ce n'est qu'en 1921 que la distance devient standard en compétition (e.g., 40 km en 1896). Du côté féminin, il a fallu attendre 1984 pour voir apparaître cette discipline aux JO de Los Angeles [1, 60, 87] (Figure 10), à la suite de la pression exercée par le mouvement féministe américain sur l'IAAF notamment avec l'affaire Switzer. En effet, en 1967, Kathy Switzer est la première femme, officiellement engagée, ayant participé au marathon de Boston sans que le règlement explicite une participation exclusivement masculine [1]. À la suite de cet événement, elle a été exclue de la fédération d'athlétisme, mais cela permit notamment quelques années plus tard, en 1981, d'ouvrir le programme des courses de fond à l'échelle internationale.

En France, l'épreuve de marathon figure au programme des championnats nationaux masculins depuis 1912, et est apparue dans le programme féminin en 1980 (Figure 11).

4. État des lieux de la pratique en France : Le développement de la FFA

Au regard de la convention de recherche obtenue par mon directeur de thèse avec la FFA, cette sous-section essentiellement destinée à présenter cette fédération, permettra notamment de faire le lien avec les travaux de recherche présentés en seconde partie de ce manuscrit.

Fondée en 1920 (Figure 8), la FFA est une association loi de 1901 de 305 914 licenciés en 2020 (161 876 et 144 038, hommes et femmes, respectivement). Dans le cadre de la délégation accordée par le Ministère chargé des sports, elle a pour principales missions : de développer, de contrôler et d'organiser la pratique de l'athlétisme sous différentes formes dont le stade (*i.e.*, courses, sauts, lancers, épreuves combinées et marche athlétique) et le hors-stade (*e.g.*, cross-country, *running*, marche nordique, *trail-running*, course en montagne) (Figure 12) ; de défendre les intérêts moraux et matériels de l'athlétisme français ; d'assurer la représentation de ces disciplines sur le plan international (*e.g.*, championnats d'Europe, du Monde, JO) ; d'assurer l'accès de tous à la pratique des activités physiques et sportives (*e.g.*, catégories d'âge, capacités de chacun, handicap) ; de former les dirigeants, animateurs, formateurs et entraîneurs fédéraux ; et de délivrer les titres fédéraux [69]. L'athlétisme souvent présenté comme une activité de performance et individuelle s'est, cependant, largement développé ces vingt dernières années [88]. La FFA a connu des changements majeurs qui ont particulièrement renouvelé l'organisation fédérale, les compétences des équipes et les activités proposées aux licenciés [88]. Les directives fédérales étaient, par exemple, d'augmenter la popularité et le nombre de courses sur route (*i.e.*, développement du secteur « hors-stade ») afin de donner aux athlètes la possibilité de courir avec plus de densité [88-90] ; mais aussi de proposer aux plus jeunes catégories des évènements qui puissent leur faire découvrir l'ensemble des activités

athlétiques (e.g., *Kid Stadium* en 1998 ou encore *Planètes Athlé Jeunes* en 2008). De plus, la FFA s'est mise à adopter une politique d'absorption des pratiques « satellites » telles que le *trail* (e.g., *Trail Tour* en 2008), les courses en nature ou la marche nordique (en 2014) éloignant ainsi certaines pratiques de la logique d'entraînement et de compétition vers une volonté de sport santé et de loisir (Figure 12) [88].

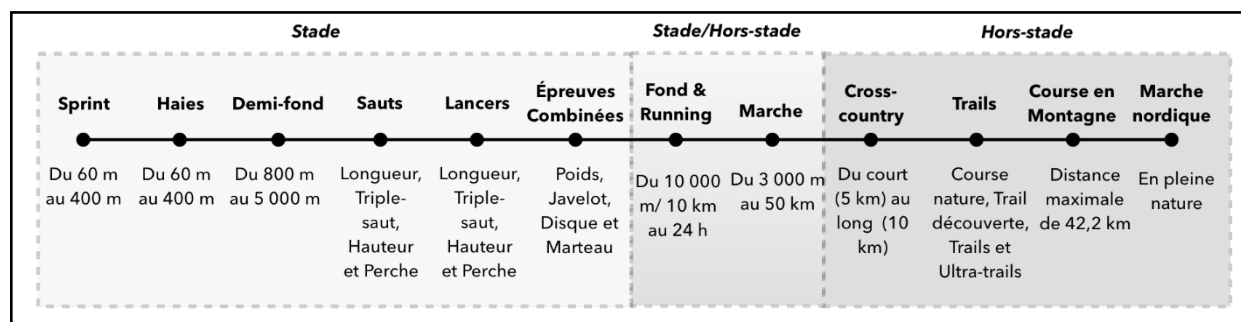


Figure 12. Disciplines de l'athlétisme sous différentes formes (i.e., stade et hors-stade)

a. Projet fédéral et plan de développement de la FFA

Depuis une vingtaine d'années, la FFA connaît une évolution sensible des formes de pratiques et du nombre de ses pratiquants [88]. En effet, l'ambition de la FFA est de pouvoir se développer tout en proposant des formes variées d'accompagnement et adaptées à son public. La conception et la mise en place de plans de développement (e.g., 2013-2017, 2018-2024), conçus à partir d'un projet fédéral [90], reposent notamment sur cinq grands secteurs d'activité au fort potentiel dont : l'éducation athlétique (i.e., jeunes de moins de 16 ans, catégorie à fort potentiel de recrutement), le stade (i.e., disciplines à finalité compétitive qui se déroulent sur un stade, qu'il soit de plein air ou couvert, avec pour but de favoriser l'accès à tout niveau de performance), le haut niveau (i.e., projet de performance fédérale), le *running* (i.e., pratique libre sans contrainte de distance ou de temps, ou pratique de manière structurée dans un club et lors de courses chronométrées) et

les pratiques de forme-santé (*i.e.*, pratique axée principalement sur la forme et la condition physique) [89, 91]. La FFA s'organise en fonction de l'émergence de nouveaux métiers avec un budget de 22,6 M€ en 2015 contre 17,7 M€ en 2012, par exemple, et de la mise en œuvre de nouveaux moyens pour répondre à des objectifs précis (*e.g.*, développement territorial et restructuration des clubs d'athlétisme, rénovation du site internet de la FFA, communications pour promouvoir l'athlétisme avec une image plus attractive) [88]. L'objectif étant de s'adapter à l'évolution de la société et aux attentes des pratiquants en offrant une diversification de l'offre tout en étant bien évidemment performant [89-91].

b. Évolution des pratiques : les différentes licences et catégories

En adhérant à la FFA, le licencié peut avoir accès à différentes formes de pratique (*e.g.*, découverte, compétition, *running*, santé) en fonction de sa catégorie d'âge (Tableau 1) et de ses besoins. À l'heure actuelle, six licences « athlé » et dix catégories allant de *baby* (moins de 6 ans) à *master* (plus de 35 ans) sont proposées aux adhérents de la FFA (Figure 13, Tableau 4). Ces catégories peuvent être regroupées en deux sous-catégories avec les jeunes « non spécialistes » (*i.e.*, découverte de l'ensemble des activités athlétiques) et les jeunes-adultes « spécialistes » (*i.e.*, orientés vers une famille athlétique puis spécialisés dans une discipline). Les jeunes athlètes, encore non spécialistes (*i.e.*, moins de 16 ans), sont l'une des priorités de la FFA. En effet, faire en sorte que les enfants et adolescents puissent découvrir les activités athlétiques (*e.g.*, sauts, courses et de lancers) et développer des compétences athlétiques en privilégiant le plaisir afin qu'ils puissent continuer à pratiquer l'athlétisme après l'âge de 16 ans, en se spécialisant selon leurs envies et leurs qualités, est un des leviers de développement et de performance de la FFA [88, 89, 91]. Piasenta, reconnu comme un des grands

entraîneurs français, considère notamment la catégorie minimale (*i.e.*, 14-15 ans) comme l'âge d'or [92].

Dans le cadre de nos études, nous nous sommes principalement intéressés à la catégorie « *senior* » (*i.e.*, 23-34 ans). Bien que la pratique puisse être pour le loisir, cette catégorie reste l'une des plus compétitives et concerne majoritairement un public désireux de pratiquer en compétition. Les athlètes *seniors*, en théorie plus matures physiologiquement que les catégories jeunes, pour les courses de demi-fond et de fond notamment [13, 93-95], se préparent aux échéances de compétitions selon des calendriers dédiés afin de pouvoir se dépasser et/ou performer.

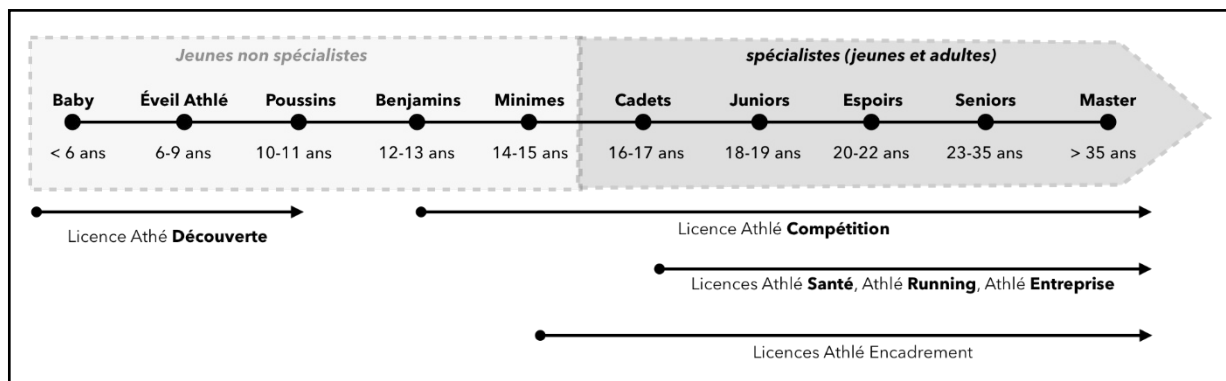


Figure 13. Licences et secteurs d'activités en fonction des catégories d'âge

Tableau 4. Détails des licences et secteurs d'activités en fonction des catégories d'âge proposées par la FFA [96]

Licences	Secteurs d'activités	Catégories
Découverte - École d'athlétisme	Découverte des habiletés motrices	Baby athlé
	Jeux de motricité Découverte des différentes disciplines	Éveil athlétique
	Approche ludique multidisciplinaire	Poussins
Compétition - Apprentissage des techniques d'Athlé	Découverte des disciplines	Benjamins
	Approche multidisciplinaire et collective des compétitions, développement des qualités physiques utiles aux pré-ados.	Minimes
Compétition - Performance	Entraînement, début de spécialisation	À partir de cadets (CA-JU-ES-SE-M0 à M10)
	Compétitions individuelles et collectives	
	Entraînement, spécialisation	À partir de <i>juniors</i> (CA-JU-ES-SE-M0 à M10)
Santé	Activités encadrées par un entraîneur diplômé ou Coach Athlé Santé, adaptées à un public qui recherche une pratique tournée vers le bien-être, la détente, la lutte contre la sédentarité et ses effets (marche nordique, <i>running</i> entretien, remise en forme, condition physique)	Toutes catégories à partir de cadet (CA-JU-ES-SE-M0 à M10)
	Running	Activités (footing, étirements, condition physique, fractionnés) visant à faire débiter ou reprendre la course à pied à des personnes à la recherche de la convivialité et de l'émulation d'un groupe et de conseils d'un entraîneur, ou à faire progresser des coureurs préparant des courses (hors championnats)
Encadrement - Dirigeants, Entraîneurs, Officiels, Spécialistes	Elle permet d'exercer des responsabilités officielles au sein d'un club, et/ou de postuler un poste de responsabilité départemental, régional ou national.	À partir de minimes (MI-CA-JU-ES-SE-M0 à M10)
Entreprise - Clubs rattachés à une Entreprise	Elle permet à son titulaire de participer à toutes les compétitions autorisées et aux Championnats de Sport en Entreprise	Cadets à <i>masters</i> (CA-JU-ES-SE-M0 à M10)

Note : Baby athlé (BA), éveil athlétique (EA), poussin (PO), benjamin (BE), minime (MI), cadet (CA), *junior* (JU), espoir (ES), *senior* (SE), *master* (M).

c. Évolution du nombre de pratiquants : une croissance importante

Alors que la FFA ne comptait que 40 000 licenciés en 1960, puis 160 000 en 1999, l'un des premiers constats majeurs est qu'elle dépassait les 300 000 licenciés en 2016 (Figure 14) [88, 96]. Si l'on s'intéresse, plus particulièrement aux vingt dernières années (en lien avec la convention FFA obtenue par mon directeur de thèse pour réaliser ce projet de recherche), la FFA a vu son nombre de licenciés augmenter de façon relativement importante [88], malgré le fait qu'elle ait parfois rencontré une période de stagnation comme au début des années 2000. En effet, en 2000, la FFA semblait traverser une période de crise. Son image n'était pas « au beau fixe », les résultats sportifs n'étaient pas non plus ceux espérés [88] et en 2000 le nombre de licences n'augmentait plus. Toutefois, dès 2008, le nombre de licenciés est reparti à la hausse et ne cessera de croître jusqu'en 2019 (Figure 14) [96]. En 2020, la FFA a dû de nouveau faire face à une période de crise, cette fois-ci d'ordre sanitaire, qui a impacté inévitablement le monde du sport avec un accès rendu difficile aux installations (sauf dérogation) et l'arrêt de la plupart des compétitions. La FFA a constaté, en conséquence, une baisse du nombre de ses licenciés.

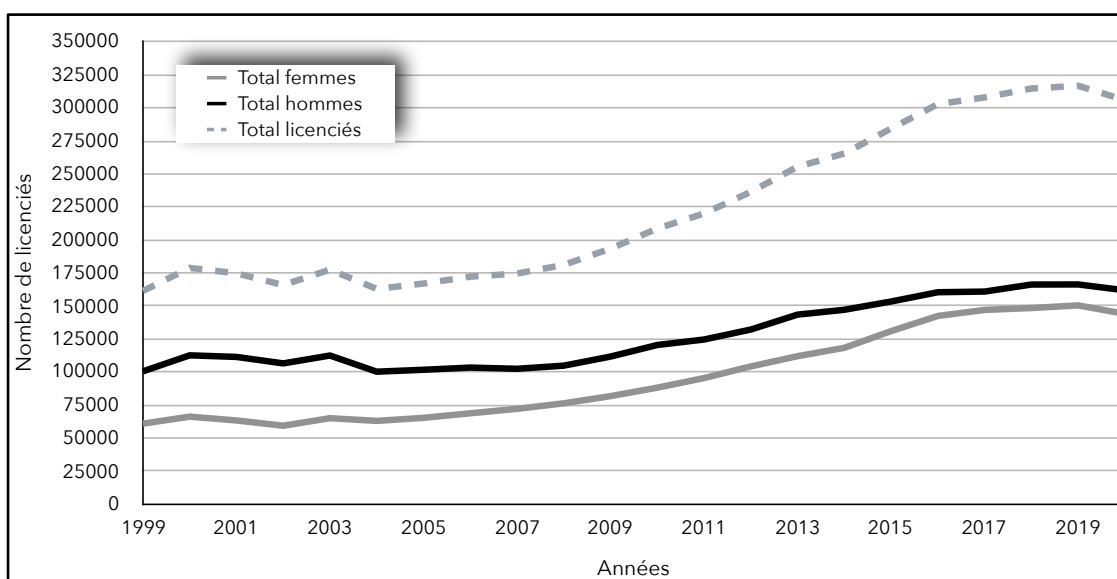


Figure 14. Évolution des licenciés FFA de 1999 à 2020 [96]

Il est également possible de constater que la part des femmes, parmi les licenciés, est devenue de plus en plus importante au fil des années (Figure 14) [96]. Elles représentent désormais près de 47% des effectifs. Il y a donc une véritable « féminisation » de l'effectif. De plus, à la suite de chaque cycle olympique, la FFA a pu constater une forte croissance des femmes *seniors* avec 7 622 licenciées en 2004, 9 965 en 2008 (*i.e.*, +30,7%), 12 969 en 2012 (*i.e.*, +30,1%) et 17 932 en 2016 (*i.e.*, +38,3%) [97].

d. Évolution dans une optique de performance et d'efficience

Toutefois, la FFA ne s'est pas développée seulement en nombre de licenciés, elle a progressé également dans d'autres secteurs (*e.g.*, haut niveau). Le développement se doit d'être une fonction transversale [89]. En effet, si la FFA propose une offre adaptée aux licenciés de niveau intermédiaire qui pratiquent pourtant dans le cadre compétitif (*e.g.*, les coureurs sur route, mais aussi les athlètes de niveaux départementaux et régionaux), elle veut également renforcer et pérenniser les acquis des olympiades précédentes en amenant notamment le plus grand nombre d'athlètes en finale des grands championnats [88, 89, 91].

En 2007, elle crée, par exemple, la ligue professionnelle (*i.e.*, la ligue nationale d'athlétisme ; LNA) afin d'accompagner ses meilleurs athlètes en leur apportant notamment un statut professionnel et une protection sociale dans le cadre de leur pratique [88].

La stratégie de performance de la FFA repose notamment sur la mise en place de différentes actions (*e.g.*, suivi médical efficace, critères de sélection aux grands championnats, équilibre sportif et professionnel, mise en place de stages nationaux)

afin d'optimiser les paramètres de la haute performance dans le secteur du haut niveau [88, 89].

5. Conclusion

Ce premier chapitre avait pour objectifs de présenter l'activité de course à pied et de relater son cheminement de l'Antiquité jusqu'à aujourd'hui. Au regard des évolutions politiques, culturelles et sociales qui ont bouleversé cette activité, à l'échelle nationale (française) et internationale, et au regard des études menées durant ce doctorat, il semblait intéressant d'exposer les principaux faits marquants. D'un simple moyen de locomotion à une activité sportive de compétition et de loisir, la course à pied a évolué en laissant place à l'athlétisme moderne. La fin du XIX^{ème} siècle fut notamment marquée par la création de clubs et la mise en place de compétitions nationales et internationales exclusivement réservées, dans un premier temps, à la gent masculine. Toutefois, sous l'impulsion d'Alice Milliat, les femmes ont progressivement réussi à accéder aux évènements sportifs et peuvent désormais s'exprimer dans tous les registres d'efforts athlétiques. Si les épreuves de course à pied, offrant un panel d'efforts continus, ont toujours été à l'origine des programmes athlétiques ; ces derniers n'ont cessé d'évoluer au fil des années témoignant ainsi des changements de société et de l'évolution des représentations des capacités athlétiques des hommes et des femmes (*i.e.*, performances ou records). Quelle que soit la nature des courses, même si l'on privilégie une importance à la place et au classement, l'athlétisme moderne accorde une place considérable à la chasse aux records et aux performances chronométriques. Si la référence chronométrique permet de comparer les performances et d'évaluer les progrès sur des distances données, la notion de record permet quant à elle d'étalonner ces performances et progrès à travers les âges. Ainsi, au-delà du contexte historique et des facteurs socioculturels, abordés dans ce premier

chapitre, et au regard de l'aspect multifactoriel qui définit la performance sportive et plus particulièrement la performance en course à pied, il semble désormais important d'analyser l'ensemble des facteurs permettant d'appréhender cette dernière et ses évolutions.

Tel est l'objectif du chapitre suivant où seront abordés dans un premier temps la notion de performance en course à pied et ses évolutions, s'exprimant au travers de références chronométriques, puis dans un second temps les différents facteurs propices à la réalisation de cette performance tout en ciblant encore une fois les disciplines ayant fait l'objet de ce travail de recherche.

CHAPITRE II. PERFORMANCE EN COURSE À PIED

Le fondement de l'athlétisme repose sur la notion de performance. En demi-fond et fond, les athlètes s'opposent en confrontation directe ou indirecte sur la distance disputée dans le cadre de compétitions, considérées comme un objectif ultime dans la pratique. Au-delà de la place très prisée en championnats nationaux, internationaux, olympiques (e.g., gagner une course, accéder à un podium), la référence chronométrique prend une place considérable dans la pratique de la course à pied. Cette référence chronométrique est un moyen de comparaison et d'évaluation de la progression sur des distances données ; et lorsque cette référence chronométrique est associée au plus petit temps réalisé sur la distance donnée (*i.e.*, meilleure performance, record), cela permet notamment d'étalonner les différentes références et les différents progrès à travers le temps (*i.e.*, diminution du temps pour une même distance).

1. Performance en course à pied et évolution des performances

a. Définition

La performance sportive peut être définie comme étant une action motrice dont les règles sont fixées par l'institution sportive qui permet à l'athlète d'exprimer ses potentialités physiques et mentales [8]. En effet, quel que soit le niveau d'exécution de l'action, dès l'instant où le rapport entre les capacités physiques de l'athlète et la tâche sportive à accomplir est optimisé, la notion de performance peut être appréhendée. Les conditions et les normes selon lesquelles les athlètes vont se confronter et réaliser des performances sont dictées par le règlement [60]. Le règlement, qui est à la fois la base et le fruit de l'évolution d'une pratique sportive, permet de structurer la performance, de conditionner la victoire, mais aussi

d'adapter les conditions de pratique d'une épreuve ou discipline afin de les rendre accessibles aux différents pratiquants de course à pied (e.g., catégories d'âges, sexe, handicap) [60, 83, 59, 3].

En course à pied, la performance, structurée par le règlement, peut s'exprimer sous deux systèmes de confrontation « directe » ou « indirecte », c'est-à-dire au travers d'un chronomètre (*i.e.*, record, qualification, niveau de performance) ou *via* un classement (*i.e.*, titre, podium, finale), respectivement ; et dans deux cadres de compétitions coexistants (*i.e.*, *meetings* ou championnats) (Figure 15) [60, 90].

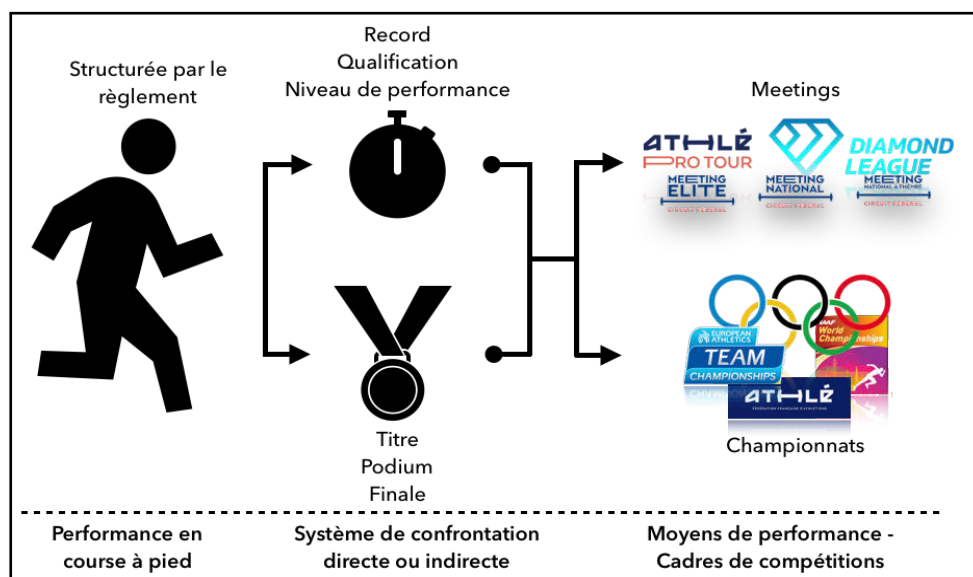


Figure 15 Performance en course à pied

Si, à l'heure actuelle, la performance peut être définie en fonction du temps nécessaire pour parcourir une distance (correspondant en l'occurrence à la vitesse moyenne de déplacement), il est important de souligner le fait que seul le classement avait de l'importance jusqu'aux années 1870.

En effet, le terme « record », emprunté à l'anglais (*i.e.*, *record*) et issu de l'ancien français (*i.e.*, se recorder : se souvenir), fait son apparition dans le dictionnaire de référence « *Oxford English Dictionary* » en 1880 où il est défini de la façon suivante : « *a record is a performance or occurrence remarkable among, or going*

beyond, others of same kind : especially, the best recorded achievement in any competitive sport ».

La comparaison des performances entre les athlètes et le recensement de records à l'échelle nationale, voire internationale était alors irréalisable auparavant et cela potentiellement en lien avec l'absence de législation, de norme ou de chronomètre fiable, ce qui par conséquent rend la notion de record relativement moderne dans le cadre de la course à pied. Le chronométrage, initialement manuel, a subi et connu de nombreuses expérimentations et évolutions technologiques afin de gagner en précision entre le XIX^{ème} et XX^{ème} siècle (e.g., chronométrage manuel en 1/5^e de s en 1895 ; 1/10^e de s en 1912) avant de devenir officiellement électrique au 1/100^e de s dans les années 1970 [1, 2, 60].

À l'échelle nationale, en France, les premiers records sont homologués à partir de 1892 par l'USFSA, mais les critères de réalisation de meilleure performance variaient et les listes des records successifs n'ont pas toujours été mises à jour avant 1905 [98]. À l'échelle mondiale, l'homologation des premiers records, soumis à un certain nombre de règles fédérales, évoluant avec le temps et influant la pratique, est apparue en 1912. L'IAAF a reconnu alors officiellement des records du monde d'athlétisme masculin. Toutefois, il aura fallu attendre 1936 avant de voir les premières diffusions de performances féminines par l'IAAF, régies auparavant par la FSFI.

Au-delà de la notion de record, toute référence chronométrique, obtenue dans le cadre d'une compétition officielle, permet aux concurrents de figurer dans des classements ou bilans, à différentes échelles (e.g., bilan national, européen mondiale ; en fonction du sexe, des catégories d'âge et/ou du handicap). A l'heure actuelle, il est possible d'expertiser des niveaux de performance (i.e., départemental, régional, interrégional, national, international) et de classer les

athlètes à l'aide de barèmes et tables de cotation mis en place par les fédérations [99, 100] (Figure 16). À chaque distance de course correspond une référence chronométrique définissant le niveau de performance de l'athlète (en nombre de points).

Epreuves	Cat.	IA 40 pt	IB 35 pt	N1 30 pt	N2 28 pt	N3 26 pt	N4 24 pt	IR1 21 pt	IR2 20 pt	IR3 19 pt	IR4 18 pt
800m	CJESV	1'59"84	2'01"84	2'06"00	2'10"00	2'13"00	2'17"00	2'20"00	2'22"50	2'25"00	2'28"00
1 000m	CJESV	2'35"70	2'38"50	2'42"00	2'46"00	2'50"00	2'54"00	2'58"00	3'01"00	3'04"00	3'08"00
1 500m	CJESV	4'06"00	4'11"00	4'21"00	4'28"00	4'35"00	4'41"00	4'47"00	4'55"00	5'00"00	5'05"00
3 000m	CJESV	8'53"00	9'05"00	9'20"00	9'35"00	9'50"00	10'05"00	10'20"00	10'40"00	10'50"00	11'00"00
5 000m - 5km	JESV	15'10"00	15'30"00	16'30"00	17'00"00	17'20"00	17'40"00	18'00"00	18'20"00	18'40"00	19'00"00
10 000m - 10km	CJESV	31'40	32'40	34'00	35'00	36'00	37'00	38'00	39'00	40'00	41'00
Marathon	ESV	2h29'00	2h32'00	2h40'00	2h48'00	2h56'00	3h04'00	3h12'00	3h18'00	3h24'00	3h32'00

Epreuves	Cat.	R1 15 pt	R2 14 pt	R3 13 pt	R4 12 pt	R5 11 pt	R6 10 pt	D1 8 pt	D2 7 pt	D3 6 pt	D4 5 pt	D5 4 pt	D6 3 pt	D7 2 pt
800m	CJESV	2'31"00	2'34"00	2'37"00	2'40"00	2'44"00	2'46"00	2'48"00	2'50"00	2'53"00	2'56"00	2'59"00	3'03"00	3'08"00
1 000m	CJESV	3'12"00	3'16"00	3'19"00	3'23"00	3'29"00	3'32"00	3'35"00	3'38"00	3'42"00	3'46"00	3'50"00	3'55"00	4'02"00
1 500m	CJESV	5'10"00	5'15"00	5'20"00	5'25"00	5'30"00	5'40"00	5'50"00	6'00"00	6'10"00	6'20"00	6'30"00	6'40"00	6'50"00
3 000m	CJESV	11'10"00	11'20"00	11'30"00	11'45"00	12'00"00	12'15"00	12'30"00	12'45"00	13'00"00	13'20"00	13'40"00	14'10"00	14'40"00
5 000m - 5km	JESV	19'20"00	19'40"00	20'00"00	20'20"00	20'40"00	21'00"00	21'30"00	22'00"00	22'45"00	23'30"00	24'00"00	25'00"00	26'00"00
10 000m - 10km	CJESV	42'00	43'00	44'00	45'00	46'00	47'00	48'00	49'00	50'00	52'00	54'00	56'00	1h00'00
Marathon	ESV	3h38'00	3h44'00	3h50'00	3h56'00	4h02'00	4h08'00	4h14'00	4h20'00	4h26'00	4h32'00	4h40'00	4h50'00	5h00'00

Figure 16. Exemple d'épreuves du barème Femme au 1er janvier 2010, d'après la FFA [99]

Note : D : départemental ; R : régional ; IR : interrégional ; N : national ; I : International ; CJESV : Cadet-Junior-Espoir-Sénior-Vétérans ; A-B et 1-4 : sous-catégories par niveau.

PS : À compter de novembre 2015, la catégorie « Vétérans » a laissé place à la catégorie « Master ».

Cependant, si la notion de nouveau record, synonyme d'amélioration et potentiellement de progrès, est l'un des objectifs majeurs pour les athlètes et entraîneurs, il n'en demeure pas moins que les notions de classement et/ou de niveau, qui en découlent, peuvent être toutes aussi importantes pour se qualifier en championnats. En effet, au-delà d'un certain niveau de performance requis et défini par les fédérations en lien avec le niveau de la compétition disputée (*i.e.*, minima qualificatifs), le nombre de places allouées en championnats ou dans certaines compétitions demeure relativement limité. Par conséquent, outre la référence chronométrique, la place au bilan ou le classement obtenu lors des courses de sélection font partie intégrante du mode de qualification, notamment lorsque le nombre de concurrents est trop élevé.

b. Modélisation et évolution des performances

Comme évoqué précédemment, pour l'ensemble des disciplines de course à pied, la performance, régie par le règlement fédéral international, peut être définie comme le temps mis sur une distance donnée, et peut par conséquent être exprimée en vitesse moyenne.

En 1906, Kennelly a publié une étude prospective abordant pour la première fois la forme de la relation vitesse-temps (ou intensité-durée) en s'appuyant sur des records du monde établis en course à pied et en natation [30]. Par la suite, ces travaux furent repris par Hill qui expliqua la forme de la relation vitesse-temps en course à pied et en natation également, pour les hommes et pour les femmes (Figure 17), en développant notamment de nombreux concepts tels que le $\dot{V}O_{2max}$ et la dette d'oxygène [101]. Il apporta notamment à l'analyse des records une approche physiologique, pionnière en la matière, soit l'aspect bioénergétique de l'exercice musculaire. La courbe, exprimant la relation vitesse-temps, exposée ci-après, démontre une diminution de la vitesse en fonction de l'allongement de la durée de l'exercice [101, 30, 102, 103].

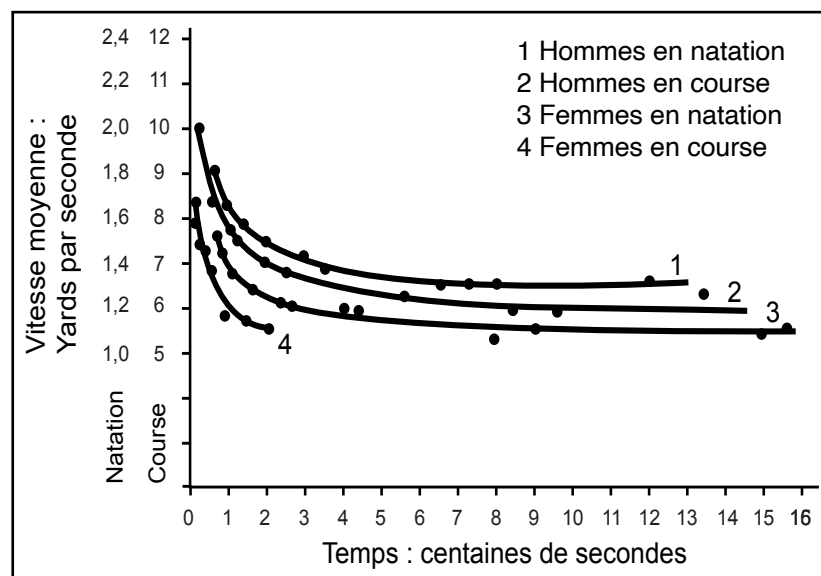


Figure 17. Relation vitesse-temps en course à pied et en natation [8, 101]

Pour établir cette courbe, il est possible de prendre, pour exemple, les records du monde actuels masculins de course à pied sur des distances allant du 100 m au 100 km (Figure 18) [104]. La forme de cette courbe permet notamment d'appréhender les facteurs limitatifs et les capacités physiques de l'homme. S'il est possible, de constater une vitesse sensiblement équivalente entre 10 et 20 s d'effort (*i.e.*, environ $10,4 \text{ m}\cdot\text{s}^{-1}$) correspondant respectivement aux épreuves de *sprint* de 100 m et 200 m ; il est également possible de constater des diminutions de vitesse importantes (*i.e.*, généralement entre $9,3$ et $7,9 \text{ m}\cdot\text{s}^{-1}$) entre des efforts compris entre environ 45 s et 1 min 40 s, soit deux temps de course relativement peu éloignés (*i.e.*, moins d'une minute de différence), correspondant respectivement aux épreuves de 400 m et de 800 m (Figure 18). Ces diminutions de vitesses sont liées à l'inertie et au délai de mise en route de chaque filière énergétique, car si ces trois filières interviennent lors de la production de l'effort, l'une d'elle prédomine [8, 59, 60, 105]. Par exemple, dans les épreuves de 100 m et de 200 m (*i.e.*, *sprint court*), courues à des intensités très élevées, la voie des phosphagènes (*i.e.*, métabolisme anaérobie alactique), contribue très fortement à la production d'énergie ; alors que dans les épreuves de 400 m (*i.e.*, *sprint long*) et de 800 m (*i.e.*, demi-fond court), les voies glycolytique et oxydative (*i.e.*, métabolisme anaérobie lactique et métabolisme aérobie) interviennent de façon non négligeable, dans la production d'énergie, expliquant ainsi une perte de vitesse en course à pied en lien avec les caractéristiques propres de puissance et de capacité de chacune des filières [65, 106] (Figure 9). En effet, pour les épreuves de 100 m, 200 m, 400 m et 800 m, les résultats de Duffield et al. [65, 106] indiquent une contribution relative du système énergétique aérobie-anaérobie basée sur les mesures de la concentration de lactate sanguin et de la dégradation estimée de la phosphocréatine de 9-91% et 11-89%, 21-79% et 22-78%, 35-65% et 37-63%, et 63-37% et 69-31%, pour les hommes et les femmes respectivement.

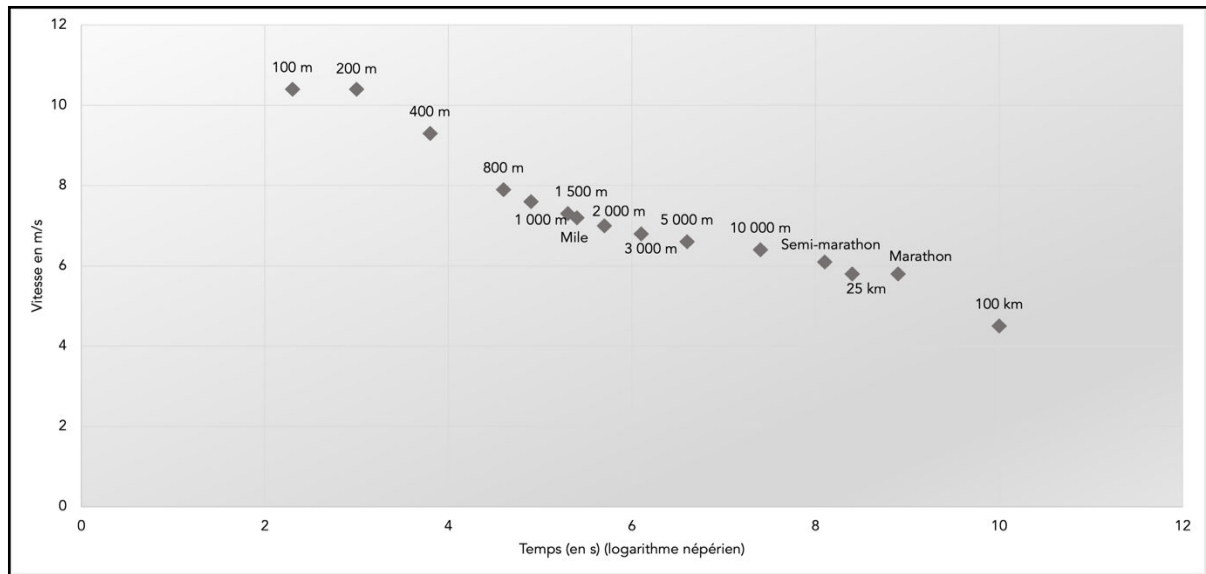


Figure 18. Exemple de la relation vitesse-temps des records du monde hommes actuels (au premier juillet 2021) en course à pied, allant des épreuves de 100 m au 100 km

La description, la comparaison, la prédiction, la recherche d'explications physiologiques ou socioculturelles sont autant d'arguments qui semblent avoir motivé un grand nombre de chercheurs à analyser les records du monde ou olympiques. Les analyses de ces records ont notamment permis d'élaborer, des modèles empiriques et théoriques de la relation entre la vitesse (ou la distance) et le temps afin d'appréhender la performance en course à pied et la capacité d'endurance aérobie [29, 30, 74, 101, 107-114].

La relation vitesse-temps (ou intensité-durée) en course à pied a fait l'objet de nombreuses investigations et plusieurs types de graphiques ont été utilisés pour analyser les meilleures références chronométriques à l'échelle mondiale voire olympique (*i.e.*, relations temps-distance, distance-temps, puissance/vitesse-distance, temps-date) (Figure 19) [102, 8]. La relation temps-date, associée à un évènement en particulier soit l'épreuve de course dans le cadre de nos recherches, illustre comment la référence chronométrique de cette épreuve a évolué dans le temps (Figure 19A). La relation temps-distance, montre pour l'ensemble des

épreuves étudiées, l'évolution de la relation entre le temps réalisé et la distance parcourue (Figure 19B). Pour la relation puissance-distance ou vitesse-distance, les vitesses (*i.e.*, distance/temps) ou puissances moyennes des différentes épreuves étudiées sont tracées en fonction de la distance de course (Figure 19C). Enfin, la dernière relation distance-temps illustre en fonction de l'allongement de la distance, une augmentation du temps de course (Figure 19D).

Par conséquent, selon les plages de temps, de distances, de puissance/vitesses ou de dates, étudiées en fonction des épreuves de course, plusieurs relations ont pu être modélisées sous forme parabolique (*i.e.*, $y = ax^n$), hyperbolique (*i.e.*, $y = \frac{a}{x}$), exponentielle (*i.e.*, $y = a[\exp (bx)] + c[\exp (dx)] + \dots$), correspondant chacune à des équations mathématiques pouvant permettre l'ajustement des différentes courbes présentées et ainsi décrire la courbe de la performance humaine en course à pied (Figure 19) [8, 102].

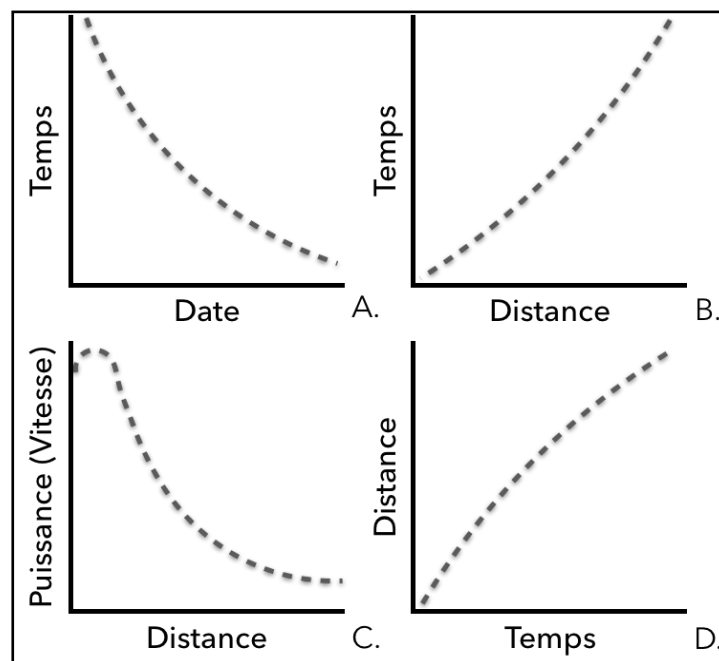


Figure 19. Illustration des différentes courbes utilisées couramment dans l'analyse des records (olympiques et du monde), d'après King et Black [102] et Billat [8]

Légende : Relations temps-date (A), temps-distance (B), puissance/vitesse-distance (C) et distance-temps (D).

À l'heure actuelle, si l'on compare les références chronométriques féminines et masculines à l'échelle mondiale, le record du monde féminin se situe entre 87,3 et 91,4% du record du monde masculin, excepté pour la distance d'ultra-fond de 100 km correspondant à un plus faible écart, avec une valeur de 93,9% du record des hommes (Tableau 5) [104].

Tableau 5. Records du monde actuels masculins et féminins en course à pied, du 100 m au 100 km (au 1^{er} juillet 2021)

Distance	Record masculin	Record féminin	Différentiel (en %) entre records masculin et féminin	Date du record actuel masculin	Date du record actuel féminin
100 m	9 s 58 ms	10 s 49 ms	9,9	2009	1988
200 m	19 s 19 ms	21 s 34 ms	9,6	2009	1988
400 m	43 s 3 ms	47 s 60 ms	8,6	2016	1985
800 m	1 min 40 s 91 ms	1 min 53 s 28 ms	11,4	2012	1983
1 000 m	2 min 11 s 96 ms	2 min 28 s 98 ms	11,5	1999	1996
1 500 m	3 min 26 s	3 min 50 s 7 ms	10,4	1998	2015
Mile	3 min 43 s 13 ms	4 min 12 s 33 ms	11,5	1999	2019
2 000 m	4 min 44 s 79 ms	5 min 23 s 75 ms	12,1	1999	2017
3 000 m	7 min 20 s 67 ms	8 min 6 s 11 ms	9,5	1996	1993
3 000 m steeple	7 min 53 s 63 ms	8 min 44 s 32 s	9,7	2004	2018
5 000 m	12 min 35 s 36 ms	14 min 6 s 62 ms	10,8	2020	2020
10 000 m	26 min 11 s	29 min 1 s 3 ms	9,8	2020	2021
5 km Route	12 min 51 s	14 min 43 s	12,7	2020	2021
10 km Route	26 min 24 s	29 min 43 s	11,2	2020	2017
Semi-Marathon	57 min 32 s	1 h 4 min 2 s	10,2	2020	2021
25 km Route	1 h 11 min 18 s	1 h 19 min 53 s	10,7	2012	2010
Marathon	2 h 1 min 39 s	2 h 14 min 4 s*	9,3	2018	2019
100 km Route	6 h 9 min 14 s	6 h 33 min 11 s	6,1	2018	2000

Note : *Performance réalisée en course mixte.

On peut également observer que le différentiel moyen entre les records du monde masculins et féminins est de l'ordre de $10,3 \pm 1,5\%$ et que les records réalisés par la gent féminine sont davantage datés (Tableau 5) [104]. En effet, pour l'ensemble des distances sur piste et sur route qui sont présentées dans le Tableau 5, les records du monde féminins ont été réalisés entre 1983 et 2021 et remontent en moyenne à

13,7 ± 14,5 ans (2007), alors que les records du monde masculins ont été réalisés entre 1996 et 2020 et datent en moyenne de 10,5 ± 9,1 ans (2011)(Tableau 5)[104].

A l'échelle nationale française (échelle peu étudiée en raison du plus modeste niveau des athlètes), le différentiel chronométrique, entre les records féminins et masculins, est légèrement supérieur au différentiel mondial. Le record de France féminin se situe entre 84,9 et 91,7% du record du monde masculin (Tableau 6)[104].

Tableau 6. Records de France actuels masculins et féminins en course à pied, du 100 m au 100 km (au 1^{er} juillet 2021)

Distance	Record masculin	Record féminin	Différentiel en % entre records masculin et féminin	Date du record actuel masculin	Date du record actuel féminin
100 m	9 s 86 ms	10 s 73 ms	9,8	2016	1998
200 m	19 s 80 ms	21 s 99 ms	9,6	2011	1993
400 m	44 s 46 ms	48 s 25 ms	8,3	2007	1996
800 m	1 min 42 s 53 ms	1 min 56 s 53 ms	12,1	2014	1995
1 000 m	2 min 13 s 96 ms	2 min 31 s 93 ms	11,9	2003	1995
1 500 m	3 min 28 s 98 ms	3 min 59 s 76 ms	13,0	2003	2010
Mile	3 min 50 s 33 ms	4 min 27 s 43 ms	13,9	1995	1996
2 000 m	4 min 53 s 12 ms	5 min 39 s	13,6	2005	1986
3 000 m	7 min 30 s 78 ms	8 min 35 s 41 ms	12,6	1998	2005
3 000 m steeple	8 min 00 s 09 ms	9 min 25 s 62 ms	15,1	2013	2009
5 000 m	12 min 58 s 83 ms	14 min 43 s 90 ms	11,9	2000	2004
10 000 m	27 min 17 s 29 ms	31 min 35 s 81 ms	13,6	2019	2012
5 km Route	13 min 18 s	15 min 31 s	14,3	2020	2020
10 km Route	27 min 13 s	31 min 15 s	12,9	2020	2019
Semi- Marathon	59 min 13 s	1 h 8 min 34 s	13,6	2019	2010
Marathon	2 h 6 min 36 s	2 h 24 min 22 s	12,3	2003	2010
100 km Route	6 h 23 min 15 s	7 h 26 min 44 s	14,2	2000	2007

Le différentiel moyen actuel entre les records de France masculins et féminins est de l'ordre de 12,5 ± 1,8% et les records féminins sont, comme pour les records du monde, plus vieux (Tableau 6). En effet, pour l'ensemble des distances présentées dans le tableau, les records de France féminins ont été réalisés entre 1986 et 2020 et remontent en moyenne à 17,2 ± 9,5 ans (2004), alors que les records du monde

masculins ont été réalisés entre 1995 et 2020 et datent en moyenne de $12,4 \pm 8,4$ ans (2009) (Tableau 6) [104].

Au-delà de ces constatations, il est également intéressant d'analyser et de comparer les évolutions de ces références chronométriques hommes et femmes, en course à pied. À l'échelle mondiale, l'évolution des performances en course à pied, sur différentes échelles de temps, a fait l'objet de plusieurs travaux et publications [8, 10, 13, 16, 59, 73, 74, 102, 111].

Si l'on s'intéresse aux épreuves de demi-fond et fond plus particulièrement, le constat est que les meilleures références chronométriques à l'échelle mondiale se sont considérablement et rapidement améliorées entre le XIX^{ème} et XX^{ème} siècle [59, 10, 16, 8, 102, 111, 115] (Tableaux 7 et 8). En effet, sur 100 ans (1882-1982), les analyses de Dyer et Dwyer [59], par exemple, ont indiqué des taux de progression des performances masculines de 12,3 à 25,2% pour des épreuves allant du 800 m au marathon (Tableau 7). Cette forte progression, possiblement due à la professionnalisation du sport et aux progrès technologiques et scientifiques, est cependant suivie d'un ralentissement de l'évolution dès la fin du XX^{ème} siècle [111]. Toutefois, malgré un ralentissement de l'évolution ces dernières années, la progression des performances à l'échelle mondiale reste notable [8, 10, 16, 74, 13, 111] (Tableaux 7 et 8). En effet, malgré des évolutions qui semblent dérisoires ces vingt dernières années, ces progressions demeurent importantes sur la ligne d'arrivée (Tableaux 7 et 8).

Tableau 7. Améliorations des records du monde masculins et féminins en demi-fond et fond lors de différentes périodes (de 1882 à 2017)

Améliorations (en %) des records du monde sur différentes périodes en fonction des distances						
Distance	Records du monde masculins (Dyer et Dwyer [59])	Records du monde masculins (Dyer et Dwyer [59])	Records du monde masculins (Dyer et Dwyer [59])	Records du monde masculins (Lippi et al. [115])	Records du monde féminins (Lippi et al. [115])	Records du monde masculins (Billat [8])
	1882-1982 (1952-1982)		1972-1982	1900-2007		1997-2017
800 m	13,7 (4,7)	2,4	4,5	-	-	0,2
1 500 m	25,2 (5,5)	0,8	3,8	12,4	10,3	0,7
5 000 m	12,3 (5,9)	1,6	4,4	-	-	3,0
10 000 m	13,2 (6,1)	1	8,0	15,1	8,5	0,006
Marathon	21,7	0,2	16,2	21,5	38,6	2,9

Tableau 8. Évolutions des records du monde masculins et féminins en demi-fond et fond, du premier record homologué à 2001 et à aujourd'hui (2021), et évolutions sur ces vingt dernières années (2001-2021)

Distance	Record du monde et date						Évolution du record du monde (%)					
	Hommes		Femmes		Hommes		Femmes		Hommes		Femmes	
	1 ^{er} record		Record en 2001		Record actuel en 2021		Du 1 ^{er} record à 2001 (1 ^{er} juillet 2021)		Du 1 ^{er} record à aujourd'hui (1 ^{er} juillet 2021)		De 2001 à aujourd'hui (1 ^{er} juillet 2021)	
800 m	1 min 51 s 9 ms (1912)	2 min 30 s 4 ms (1922)	1 min 41 s 11 ms (1997)	1 min 53 s 28 ms (1983)	1 min 40s 91 ms (2012)	1 min 53 s 28 ms (1983)	9,0	24,7	9,8	24,7	0,9	0,0
1 500 m	3 min 55 s 8 ms (1912)	4 min 17 s 3 ms (1967)	3 min 26 s (1998)	3 min 50 s 46 ms (1993)	3 min 26 s (1998)	3 min 50s 7 ms (2015)	12,3	10,5	12,3	10,5	0,0	0,0
3 000 m	8 min 36 s 8 ms (1912)	8 min 52 s 8 ms (1974)	7 min 20 s 67 ms (1996)	8 min 6 s 11 ms (1993)	7 min 20s 67 ms (1996)	8 min 6 s 11 ms (1993)	14,7	8,6	14,7	8,6	0,0	0,0
3 000 m steeple	8 min 49 s 6 ms (1954)	9 min 48 s 88 ms (1999)	7 min 55 s 28 ms (2001)	9 min 25 s 31 ms (2001)	7 min 53 s 63 ms (2004)	8 min 44 s 32 ms (2018)	10,2	3,9	10,6	10,9	0,4	7,3
5 000 m	14 min 36 s 6 ms (1912)	15 min 14 s 51 ms (1981)	12 min 39 s 36 ms (1998)	14 min 28 s 9 ms (1997)	12 min 35 s 36 ms (2020)	14 min 6 s 62 ms (2020)	13,4	5,0	13,8	7,4	0,5	2,5
10 000 m	30 min 58 s 8 ms (1911)	32 min 17 s 20 ms (1981)	26 min 22 s 75 ms (1998)	29 min 31 s 78 ms (1993)	26 min 11 s (2020)	29 min 1 s 3 ms (2021)	14,9	8,6	15,4	10,1	0,7	1,7
Marathon	2 h 4 min* (2003)	2 h 17 min 42 s (2005)	-	-	2 h 1min 39 s* (2018)	2 h 17 min 1 s (2017)	9,0	24,7	1,9	0,5	-	-
Marathon course mixte	-	2 h 15 min 25 s (2003)	-	-	-	2 h 14 min 4 s* (2019)	-	-	-	1,0	-	-

*Note : L'homologation des records du monde sur la distance marathon s'est effectuée relativement tard expliquant ainsi une faible progression des records ces dernières années. Toutefois, si l'on s'intéresse aux premières marques de meilleure performance mondiale qui ont pu être enregistrées, en 1908 en 2 h 55 min 18 s et en 1926 en 3 h 40 min 22 s, chez les hommes et les femmes, respectivement, mais qui ne faisaient pas hélas office de records du monde à l'époque, nous pouvons constater une forte progression entre cette première marque et aujourd'hui (2021) à hauteur de 30,6% et 39,2%, respectivement. De plus, il existe actuellement deux classements recensant les records du monde femmes sur marathon, l'un pour les performances réalisées lors de course mixte et un autre pour les courses femmes uniquement.

À l'échelle nationale française, il est possible de constater des évolutions de performances relativement similaires à celle de l'échelle mondiale (Tableau 9).

Tableau 9. Évolutions des records de France masculins et féminins en demi-fond et fond, du premier record homologué à aujourd'hui (2021)

Distance	Record de France et date						Évolution du record de France (en %)			
	Hommes	Femmes	Hommes	Femmes	Hommes	Femmes	Hommes	Femmes	Hommes	Femmes
	1 ^{er} record		Record en 2001		Record actuel		Du 1 ^{er} record à aujourd'hui (1 ^{er} juillet 2021)		De 2001 à aujourd'hui (1 ^{er} juillet 2021)	
800 m	2 min 14 s 0 ms (1892)	2 min 45 s (1921)	1 min 43 s 9 ms (1979)	1 min 56 s 53 ms (1995)	1 min 42 s 53 ms (2014)	1 min 56 s 53 ms (1995)	23,8	29,7	0,9	0,0
1 500 m	4 min 21 s (1893)	5 min 54 s 5 ms (1918)	3 min 31 s 51 ms (1999)	4 min 2 s 26 ms (1996)	3 min 28 s 98 ms (2003)	3 min 59 s 76 ms (2010)	20,3	32,5	1,4	1,2
3 000 m	9 min 22 s 4 ms (1895)	9 min 56 s 6 ms (1972)	7 min 30 s 78 ms (1998)	8 min 38 s 13 ms (1999)	7 min 30 s 78 ms (1998)	8 min 35 s 41 ms (2005)	19,9	13,6	0,0	0,6
3 000 m steeple	9 min 20 s 4 ms (1954)	9 min 55 s 69 ms (1999)	8 min 7 s 62 ms (1984)	9 min 44 s 68 ms (2001)	8 min 00 s 9 ms (2013)	9 min 25 s 62 ms (2009)	14,3	5,0	1,4	3,3
5 000 m	18 min 7 s (1894)	16 min 42 s (1981)	12 min 58 s 83 ms (2000)	14 min 47 s 79 ms (2000)	12 min 58 s 83 ms (2000)	14 min 43 s 90 ms (2004)	28,4	11,9	0,0	0,4
10 000 m	37 min 47 s (1894)	33 min 42 s 5 ms (1982)	27 min 22 s 78 ms (1992)	31 min 42 s 83 ms (1992)	27 min 17 s 29 ms (2019)	31 min 35 s 81 ms (2012)	27,8	6,3	0,3	0,4
Marathon	2 h 32 min 57 s* (1928)	3 h 16 min 13 s* (1972)	2 h 7 min 2 s* (2001)	2 h 29 min 4 s* (2001)	2 h 6 min 36 s (2003)	2 h 24 min 22 s (2010)	17,2	26,4	0,3	3,2
	2 h 6 min 36 s (2003)	2 h 28 min 28 s (2002)	-	-	-	-	0	2,8	-	-

Note : D'après les données du site FFA et *World Athletics* [104]. *meilleure performance française et non « record » de France, car distance sur marathon pas encore homologuée par la FFA en 2001.

En effet, à l'image des analyses de Dyer et Dwyer [59] portant sur les records du monde masculins (entre 1882 et 1982) ou de Lippi et al. [115], nous pouvons constater, entre la réalisation du premier record de France et le record de France actuel, une nette évolution du taux de progression, quel que soit le sexe (Tableau 9). De plus, en accord avec les taux de progression, relevés par Billat [8] entre 1997 et 2017 ou Dyer et Dwyer [59] entre 1972 et 1982 à l'échelle mondiale masculine, nous pouvons également constater une faible évolution du taux de progression du record de France depuis ces vingt dernières années, indépendamment du sexe (Tableaux 7 à 9).

Toutefois, il est important de souligner que malgré une amélioration des records mondiaux ou nationaux (caractérisée par des vitesses moyennes de course à pied plus élevées) que la relation vitesse-distance semble relativement similaire avec les années de pratique (*i.e.*, forme de la courbe) [102] (Figure 20).

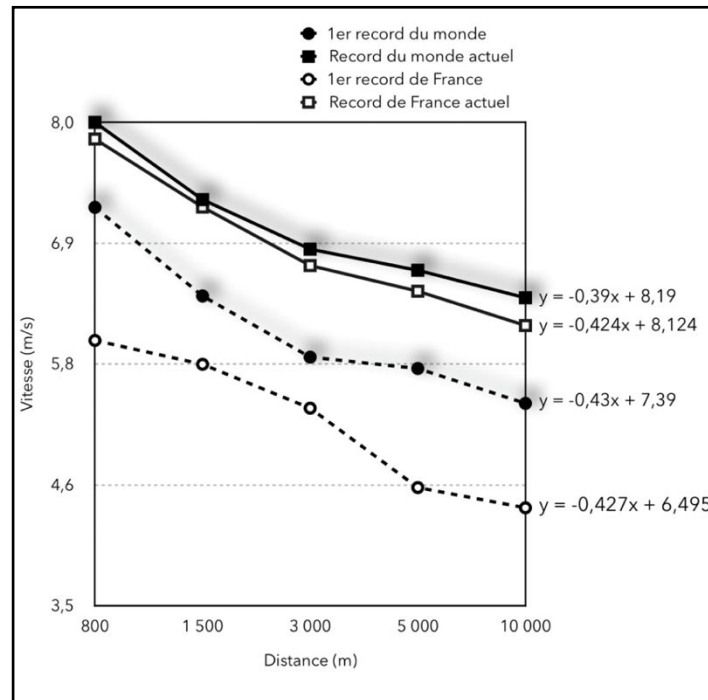


Figure 20. Courbes vitesse-distance illustrant l'évolution des premiers records du monde sur piste masculins (1908-1912) et les premiers records de France masculins (1892-1895) avec les records du monde et de France actuels masculins (au 1^{er} juillet 2021)

De plus, si l'on s'intéresse désormais à l'évolution des performances et plus précisément aux taux de progression entre les sexes (*i.e.*, hommes et femmes) quelle que soit l'échelle (*i.e.*, mondiale ou nationale), nous pouvons constater que les performances féminines se sont améliorées et ont encore tendance à s'améliorer davantage que les hommes ces dernières années (Tableaux 7 à 9). Les différences physiologiques existantes entre les sexes [59, 116-119] expliquent le fait que les hommes ont de meilleures références chronométriques par rapport aux femmes et qu'il soit peu probable que ces dernières les dépassent un jour [16, 111, 120]. Toutefois, il est possible de constater que les femmes réussissent avec le temps à

comblent les écarts sur les meilleurs temps masculins sur certaines distances de course à pied [10, 13, 16, 59]. Dans leur étude de 2005, Cheuvront et al. [10] ont constaté, par exemple, un différentiel entre les records masculins et féminins à hauteur de 11,9%, 14,1% et 12,1% sur 1 500 m, 5 000 m et 10 000 m respectivement ; alors qu'aujourd'hui (en 2021) nous constatons un différentiel de 10,4%, 10,8% et 9,8% respectivement, soit un différentiel réduit sur ces mêmes distances. Les différences entre les records du monde masculins et féminins sont par conséquent toujours existantes à l'heure actuelle, mais ces dernières ont tendance à s'estomper au fil des années. Sur marathon par exemple, les records du monde féminins ont été améliorés à hauteur de 4,0% contre 1,5% pour les records masculins entre 1985 et 2004 [10]. L'une des explications plausibles de ces évolutions chez les femmes pourrait être d'ordre historique et social en lien avec leur arrivée tardive dans le monde de la course à pied et des compétitions avec notamment une participation croissante au fil des ans [59, 1] (cf. première partie, chapitre I).

Ainsi, au-delà de l'analyse historique des records, la compréhension des circonstances dans lesquelles les performances sportives sont exécutées peut présenter un véritable intérêt pour les athlètes, les entraîneurs et les fédérations afin d'identifier des priorités de développement. Si de nombreux records ont été améliorés avec les années (Tableaux 7 à 9), nous pouvons également constater des progressions récentes plus modérées avec parfois l'atteinte d'un plateau (*i.e.*, stagnation chronométrique) à un moment donné de certains records selon les disciplines [16, 73, 74, 115, 120, 121]. De ces principaux constats, nous sommes amenés à nous questionner sur les conditions de réalisation qui auraient pu ou qui pourraient impacter ces évolutions chronométriques en nous demandant, par exemple, si les modifications du règlement (édicteant les paramètres, les conditions, et les normes) avec l'autorisation de courses mixtes, notamment sur route, n'auraient pas permis d'améliorer les performances féminines en offrant davantage de densité

; si physiologiquement l'homme n'avait pas atteint ses propres limites dans certaines épreuves [3, 59, 74, 111, 115, 121] ; ou encore, si des athlètes n'auraient pas eu recours à des techniques de conduite dopante, avec le détournement de certaines substances développées à des fins médicales (e.g., hormones de croissance, érythropoïétine) puis détournées dans une optique de performance sportive, rendant les records difficilement atteignables à l'heure actuelle [3, 59, 115, 122-124]. Par conséquent, lors de l'analyse historique des records, il semble important de ne pas négliger les différents facteurs pouvant limiter ou optimiser la performance sportive [3, 9, 16, 59, 70, 74, 102, 115]. En d'autres termes, il semble essentiel d'identifier et de prendre en considération l'ensemble des paramètres qui peuvent faire partie prenante de ces évolutions chronométriques afin d'en appréhender la performance en course à pied (i.e., paramètres physiologiques, biomécaniques, psychologiques, environnementaux, technologiques).

2. Facteurs de performance

La performance en course à pied est multifactorielle et déterminée par l'intégration de nombreux facteurs d'ordre physiologique, biomécanique, psychologique, environnemental et technologique (Figure 21) [59, 83, 9, 3, 74, 70, 18]. Par conséquent, la performance exige la prise en considération de multiples facteurs qui peuvent être dépendants de l'entraînement (e.g., $\dot{V}O_{2\max}$, capacité d'endurance aérobie...) mais aussi indépendants et hors de contrôle de l'athlète et de l'entraîneur (e.g., génétique, âge, conditions climatiques) [3, 9, 18, 19, 74, 83].

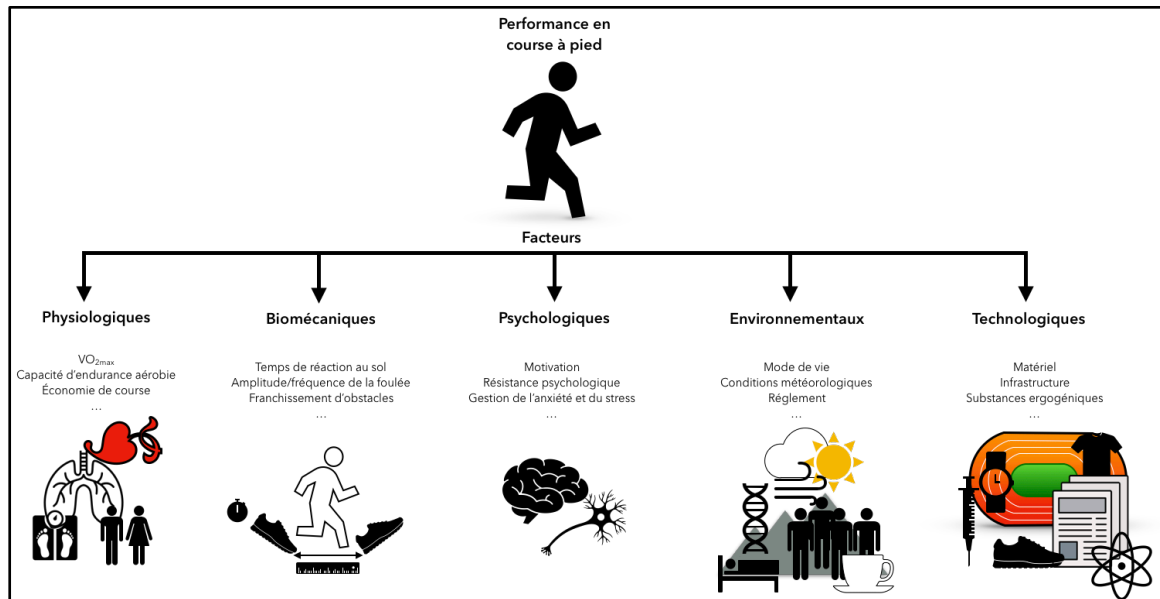


Figure 21. Facteurs de performance en course à pied

a. Paramètres physiologiques

Sur le plan physiologique, les épreuves de course à pied offrent un panel d'efforts continus relativement étendu. Concernant le demi-fond et le fond, allant des distances les plus courtes comme le 800 m aux distances les plus longues comme le marathon, l'ensemble des filières énergétiques (*i.e.*, voie des phosphagènes, voie glycolytique et voie de la phosphorylation oxydative) interviennent, mais l'une prédomine pour produire de l'énergie (Figure 9) [8, 59, 60]. Lors de l'effort, même si la contribution relative du métabolisme aérobie (*i.e.*, voie de la phosphorylation oxydative) est considérable en demi-fond et fond, les trois systèmes énergétiques entrent en action et chaque métabolisme, caractérisé par sa puissance (*i.e.*, quantité maximale d'énergie utilisable par unité de temps) et sa capacité (*i.e.*, quantité totale d'énergie pouvant être produite), possède son inertie et son délai d'intervention en fonction de la durée et de l'intensité de l'effort produit (Figure 9) [8, 59, 60, 105]. Par exemple, dans les épreuves de 800 m et de 1 500 m (*i.e.*, demi-fond court), courues à des intensités élevées, la voie glycolytique et la voie de phosphorylation oxydative (*i.e.*, métabolismes anaérobie lactique et aérobie, respectivement), contribuent

fortement toutes les deux à la production d'énergie ; alors que dans les épreuves allant du 5 000 m au marathon (*i.e.*, du demi-fond long au fond), la voie de la phosphorylation oxydative est prépondérante (Tableau 10, Figure 22) [8, 59, 60, 63, 64, 67, 125-127].

Tableau 10. Exemple de contribution des métabolismes énergétiques en fonction de la distance de course et du sexe

Distance (m)	Part des métabolismes d'énergie requis (%)							
	Dyer & Dwyer [59]		Spencer et al. [63]		Hill [64]		Duffield et al. [65, 66]	
	Anaérobie	Aérobie	Anaérobie	Aérobie	Anaérobie	Aérobie	Anaérobie	Aérobie
800	50	50	31	69	33 (F)	67 (F)	30 (F)	70 (F)
					39 (H)	61 (H)	40 (H)	60 (H)
1 500	-	-	17	83	17 (F)	83 (F)	23 (H)	77 (H)
					20 (H)	80 (H)	14 (F)	86 (F)
3 000	-	-	-	-	-	-	14 (H)	86 (H)
10 000	15	85	-	-	-	-	6 (F)	94 (F)
Marathon	5	95	-	-	-	-	-	-

Note. H : Homme ; F : Femme

Selon la durée et l'intensité de l'effort prédomine un type de fibre musculaire se distinguant en fonction de ses caractéristiques physiologiques. Lors de courses à dominante aérobie, les fibres musculaires à contraction lente (*i.e.*, fibres de type I), riches en capillaires sanguins et en myoglobine, sont principalement sollicitées [8, 60, 70, 128, 129]. Elles sont notamment très résistantes à la fatigue, contrairement aux fibres musculaires à contraction rapide de type II beaucoup plus fatigables [129]. Si les fibres rapides de type IIb sont recrutées dans les efforts de courte durée et à intensité élevée, les fibres rapides de type IIa sont, quant à elles, recrutées lors d'efforts de type intermédiaire [128]. En d'autres termes, les coureurs de très courtes distances (*i.e.*, épreuves de *sprint*) solliciteront davantage les fibres de type IIb alors que les coureurs de demi-fond ou de fond solliciteront davantage les fibres de type IIa et de type I, respectivement [128, 70, 130].

Prédiction de la performance en course à pied : demi-fond et fond

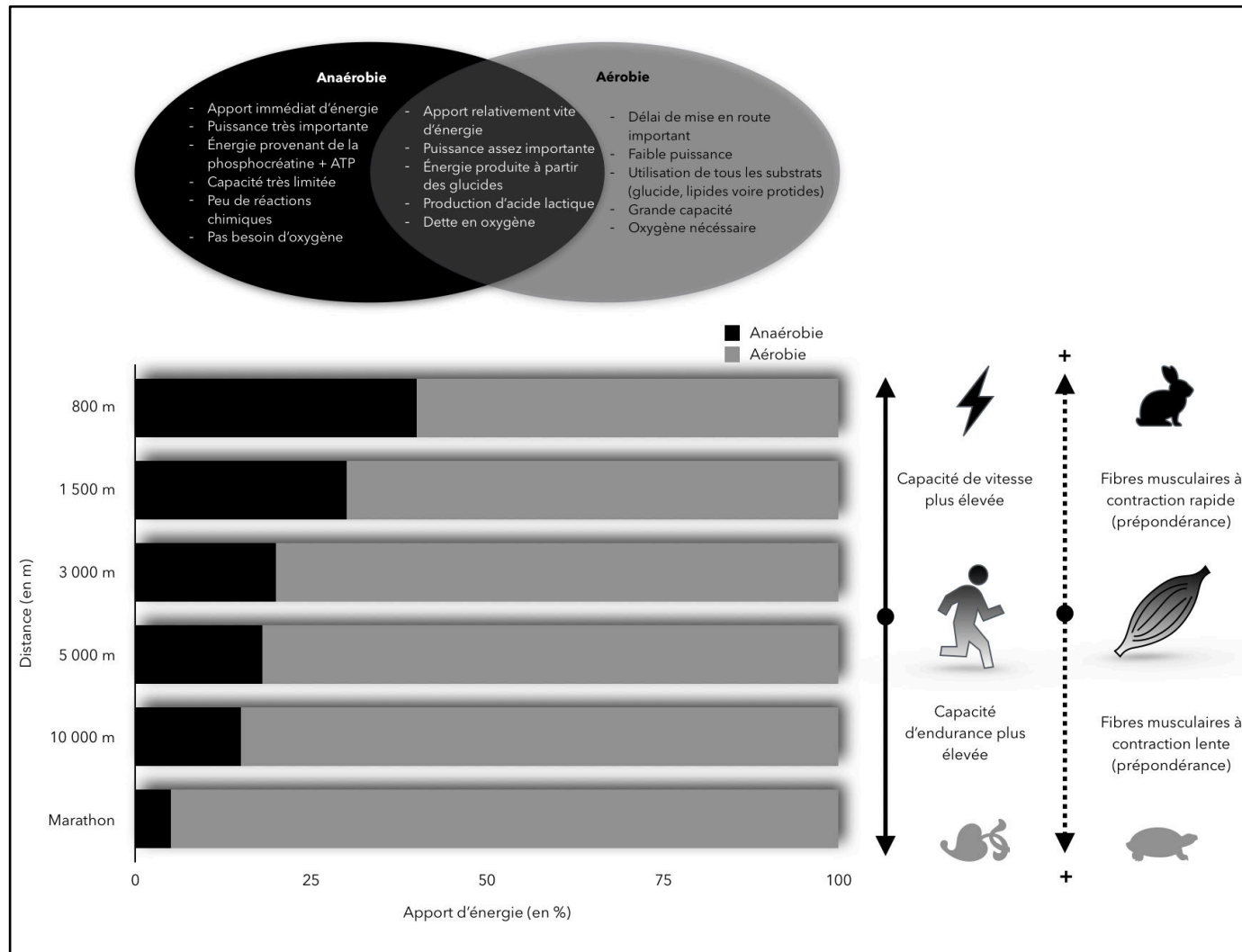


Figure 22. Contributions relatives des filières énergétiques en fonction des distances de courses en demi-fond et fond

Au-delà de ce rappel, la suite de cette section est consacrée à une brève revue bibliographique des principaux facteurs physiologiques fondamentaux de la performance en course de demi-fond et de fond, soit le $\dot{V}O_{2max}$, l'économie de course et la capacité d'endurance aérobie [11, 14, 68, 70, 125, 131].

Le $\dot{V}O_{2max}$ définit les limites de la capacité des systèmes cardiovasculaire, respiratoire et musculaire, et renseigne sur la consommation maximale d'oxygène qu'une personne est capable d'utiliser par minute lors d'un effort maximal [132, 133]. Le $\dot{V}O_{2max}$ détermine la puissance du système aérobie (correspondant généralement à une intensité de course maximale à hauteur d'environ 6 min) et permet notamment d'évaluer l'efficacité des systèmes respiratoire et cardiovasculaire, et musculaires de l'athlète. Le $\dot{V}O_{2max}$ peut être limité et régulé par des facteurs liés au débit sanguin maximal de l'individu et à la quantité maximale d'oxygène que le muscle peut utiliser. L'équation de Fick (*i.e.*, $\dot{V}O_2 = \dot{Q}c \times [Ca_{O_2} - Cv_{O_2}]$) permet de faire le lien avec le débit cardiaque (*i.e.*, $\dot{Q}c$ en $ml \cdot min^{-1}$, soit le produit de la FC et du volume d'éjection systolique : VES) et l'utilisation périphérique de l'oxygène (*i.e.*, différence artérioveineuse en oxygène : $Ca_{O_2} - Cv_{O_2}$ en ml). Le $\dot{V}O_{2max}$, limité par les valeurs maximales de $\dot{Q}c$ et $[Ca_{O_2} - Cv_{O_2}]$ [134], est par conséquent l'un des principaux reflets du potentiel de performance dans les efforts aérobies, soit un facteur clé de la performance en demi-fond et de fond [11, 33, 68, 108, 108, 135-137]. Généralement, dans une population hétérogène, plus le $\dot{V}O_{2max}$ est important, plus le coureur est capable de maintenir une vitesse (absolue) importante sur une longue durée [8, 130, 136]. Cependant, des athlètes entraînés (et ayant le même $\dot{V}O_{2max}$) n'auront pas nécessairement les mêmes performances en course à pied [85, 138]. De même certains athlètes ayant un $\dot{V}O_{2max}$ inférieur pourraient être capable de compenser en utilisant un $\% \dot{V}O_{2max}$ plus élevé lors de l'effort [139].

D'autres facteurs tels que l'économie de course (*i.e.*, $\dot{V}O_2$ pour une vitesse sous-maximale donnée) ainsi que la capacité d'endurance aérobie (*i.e.*, capacité à maintenir un $\% \dot{V}O_{2max}$ pendant la durée de l'effort) sont à prendre en considération pour expliquer les différences interindividuelles dans les performances [11, 68, 85, 140-142].

L'économie de course, de l'anglais *running economy*, est la consommation d'oxygène par min à l'état d'équilibre (conditions métaboliques stables, sous-maximales) pour une vitesse de course donnée [85]. L'économie de course peut être altérée par des mouvements parasites (*e.g.*, mouvements rotatoires, déficit de force musculaire), une foulée inadaptée (*e.g.*, rapport amplitude-fréquence), et/ou une importante co-contraction des muscles agonistes-antagonistes, entraînant ainsi un surcoût énergétique. En d'autres termes, l'économie de course est extrêmement dépendante de facteurs physiologiques (*i.e.*, fatigue musculaire, âge) et biomécaniques (*e.g.*, oscillation excessive du centre de gravité, temps d'appui au sol, inclinaison du buste) pouvant être la cause d'une augmentation du coût énergétique en course à pied [33, 81, 143-145]. En effet, la notion de coût énergétique (CE), principalement développée en course à pied, définie comme la quantité d'oxygène consommée par unité de distance parcourue (en $ml.kg^{-1}.m^{-1}$) [11, 68], est un des paramètres explicatifs de la performance. Lorsque les vitesses de déplacement sollicitent des intensités sous-maximales, la faculté que possède l'athlète de minimiser sa dépense en énergie pour une intensité d'effort donnée joue indéniablement un rôle sur la performance [85, 146]. Par conséquent le CE renseigne sur la dépense énergétique occasionnée au cours d'un exercice sur une distance donnée, et moins il y aura d'oxygène consommé à une allure donnée, meilleure sera l'économie de course du coureur. Il a été démontré que les coureurs de demi-fond (*e.g.*, 800 m, 1 500 m) étaient plus économes que les coureurs de fond (*e.g.*, marathon) à des vitesses supérieures à $19 km.h^{-1}$, mais qu'ils étaient

moins économes que les coureurs de fond à des vitesses inférieures [141]. En effet, en raison d'un déplacement vertical du centre de masse plus faible (possiblement en lien avec des adaptations neuromusculaires induites par l'entraînement sur de longues distances) [147] ou des valeurs d'indice de masse corporelle (*i.e.*, exprimant le rapport masse/taille² ; IMC) plus faibles également [15], plusieurs études ont notamment démontré que les coureurs de fond étaient plus économes que les coureurs de demi-fond [148, 149] à des vitesses de déplacement sous-maximales. À noter également que dans les courses techniques comme le 3 000 m *steeple*, qu'au-delà de l'économie de course, l'économie de franchissement (en lien avec la technique de course) apparaît comme un élément primordial afin de ne pas générer ou limiter des changements d'allure ou une perte de vitesse, pouvant impacter directement la performance chronométrique [81, 82]. Comme suggéré par Cavanagh [150], de modestes améliorations de l'économie de course pourraient se traduire par des gains de performance substantiels, notamment chez les coureurs élites. De ce fait, au sein d'une population athlétiquement homogène, la capacité que possède le coureur à dépenser peu d'énergie pour une vitesse de déplacement donnée a été citée comme un des principaux indicateurs de la performance [85, 144, 148]. En outre, au regard du fait que la VMA (*i.e.*, la vitesse de course minimale à laquelle est atteint $\dot{V}O_{2max}$) puisse être obtenue par la division du $\dot{V}O_{2max}$ et du CE, cette dernière peut refléter également le potentiel de performance dans les efforts aérobies.

Enfin, la capacité d'endurance aérobie est également l'un des facteurs importants pour la performance [11, 29, 33, 142]. À la différence du $\dot{V}O_{2max}$ qui caractérise la puissance aérobie, la capacité d'endurance aérobie est la capacité à maintenir une fraction très élevée de $\dot{V}O_{2max}$ pendant une durée donnée [140]. Généralement, plus le $\dot{V}O_{2max}$ (voire la VMA) et la capacité d'endurance aérobie sont élevés, meilleure est la performance en course à pied. Par conséquent, à $\dot{V}O_{2max}$ ou VMA

équivalent, l'athlète ayant une plus grande capacité d'endurance sera capable de réaliser une meilleure performance (*i.e.*, courir plus longtemps, maintenir une vitesse relative plus élevée) [60]. De plus, comme évoqué dans le chapitre précédent (*cf.* première partie, chapitre I), le pourcentage de $\dot{V}O_{2max}$ ou de VMA sollicités seront dépendants de l'effort réalisé (Tableau 2 et Figure 23) [8, 33, 59, 60, 126, 135, 151-153]. Par exemple, les coureurs de 10 000 m sollicitent environ 95% de leur $\dot{V}O_{2max}$ sur la totalité de la distance, tandis que les marathoniens maintiennent plutôt 80-90% de leur $\dot{V}O_{2max}$ durant un marathon [142].

Ainsi, en augmentant le $\dot{V}O_{2max}$ ou la fraction du $\dot{V}O_{2max}$ soutenue lors de l'effort, ou en diminuant le CE de la course, ou encore en combinant ces trois paramètres, la performance peut être améliorée.

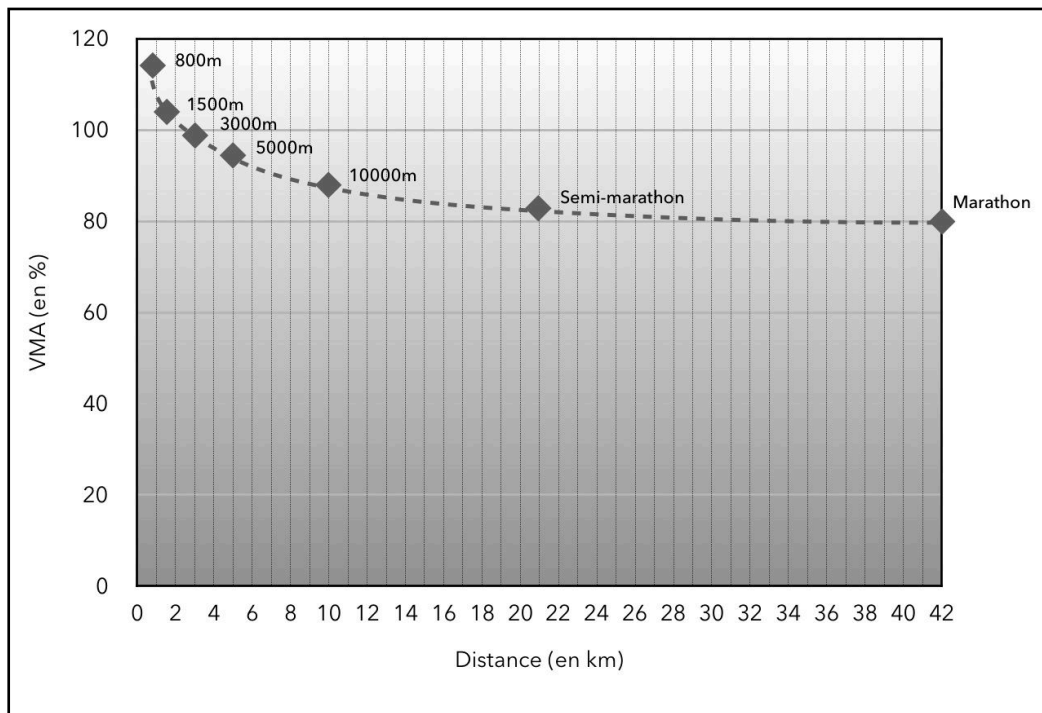


Figure 23. Évolution de la VMA en fonction de la distance d'épreuves de demi-fond et fond, d'après Aubert et Choffin [60]

Toutefois, ces facteurs physiologiques clés de la course à pied (*i.e.*, le $\dot{V}O_{2max}$, l'économie de course et capacité d'endurance aérobie) dépendent également d'adaptations liées à l'entraînement (*e.g.*, augmentation du VES, de la densité capillaire...) où les athlètes pourront être en mesure de développer une plus grande vitesse de course ou de maintenir une vitesse donnée sur une plus longue durée [33, 68, 130, 137, 154, 155] ; mais aussi de caractéristiques individuelles comme l'âge, le sexe et l'IMC [15, 17, 22, 116, 141, 144].

L'IMC est, aussi considéré comme un indicateur pertinent permettant de différencier les capacités des athlètes (de la population générale) et leur niveau de performance [15]. Le constat est tel que la masse et la taille des athlètes de niveau élite diminuent avec l'allongement de la distance [15]. Sur marathon, la vitesse de course est proportionnellement inverse à l'IMC (*i.e.*, généralement, plus l'IMC est faible, plus la vitesse de course est importante) [7]. Sur le plan énergétique, l'excédent pondéral a une conséquence directe sur le $\dot{V}O_{2max}$ exprimé en fonction de la masse corporelle) et de ce fait, il peut-être pénalisant en course à pied, notamment en demi-fond et fond [156].

L'âge, hors de contrôle de l'athlète et de l'entraîneur, est également un facteur à prendre en considération puisqu'en vieillissant, les capacités physiologiques se dégradent (*e.g.*, capacité à produire une certaine vitesse, diminution de la FC maximale : FC_{max} , diminution du $\dot{V}O_{2max}$) et l'âge peut alors devenir un facteur limitant [95, 127, 157]. Cependant, sur certaines distances de course, notamment en course de fond, il est possible d'atteindre de meilleures performances à un âge légèrement supérieur que sur des courses de plus courtes distances (*e.g.*, 25-26 ans pour des épreuves de demi-fond comme le 800 ou le 1 500 m contre 28-29 ans sur l'épreuve de marathon) [13, 94, 158, 159]. Cette différence d'âge en fonction des disciplines pourrait notamment s'expliquer par le fait que les longues distances

nécessitent davantage d'expérience, probablement dans la gestion de son effort [13, 68, 152, 153, 160].

Enfin, comme évoqué précédemment, le facteur sexe a aussi toute son importance au regard des différences physiologiques existantes et de leurs conséquences afin de pouvoir appréhender la performance en course à pied (e.g., $\dot{V}O_{2\max}$ et taux d'hémoglobine inférieurs, et taux de masse grasse plus important chez les femmes) [59, 116-119].

b. Paramètres biomécaniques

La réalisation de performance, quelle que soit la distance, impose aux athlètes et aux entraîneurs de comprendre comment dépasser la motricité usuelle de la course à pied afin d'en affiner son efficacité et d'accéder à un plus haut degré d'habileté technique. L'approche biomécanique est alors fondamentale et plurielle en course à pied au regard de la répétition des mouvements cycliques dépendants de la durée et de l'intensité de l'épreuve. Les caractéristiques de l'épreuve (spécificités techniques lors d'un 3000 m *steeple* par exemple) engendrent des adaptations motrices (e.g., organisation segmentaire, rapport amplitude-fréquence de la foulée) qui sont propres à chaque athlète [60, 161-163]. Les variables telles que la longueur (*i.e.*, amplitude, soit la distance horizontale moyenne entre le contact d'un pied avec le sol et le contact controlatéral suivant) et la fréquence de la foulée (*i.e.*, nombre de foulées par unité de temps), dont leur produit définit la vitesse de la foulée (*i.e.*, vitesse horizontale moyenne du contact d'un pied avec le sol au contact controlatéral suivant), constituent des données essentielles de la technique de course. En effet, l'optimisation du rapport amplitude/fréquence est nécessaire pour l'atteinte de la vitesse optimale pour un effort donné [60, 130, 162, 163]. De plus, en vue d'une course efficace et plus rapide, les membres supérieurs et inférieurs

reliés au buste jouent également un rôle crucial permettant notamment de stocker et de libérer l'énergie nécessaire, soit d'assurer le transfert d'énergie, le long de la chaîne cinétique [130]. Toutefois, si les bras ne peuvent compenser l'élan des jambes le coût métabolique lors de l'effort s'accroît fortement [164, 165].

Il est alors important de prendre en considération le fait que selon le type d'effort réalisé, les paramètres physiologiques (présentés précédemment dans la section a.) et biomécaniques sont étroitement dépendants et déterminants dans la réalisation de la performance en course à pied [70]. À la différence des courses de fond où l'économie de course est fondamentale [85, 166], les courses de demi-fond exigent des vitesses relativement élevées soit une puissance mécanique développée plus importante lors de l'effort [68, 70, 125]. De ce fait, des différences biomécaniques (*i.e.*, cinétiques et cinématiques) entre les performances des demi-fondeurs et les fondeurs sont observables lors de la course [162, 167]. À vitesse égale par exemple, la flexion de genou, le temps de contact au sol et le déplacement du centre de masse sont plus importants chez le coureur de fond ; et à la différence des demi-fondeurs, les fondeurs ont également une longueur de foulée réduite [162, 167]. Les différentes sollicitations physiologiques peuvent, par conséquent, permettre de comprendre les différences cinétiques et cinématiques observables dans des courses de demi-fond et de fond et les raisons potentielles des modifications de la technique de course en fonction de la distance parcourue et des spécificités techniques de la discipline [70, 81, 162, 165-170]. En course de fond et avec l'allongement de la distance, il est moins nécessaire de balancer exagérément les bras (compensant le mouvement des membres inférieurs qui est en l'occurrence moins rapide en fond qu'en demi-fond) afin d'éviter une dépense d'énergie inutile [161, 169]. De plus, avec l'augmentation de la distance (indépendamment de la distance totale à parcourir), de potentielles dégradations techniques peuvent être à l'origine de modifications biomécaniques de la course à pied (*e.g.*, augmentation

du temps de réaction au sol, orientation du buste vers l'avant, mouvements parasites des membres supérieurs) [60, 163, 170, 171].

Enfin, au-delà de tous ces éléments abordés, d'autres facteurs tels que le style de course (e.g., attaque avant ou arrière pied), les surfaces de sol (i.e., rigidité des revêtements ; e.g., piste en cendrée, en herbe, en synthétique), les chaussures de course (e.g., minimalistes, maximalistes, avec carbone et/ou mousse super élastique) ou semelles orthopédiques peuvent également engendrer des modifications de paramètres biomécaniques, soit des adaptations du *pattern* de course (e.g., fréquence, temps de contact, pose de pied) [105, 161, 166, 172-178]. Par exemple, certains modèles de chaussures peuvent notamment améliorer l'économie de course *via* des modifications cinétiques et cinématiques lorsque le pied entre en contact avec le sol, et être par conséquent à l'origine d'une optimisation et amélioration de la performance sportive [174, 177, 179-181]. Par conséquent, au regard de ces paramètres, il apparaît important de prendre en considération les diverses technologies (i.e., revêtement des sols, chaussures) pouvant influencer les paramètres physiologiques et biomécaniques afin de pouvoir appréhender la performance sportive dans son ensemble.

c. Paramètres psychologiques

La course à pied, en tant qu'activité de moyenne et de longue durée, requiert des qualités psychologiques qui apparaissent comme un véritable facteur de performance (e.g., motivation, résistance psychologique, aptitude à tolérer la douleur et à maintenir un effort pendant l'entraînement et la compétition, capacité à gérer l'anxiété et le stress...) [9, 105, 182]. Il est intéressant de différencier les « qualités psychologiques », faisant référence au résultat souhaité (e.g., envie de gagner, surmonter la pression, se dépasser, avoir la maîtrise de soi...), des

« méthodes » ou « techniques psychologiques » (e.g., imagerie, relaxation, autosuggestion, fixation d'objectifs) en tant que moyens de promouvoir les résultats souhaités par l'application de ces techniques [183, 184]. En d'autres termes la compétence et la technique peuvent être définies comme étant, respectivement, la capacité requise à effectuer une tâche spécifique et la procédure utilisée permettant d'améliorer cette capacité afin d'être en mesure d'accomplir cette tâche [185]. La mise en place de « stratégies », terme utilisé pour désigner les moyens ou le plan d'action qui sont utilisés pour améliorer les compétences psychologiques en utilisant une ou plusieurs techniques psychologiques [185], peut permettre d'améliorer la performance [153, 185, 186]. Si le stress, par exemple, peut se présenter à plusieurs niveaux (e.g., physiologique, émotionnel), le recours à de simples stratégies de gestion du stress peut diminuer significativement la consommation d'oxygène pendant l'exercice [187] et par conséquent optimiser la performance. Il existe en effet, une possible relation en forme de U inversé (loi de Yerkes-Dodson) entre la performance et le niveau d'activation (stress, anxiété) [130, 188]. Cette loi énonce le fait que des niveaux d'activation trop faibles ou trop élevés (i.e., hypo stress ou hyper stress, respectivement, associés à l'ennui et la déconcentration) se traduisent principalement par de mauvais résultats ou de contre-performance ; et qu'à la différence un niveau d'activation optimale, soit le « bon stress » (associé à une anxiété modérée, au relâchement, au calme et à la sérénité, à la gestion de ses émotions), se traduit généralement par une performance ou la réussite [130, 188].

L'effort étant contrôlé par le cerveau *via* une série de boucles neurosensorielles [189, 190], afin de faire basculer l'athlète dans une zone d'activation optimale, différents facteurs « psychologiques » peuvent contribuer à la réussite de la performance sportive [182]. La connaissance de ses points forts et l'utilisation de ces derniers sans s'enfermer dans un mode de fonctionnement apparaissent comme des paramètres tactiques individuels importants. En demi-fond et fond, si l'aspect

énergétique est essentiel pour réussir, la connaissance de la discipline ou la capacité à tolérer la douleur et maintenir l'effort, par exemple, jouent également un rôle prépondérant dans la performance [60, 185]. La gestion de l'effort, le comportement en peloton, la motivation pour l'obtention de la victoire ou la réalisation de records sont par conséquent autant de paramètres qui sont à prendre en considération afin d'appréhender la performance [191-194]. Une stratégie d'allure de course optimale peut se définir comme étant « la puissance qui permettrait de réaliser la durée de course la plus courte pour un athlète le jour J » [195]. Les décisions stratégiques, prises à l'avance, concernent l'approche globale de l'événement, alors que les décisions tactiques, prises lors de la course, répondent aux changements de l'état physiologique de l'athlète et du comportement des rivaux [196]. Pour les épreuves de longue distance, comme le marathon, si la gestion de l'effort a toute son importance [197-199] ; d'autres paramètres tels que les conditions climatiques (e.g., humidité, vent, température extérieure), la topographie de la course (e.g., dénivelé, altitude, virages) et les tactiques de course sont à prendre en compte (e.g., *positive split*, *negative split*, course en peloton) [5, 7, 20, 192, 200-202]. Sur marathon, des différences tactiques entre les sexes possiblement en lien avec les aspects psychologiques et physiologiques peuvent être observées [198, 199, 203]. Par exemple, les hommes ont tendance à opter davantage pour une stratégie à « risque » en partant sur des allures rapides quitte à moduler ou ralentir leur vitesse ensuite, alors que les femmes misent généralement sur la prudence [203]. Sur les distances courtes comme le 800 m ou le 1 500 m, courues à des allures relativement élevées, toute erreur tactique, indépendamment de la distance parcourue ou qu'il reste à parcourir, se paye généralement immédiatement alors qu'à l'inverse, sur les distances plus longues comme le 10 000 m, les aspects tactiques sont généralement décisifs en fin de course. En effet, sur piste, la stratégie d'allure diffère selon les distances et le contexte des courses [20, 194, 204]. Sur 800 m, des vitesses de course plus élevées

sont généralement atteintes dans le premier tour à la différence du second tour où la capacité d'augmentation de vitesse est limitée ; alors que dans les épreuves plus longues, comme le 5 000 m ou le 10 000 m, une accélération finale (dans le dernier tour, voire le dernier km) peut généralement être observée en raison du maintien d'une réserve d'énergie pendant la première partie de la course [194, 204]. Enfin, comme sur marathon, des différences de stratégies de course entre les sexes, en lien avec des caractéristiques physiologiques et techniques selon l'épreuve (e.g., *steeple-chase*), sont observables sur piste [12, 20, 205].

d. Paramètres environnementaux

Les facteurs environnementaux (e.g., mode de vie, infrastructure, règlement...) sont autant de paramètres qui peuvent impacter la performance et son évolution en course à pied [3, 83, 74, 6, 9, 105, 206].

Il est probable que les facteurs génétiques, hors de contrôle de l'athlète, aient une influence sur les facteurs de la performance (e.g., $\dot{V}O_{2max}$, capacité d'endurance et force musculaire) et par conséquent sur les performances en course à pied [34, 182, 207-209]. Les athlètes prédisposés génétiquement à la performance auraient plus de chances de répondre positivement à certains types d'entraînement [207]. Toutefois, la prédisposition génétique n'est pas suffisante et la réalisation de performance nécessite notamment un accompagnement de qualité ainsi qu'un mode de vie propice à l'optimisation de ces dernières, notamment à haut et très haut niveau, afin de pouvoir faire la différence en compétitions [210, 211].

L'entraînement, et plus précisément le type d'entraînement (e.g., qualitatif vs quantitatif) va conditionner la performance en générant notamment des stimuli en lien avec la charge définie (*i.e.*, intensité, volume, durée, fréquence d'entraînement, récupération) afin d'augmenter les capacités de l'athlète [8, 9, 202, 212-214]. Si

l'entraînement est bien calibré, diverses adaptations physiologiques comme l'érythropoïèse et l'angiogenèse vont permettre par exemple d'améliorer l'apport en oxygène des cellules musculaires [215-219]. Toutefois, si le programme d'entraînement n'est pas défini en fonction du profil de l'athlète, ce dernier peut s'avérer contre-productif générant des adaptations inverses (e.g., syndrome de surentraînement, blessures) [9, 220].

Au-delà de l'interaction des capacités physiques des athlètes et de l'entraînement, les performances sportives sont dépendantes et se développent dans des conditions environnementales favorables (*i.e.*, conditions météorologiques, topographie des parcours) [5-7, 201, 221, 222]. Lorsque les conditions météorologiques sont difficiles, par exemple de forte chaleur ou de fort taux d'humidité, le corps est amené à lutter pour éviter un état d'hyperthermie. Si la sudation est un moyen d'atténuer cette situation de « surchauffe », elle reste cependant moins efficace lorsque la chaleur ou l'humidité de l'air empêchent une évaporation efficace. Sur les épreuves de longue distance, comme le marathon, l'un des principaux facteurs environnementaux influençant la performance, est la température de l'air [5, 7]. Les facteurs tels que la topographie et l'altitude sont également des facteurs importants qui influencent directement les performances en course à pied [202]. Si l'altitude peut s'avérer être un avantage pour les épreuves de course à pied de courte distance (*i.e.*, *sprint*) en lien avec la diminution de la résistance à l'air ; elle peut être à l'inverse un désavantage non négligeable sur les plus longues distances (nécessitant un apport d'oxygène supérieur) en lien avec la diminution de la pression barométrique ayant pour conséquence directe une diminution de la pression partielle de l'oxygène [202, 221].

Enfin, les différents facteurs historiques, géographiques, politiques, sociologiques, culturels et économiques sont également des paramètres qui ont impacté et qui

continueront de faire évoluer les performances et leur environnement [1, 59, 74], à l'image du règlement (à la fois la source et la conséquence de l'évolution des disciplines sportives) qui édicte notamment les paramètres, les conditions et les normes selon lesquelles les athlètes doivent évoluer et se confronter pour réaliser des performances [1, 3, 59, 60, 73, 74, 83].

e. Paramètres technologiques

Pendant de nombreuses années, les athlètes ont cherché à améliorer leurs performances *via* des équipements (*i.e.*, matériels, pistes d'athlétisme) et des méthodes ergogènes (*e.g.*, compléments alimentaires, boissons énergétiques) visant à optimiser ou augmenter leurs capacités physiques (*e.g.*, force musculaire, vitesse et puissance) [8, 115, 129]. Au-delà de l'apport indéniable des connaissances en physiologie et biomécanique dans l'entraînement sportif, de nombreux progrès pharmacologiques (*i.e.*, médicaments, traitements) et technologiques (*e.g.*, textiles, chaussures, matériels connectés, revêtement des surfaces de course) ont été réalisés par les industries et équipementiers afin d'optimiser les performances et faire repousser les limites des athlètes [3, 223-227].

Parmi les nombreuses évolutions matérielles, la chaussure de course est l'un des facteurs technologiques fondamentaux à prendre en considération afin d'appréhender la performance. Cette dernière a indéniablement évolué grâce aux technologies et matériaux permettant notamment l'absorption des chocs, la prévention du risque de blessure, l'amélioration de l'économie de course dans le but d'améliorer la performance [174, 179, 181, 228]. Les technologies utilisées à l'heure actuelle dans les chaussures « nouvelles générations » (*e.g.*, plaque en fibre de carbone et mousse super élastique), permettent aux coureurs de diminuer considérablement le coût énergétique lors de l'effort (sur piste ou sur route) leur

donnant ainsi la possibilité de se préserver davantage musculairement tout en maintenant une vitesse de course élevée [226, 229]. Depuis l'intégration en compétition de ce type de chaussures, reconnues pour augmenter l'économie de course à hauteur de 4%, tous les records du monde officiels sur route féminins (depuis 2017) et masculins (depuis 2018), allant du 5 km au marathon, ont été battus [226].

Si les chaussures comptent naturellement dans l'équipement d'un coureur, d'autres technologies ont également fait leur apparition avec le temps dont les vêtements de compression avec des vertus de récupération (*i.e.*, diminution des courbatures en post effort, élimination des métabolites ou encore accélération de la réparation tissulaire *via* l'amélioration du flux sanguin) [225, 230-232] ; mais aussi des textiles respirants dotés de membranes technologiques permettant notamment le contrôle de l'humidité lors de l'effort [233].

Les outils connectés, comme les moniteurs et montres GPS (*Global Positioning System*) ou les cardiofréquencemètres, font également partie des grandes évolutions technologiques de la course à pied. Au-delà de fournir des informations détaillées sur la FC, la distance, la vitesse de course, la fréquence de la foulée, le dénivelé ; ces outils aident les athlètes à mieux juger et calibrer leur effort afin d'optimiser leur performance [223, 227].

Toujours dans une optique d'optimisation de la performance sont apparus différents procédés de récupération (*e.g.*, électrostimulation, cryothérapie, massages...) afin d'éliminer les métabolites ou favoriser le relâchement musculaire [53, 234, 235], différentes substances ergogènes afin de couvrir les besoins en fluides, en glucides et en électrolytes pendant ou suite à un effort physique, ou encore afin de prévenir du risque de blessure (*e.g.*, déshydratation, carences) [224, 236], des procédés permettant de simuler l'altitude (*e.g.*, tente et caisson hypoxiques) en appauvrissant l'air en oxygène afin d'enclencher l'érythropoïèse [219] , ou encore des procédés permettant de simuler de fortes chaleurs ou hauts

taux d'humidité (*i.e.*, *thermo training room*) afin de préparer les athlètes à supporter des conditions environnementales difficiles (*e.g.*, acclimatation des athlètes français dans le cadre de la préparation des JO de Tokyo).

Enfin, il semble également important de souligner l'évolution des surfaces du sol, de la piste en cendrée à la piste synthétique (*i.e.*, revêtements aux caractéristiques qui diffèrent) ayant permis notamment l'amélioration des références chronométriques en lien avec une restitution d'énergie accrue [176, 178].

Cependant, au-delà de toutes ces innovations autorisées (matérielles et équipements) dans le monde de la course à pied, un autre type d'innovation technologique d'ordre médical et biotechnologique, non approuvé par le système sportif, est également à souligner [124, 237]. Certaines substances (*e.g.*, érythropoïétine, stéroïdes anabolisants, hormone de croissance) développées initialement à des fins médicales pour le traitement de pathologies ont été détournées (*i.e.*, techniques de dopage) et utilisées afin d'optimiser (illégalement) les performances des athlètes en bonne santé leur procurant ainsi un avantage physiologique significatif (*e.g.*, augmentation du $\dot{V}O_{2max}$ suite à une prise d'érythropoïétine) [124, 237-241]. Le dopage, encore d'actualité et parfois institutionnalisé malgré la mise en place de processus de lutte antidopage, est une préoccupation mondiale qui touche de nombreuses pratiques sportives, dont la course à pied. Si le bénéfice du doute reste la règle, certaines progressions chronométriques laissent toutefois envisager de possibles abus de cette recherche biomédicale [16, 115, 122].

Par conséquent, au regard de toutes ces évolutions, il apparaît que l'expansion des connaissances (*e.g.*, science, médecine, technologie) a joué et jouera encore un rôle de plus en plus important dans l'optimisation et l'amélioration des performances sportives [3, 34, 115, 129].

3. Conclusion

Au fil des années, chaque discipline de course à pied s'est singularisée par sa référence chronométrique permettant de situer, au passage, le niveau de performance de l'athlète. Les épreuves de demi-fond et de fond, disputées en championnats ou hors championnats, ont fait l'objet de tentatives de records à plusieurs niveaux et sur plusieurs échelles de valeurs (e.g., record personnel ; record de *meeting* ; record de championnats ; record national, mondial, olympique) amenant ainsi à des évolutions de performances, à la fois structurelles (*i.e.*, liées aux tables de cotation et barèmes) et conjoncturelles (*i.e.*, amélioration du niveau). Ces évolutions de records du monde ou de France, pris pour exemples, sont des moyens d'évaluer le potentiel et les limites du corps humain, et reflètent également la manière dont la société a évolué, évolue et continuera d'évoluer au fil du temps. Si l'analyse historique des records a pour intérêt de comprendre les circonstances dans lesquelles les performances sont réalisées, en identifiant notamment au préalable les différents facteurs qui peuvent limiter ou optimiser ces performances sportives ; l'analyse historique des records a également pour intérêt de modéliser ces performances sportives afin de tenter de les prédire. Au-delà des modèles proposés pour analyser les records du monde en course à pied, d'autres modèles furent établis dans le but de pouvoir définir un temps sur une distance donnée, utile par exemple, pour prescrire des intensités de course à l'entraînement et progresser, ou encore détecter et orienter les athlètes au fort potentiel (*i.e.*, talent, capacités). Ces modèles de prédiction feront par conséquent l'objet du chapitre suivant.

CHAPITRE III. PRÉDICTION DE LA PERFORMANCE EN COURSE À PIED

Comme évoqué précédemment, les analyses des records du monde ont permis d'élaborer des modèles théoriques et d'étudier la progression des références chronométriques afin d'appréhender la performance en course à pied [29, 30, 74, 101, 107-114]. Au-delà des modèles utilisés pour l'analyse des performances à travers le temps, d'autres méthodes et modèles furent établis afin de pouvoir prédire (*i.e.*, réalisation d'un temps sur une distance donnée), optimiser la performance, et détecter le potentiel de l'athlète en course à pied.

La première section de ce chapitre est consacrée à une brève revue bibliographique des différentes méthodes ou outils destinés à prédire la performance ou le potentiel de performance en course à pied ; et la seconde section est destinée à présenter plus précisément les différents outils ayant fait l'objet des travaux menés durant ce doctorat.

1. Revue de littérature : Concepts et disciplines étudiés

Si la prédiction de performance en course à pied a été préalablement présentée comme étant l'action d'annoncer par avance la réalisation d'un temps de course sur une distance donnée, la détection de talent ou de potentiel de performance pourrait, quant à elle, être définie, de manière générale, comme étant la prédiction de la performance dans le temps *via* l'utilisation de références chronométriques, seules ou combinées avec des variables d'ordre physiologique par exemple.

**a. Détermination du profil et du potentiel de performance des coureurs
à l'aide de tests de laboratoire ou de terrain**

La détermination des profils de coureurs, en lien avec la connaissance des différents facteurs de performance du demi-fond et du fond, est essentielle. En effet, la connaissance du $\dot{V}O_{2max}$, de la VMA ou encore de la capacité d'endurance aérobie, par exemple, permet de construire et d'individualiser le programme d'entraînement des athlètes (*i.e.*, charge de travail) et d'identifier les progrès [8, 60]. Les valeurs de $\dot{V}O_{2max}$, de VMA et/ou de capacité d'endurance aérobie peuvent être obtenues par l'exécution de différents tests effectués en laboratoire (*i.e.*, protocole direct ou indirect sur ergomètre, sous le contrôle d'un médecin) ou sur le terrain (*e.g.*, en continu, à intensité progressive, en intermittent...) [8, 60] (Figure 24).

Par exemple, en laboratoire, pour déterminer la capacité d'endurance aérobie, les méthodes directes intègrent tous les indices permettant la représentation complète ou partielle de la relation puissance-durée (*i.e.*, intensité relative qui pourrait être maintenue ou endurée le plus longtemps possible, l'intensité relative la plus élevée pour une durée ou une distance déterminée), alors que les méthodes indirectes portent sur la détermination du seuil dit « anaérobie » (*i.e.*, méthodes qui ne mesurent la durée maximale ni l'intensité relative maximale pour une durée déterminée, mais qui sont censées refléter la capacité d'endurance) [140]. Les méthodes indirectes sont notamment basées sur une relation significative entre la fraction de $\dot{V}O_{2max}$ pour une épreuve donnée et d'autres variables physiologiques comme les paramètres ventilatoires, les concentrations de lactate sanguin ou la FC [140].

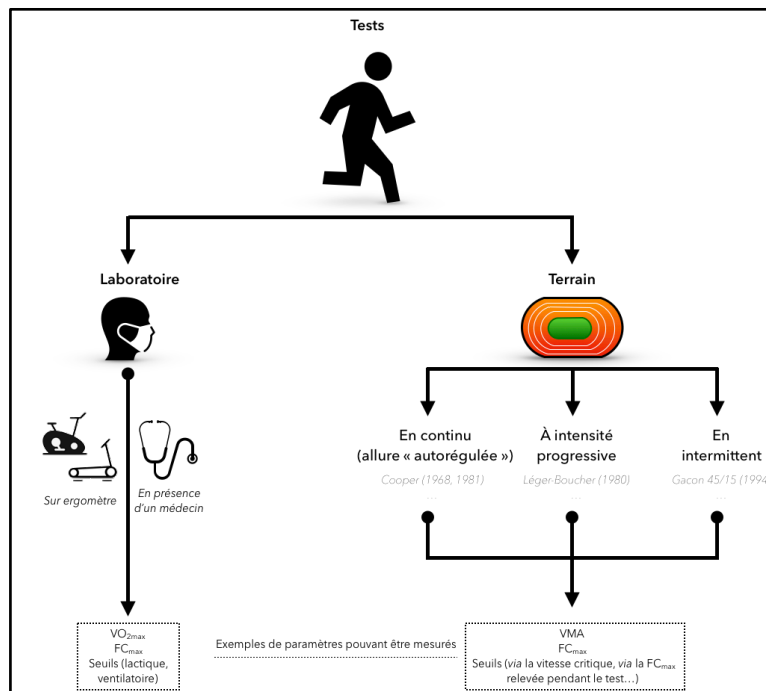


Figure 24. Exemples de tests de laboratoire ou de terrain

Suite à ces tests, la prédiction des performances, susceptibles d'être atteintes en course, peut alors être envisagée [242]. En effet, à partir de la connaissance du $\dot{V}O_{2max}$, il est possible de prédire une performance sur une distance donnée en appliquant, par exemple, le modèle de Margaria et al. [243] (*i.e.*, $m = 5(\dot{V}O_{2max} - 6) \times t + 5(\dot{V}O_{2max})$, où m est la distance maximale (en m) parcourue dans le temps t (en min) pour un $\dot{V}O_{2max}$ donné). En appliquant numériquement ce modèle sur une distance de demi-fond comme le 3 000 m, et en prenant pour exemple un athlète ayant un $\dot{V}O_{2max}$ de $65 \text{ ml} \cdot \text{min}^{-1} \cdot \text{kg}^{-1}$, le temps prédit sur cette distance est en l'occurrence de 8 min 57 s. Toutefois, Margaria et al. [243] ont précisé que l'équation du modèle pouvait être influencée par l'état nutritionnel du sujet, le coût énergétique de sa course, et sa capacité à maintenir son $\dot{V}O_{2max}$ tout au long de l'épreuve. À noter, également que si ce modèle peut permettre la prédiction d'un temps sur une distance donnée, il peut à l'inverse, estimer le $\dot{V}O_{2max}$ à partir d'une performance réalisée (à vitesse maximale). Cazorla [244] a toutefois souligné la limite potentielle de ce modèle de prédiction établis lors de courses sur tapis

roulant. La dépense énergétique théorique de la course sur tapis roulant considérée pour estimer le $\dot{V}O_{2\max}$ pourrait en effet ne pas être à sa juste valeur (*i.e.*, négligence de l'augmentation du $\dot{V}O_{2\max}$ due à la résistance de l'air lors d'une course hors laboratoire, et ceci d'autant plus la vitesse est élevée) et altérer les résultats de prédiction [245].

À partir de la VMA, il est également envisageable selon des tables de correspondance (*e.g.*, table de Mercier et Léger, table de Daniels), ou en s'appuyant sur des travaux scientifiques (ayant déterminé les pourcentages de VMA sollicités en fonction des différentes distances de course à pied) d'approximer les performances potentielles et d'orienter l'entraînement spécifique en fonction des performances visées (Tableaux 2 et 11) [8, 60, 84, 242, 246, 247]. Il a notamment été mis en évidence que les corrélations entre la VMA et les vitesses auxquelles ont été réalisées ces performances en compétition augmentent entre le 800 m et le 3 000 m (*i.e.*, $r = 0,72$, $r = 92$ et $r = 98$, pour les distances de 800 m, 1 500 m et 3 000 m, respectivement) et inversement diminuent progressivement au-delà au 5 000 m ($r = 0,98$, $r = 0,88$ et $r = 0,85$, respectivement), mais restent élevées dans tous les cas. De manière générale, ces corrélations indiquent que la distance de 3 000 m, est la distance la plus proche de la VMA (*i.e.*, 95 à 100%) à haut niveau, et que cette dernière peut éventuellement servir de critère pour déterminer la VMA chez des coureurs de demi-fond à condition qu'ils soient bien entraînés.

Tableau 11. Pourcentage de VMA sollicité en fonction de la distance de course en compétition (valeurs moyennes pour des coureurs relativement entraînés) d'après Cazorla et Leger [242]

Distance (m)	VMA (%)
800	120 à 125
1 500	101 à 111
3 000	95 à 100
5 000	90 à 95
10 000	85 à 90
20 000	80 à 88
Marathon	75 à 84

Cependant, si ces tests de laboratoire ou de terrain ont pour avantage de pouvoir prédire les performances sur une distance donnée, ces derniers présentent également quelques inconvénients. Les tests de laboratoire, généralement effectués sur tapis de course et nécessitant des outils de mesure sophistiqués, ont notamment pour désavantage d'être relativement onéreux et de devoir s'effectuer dans une structure spécialisée sous le contrôle d'un médecin. Quant aux tests de terrain, au-delà du fait que ces derniers sont moins précis lorsqu'ils sont basés sur des estimations, ils nécessitent respectivement, à allure progressive (e.g., Léger-Boucher, Vaméal) ou à allure autorégulée (e.g., Cooper, demi-Cooper), toute une mise en place matérielle (*i.e.*, plots et sifflet), ou encore la connaissance de la gestion d'allure de course par le coureur : un départ trop rapide ou trop lent peut se révéler contre-productif et biaiser le résultat. À noter également qu'en utilisant des systèmes métaboliques portables, permettant notamment la mesure des échanges gazeux, ces tests de terrain peuvent s'avérer relativement coûteux.

Le suivi et la prise en considération des séances d'entraînement (e.g., nombre d'entraînement par semaine, distance parcourue, intensité d'effort) peuvent également être des alternatives donnant une idée très approximative des références chronométriques que pourrait valoir un athlète sur une distance donnée [249, 250]. Cependant, au regard de l'aspect multifactoriel qui définit la performance en course à pied (*i.e.*, multiplicité des facteurs évoqués dans le chapitre précédent) [9, 18, 59, 70, 74], ce suivi des sessions d'entraînement peut s'avérer relativement limité pour appréhender la performance dans son ensemble. En effet, si la connaissance des différents facteurs de performance du demi-fond et fond tels que le $\dot{V}O_{2max}$, la VMA, la capacité d'endurance aérobie (quelle que soit la méthode utilisée), est importante pour déterminer le profil des coureurs [11, 14, 68, 70, 131, 251], la connaissance de ces paramètres ne suffit pas à orienter l'entraînement, déterminer le potentiel et à prédire les performances des athlètes.

La prise en considération d'autres facteurs tels que l'économie de course et les états et/ou traits psychologiques en compétition sont également déterminants et peuvent donner une bien meilleure évaluation des intensités d'entraînement et des performances réalisables sur d'autres distances [84].

b. Détermination du profil et du potentiel de performance des coureurs à l'aide des performances réalisées en compétition.

Le recours à des temps réalisés en compétition sur une distance donnée, apparaît plus représentatif de la capacité de performance des athlètes pour estimer la capacité sur d'autres distances, que l'évaluation d'intensités d'efforts réalisés lors de tests en laboratoire ou de terrain, ou lors du suivi de séances [84]. En effet, les références chronométriques établies lors de compétitions reflètent généralement les capacités maximales de l'athlète (e.g., le $\dot{V}O_{2max}$, l'économie de course, la capacité d'endurance, l'attitude, la résistance mentale...) et tout cela en une seule valeur : le temps mis pour parcourir la distance donnée [84].

En lien avec l'analyse des records du monde en course à pied, de nombreux auteurs ont élaboré des modèles ou outils afin de modéliser la relation entre le temps et la distance (ou la vitesse) et de prédire la performance en course à pied sur route ou sur piste [18, 26, 30, 31, 113, 252]. Les tables [84, 253], les équations mathématiques (e.g., logarithmiques, hyperboliques, exponentielles, linéaires...) [113, 131, 252, 254] incluant les concepts de vitesse critique [28, 29] ou de loi de puissance [30-32, 255], les nomogrammes [26, 27], et les algorithmes d'IA [256, 257] en sont des exemples notables (Figure 25).

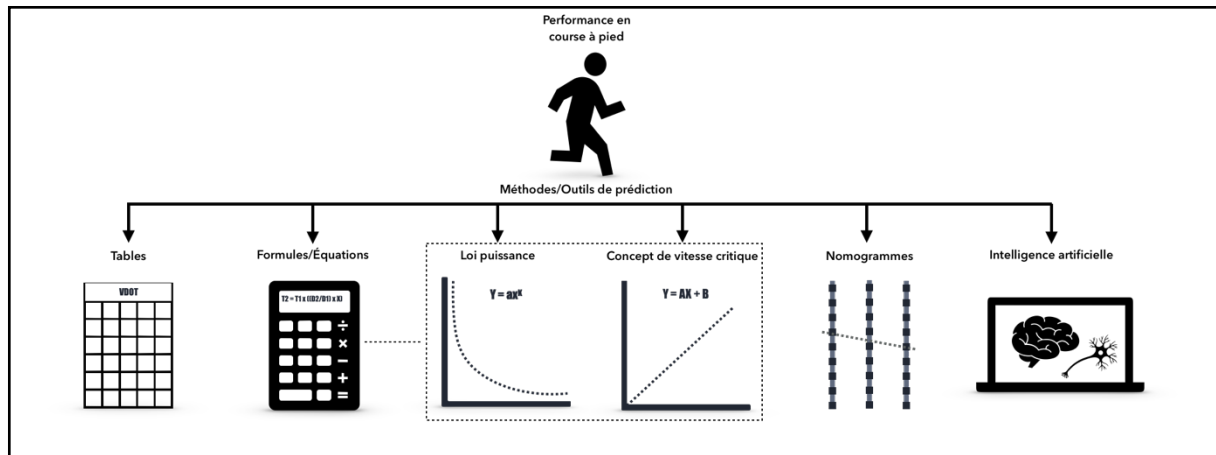


Figure 25. Méthodes de prédiction de performance en course à pied grâce à la compétition

La prédiction de la performance en course à pied peut être utile pour prescrire des intensités spécifiques d'entraînement, planifier les temps de passage en compétition, mais aussi détecter, par exemple, les futurs athlètes au fort potentiel [23-25].

De manière générale, les prédictions de performance en course à pied reposent sur le calcul théorique du temps de course en utilisant la ou les meilleures performances réalisées sur d'autres distances ; et en se basant sur une certaine équivalence entre les références chronométriques des différentes distances ayant été parcourues [258]. À titre d'exemple, le recours aux tables (e.g., temps sur marathon à partir du temps sur semi-marathon) [253] ou aux formules rapportant le temps réalisé pour parcourir une distance donnée peut permettre d'obtenir une prédiction sur la distance souhaitée (e.g., temps sur marathon = temps sur 10 km \times 4,76) [18, 258]. Cependant, comme cela a été évoqué précédemment, parmi les différentes approches proposées, certaines méthodes incluent, au-delà des références chronométriques, diverses variables d'ordre physiologique, anthropométrique, voire d'entraînement [84, 113, 131, 259, 260]. La formule de Riegel, par exemple, développée sur la base de records du monde de plusieurs disciplines d'endurance (e.g., course, marche, natation, ski de fond et cyclisme) permet de prédire une

performance en incluant notamment le facteur fatigue [113]. Elle se présente sous la forme de $t = ax^b$, où t est le temps ; x la distance ; a une constante dépendante des unités de mesure choisies qui n'a pas de valeur absolue particulière ; et b un exposant faisant référence au facteur fatigue qui conserve la même valeur, quelles que soient les unités choisies pour le temps et la distance incluse. Dans le cadre de la course à pied, au regard du fait que le paramètre de capacité d'endurance soit fixe dans la formule de Riegel (*i.e.*, loi de puissance avec un exposant fixe), la formule appliquée est la suivante :

$$t_2 = t_1 \times \frac{d_2^\beta}{d_1^\beta} \quad (1)$$

Par exemple, si un athlète parcourt le 10 km en 35 min, il devrait être capable de courir le semi-marathon en 1 h 17 min (*i.e.*, $t_{\text{semi-marathon}} = 35 \times \frac{21,1^{1,06}}{10} = 77 \text{ min}$).

Ainsi, à partir d'une référence chronométrique (*i.e.*, performance), où le temps (t_1) et la distance (d_1) sont connus, la formule permet d'estimer le temps de course (t_2) sur une nouvelle distance (d_2). Étant donné que le rapport entre les deux distances de course a été estimé à hauteur de 6-8% selon Riegel [113] pour des distances allant de 1 500 m au marathon, l'exposant β (calculé sur la base des records du monde) prend généralement la valeur de 1,06 pour les coureurs *seniors* voire 1,08 pour les coureurs élités [23, 113]. Cette valeur de l'exposant doit ainsi être adaptée selon les profils des coureurs, c'est-à-dire en fonction de leur capacité d'endurance afin de limiter les biais du modèle [261]. Les travaux de Smet d'Olbecke [261] ont indiqué, par exemple, que lorsqu'il s'agissait de prédire les performances sur marathon, en prenant une valeur d'exposant fixée à 1,06, 1,15 ou 1,20, que les erreurs de prédiction des performances obtenues étaient respectivement de 13,71%, 11,98% et 12,09%, montrant ainsi que le biais pouvait être compensé sur cette distance de fond en sélectionnant une valeur d'exposant adaptée (soit $\beta = 1,15$, en l'occurrence pour cet exemple).

Une autre approche de prédiction incluant également un seul paramètre est le modèle VDOT [84]. Le terme VDOT n'est autre que l'abréviation de V POINT (sur le V) (*i.e.*, *V DOT* en anglais) soit la valeur de $\dot{V}O_{2\max}$ à laquelle il est lié (*i.e.*, *V dot O₂*, exprimé en $\text{ml}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$). Plus précisément, il s'agit d'un $\dot{V}O_2$ spécifique sur une distance donnée en se servant d'une estimation du $\dot{V}O_{2\max}$ du coureur à partir d'une performance de terrain sans avoir le besoin de réaliser un test (*e.g.* Vaméal) [262]. Si ce modèle a été relativement peu discuté dans la littérature, il est cependant plus répandu dans le monde de l'entraînement où il est communément présenté sous la forme de tables (ou sous forme d'application « *VDOT Running Calculator* » dans le cadre du « *Run SMART Project* ») et utilisé dans les programmations pour établir des intensités de course [263] (Tableaux 12 et 13) [84]. Ces tables VDOT ont été élaborées à partir de données échelonnées sur plusieurs années chez de nombreux athlètes pratiquant la course à pied à différents niveaux en compétition. Les trois principales variables générées étaient le $\dot{V}O_{2\max}$, l'économie de course (avec au minimum quatre vitesses de course sous-maximales) et la fraction de $\dot{V}O_{2\max}$ (*i.e.*, $\frac{\dot{V}O_2}{\dot{V}O_{2\max}}$) respective à laquelle chaque athlète courait en compétition pour un ensemble de distances [84]. La Figure 26 ci-après, représente respectivement, les résultats qui ont été obtenus *via* des expressions mathématiques, dont la relation entre le $\dot{V}O_2$ et la vitesse (v) du coureur (*i.e.*, $\dot{V}O_2 = f_{O_2}(v) = 10,93548 \times v + 0,3744 \times v^2 - 4,60$), soit l'économie de course (à vitesse sous-maximale) (Figure 26A) ainsi que la fraction de $\dot{V}O_{2\max}$ associée aux durées des distances courues en compétition (*i.e.*, $\frac{\dot{V}O_2}{\dot{V}O_{2\max}} = f_{s.i}(t) = 0,2990e^{-0,003221t} + 0,1894e^{-0,0002130t} + 0,8$), soit l'intensité pouvant être maintenue pendant un temps (t) donné (Figure 26B) ont permis l'élaboration des tables VDOT [261]. Le rapport entre le $\dot{V}O_{2\max}$ et le CE est à l'origine du concept de vitesse à $\dot{V}O_{2\max}$ (*i.e.*, $v\dot{V}O_{2\max}$ soit la VMA). Cette $v\dot{V}O_{2\max}$ (VMA), est alors déterminée soit à partir de paramètres physiologiques ($\dot{V}O_{2\max}$ et CE), soit *via* la mesure de l'allure de course la plus élevée atteinte lors d'un effort à vitesse progressivement augmentée et poursuivie jusqu'à épuisement.

Par exemple, en utilisant la première formule représentant la Figure 26A le $\dot{V}O_2$, pour réaliser un 10 km en 37 min 30 s (soit $4,44 \text{ m.s}^{-1}$), est d'environ $51,4 \text{ ml.min}^{-1}.\text{kg}^{-1}$ (i.e., $\dot{V}O_2 = 10,93548 \times 4,44 + 0,3744 \times 4,44^2 - 4,60 = 51,4 \text{ ml.min}^{-1}.\text{kg}^{-1}$). Maintenant, si un coureur a effectué une course de 10 km en 37 min 30 s, la formule représentant la courbe de la Figure 26B indique que la fraction de VDOT $O_{2\text{max}}$ sur une course est de 0,92 (soit 92% $\dot{V}O_{2\text{max}}$). En conséquence, si le coût de la course est de $51,4 \text{ ml.min}^{-1}.\text{kg}^{-1}$ et que le coureur travaille à 92% $\dot{V}O_{2\text{max}}$ alors la valeur VDOT de ce coureur (pseudo $\dot{V}O_{2\text{max}}$) sera de $\frac{51,4}{0,92}$, soit environ 56.

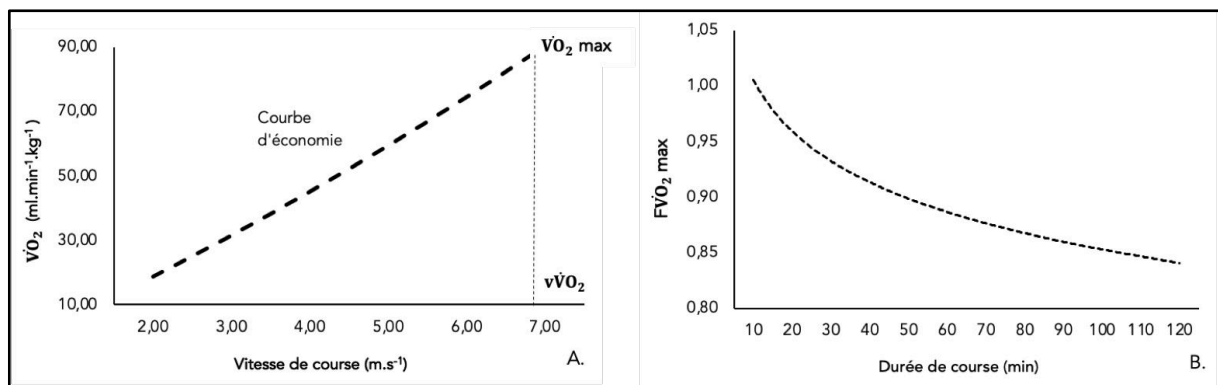


Figure 26. Illustration des deux expressions utilisées dans le modèle VDOT

Légende : Courbe d'économie moyenne et $v\dot{V}O_{2\text{max}}$ (vitesse minimale qui sollicite $\dot{V}O_{2\text{max}}$) associée à cette courbe et à $\dot{V}O_{2\text{max}}$ (A); et courbe de la relation entre la durée de course et la fraction de $\dot{V}O_{2\text{max}}$ ($F\dot{V}O_{2\text{max}}$), d'après Daniel's [84].

Dès lors que la valeur VDOT du coureur est estimée à partir d'une performance (représentative de son niveau actuel), les prédictions pour n'importe quelle distance peuvent être réalisées en utilisant les tables de VDOT (Tableaux 12 et 13).

Par exemple, si un coureur de 10 km a pour meilleure performance actuelle 37 min 30 s et souhaite s'orienter sur une distance de fond comme le marathon ; en se référant au Tableau 12 indiquant les valeurs VDOT associées à sa performance sur 10 km, soit une valeur VDOT de 56 ce coureur peut envisager un temps de 2 h 53 min 20 s. Puis en reportant cette valeur dans le Tableau 13 indiquant les intensités d'entraînement, ce coureur aura pour indication de s'entraîner à une allure spécifique de 4 min 6 s au km afin de préparer cette nouvelle distance.

Tableau 12. Valeurs de VDOT associées aux chronos sur les distances les plus répandues, d'après Daniel's [84]

VDOT	1 500 m	Mile	3 000 m	2 mile	5 000 m	10 km	15 km	Semi	Marathon	VDOT
30	8:30	9:11	17:56	19:19	30:40	63:46	98:14	2:21:04	4:49:17	30
31	8:15	8:55	17:27	18:48	29:51	62:03	95:36	2:17:21	4:41:57	31
32	8:02	8:41	16:59	18:18	29:05	60:26	93:07	2:13:49	4:34:59	32
33	7:49	8:27	16:33	17:50	28:21	58:54	90:45	2:10:27	4:28:22	33
34	7:37	8:14	16:09	17:24	27:39	57:26	88:30	2:07:16	4:22:03	34
35	7:25	8:01	15:45	16:58	27:00	56:03	86:22	2:04:13	4:16:03	35
36	7:14	7:49	15:23	16:34	26:22	54:44	84:20	2:01:19	4:10:19	36
37	7:04	7:38	15:01	16:11	25:46	53:29	82:24	1:58:34	4:04:50	37
38	6:54	7:27	14:41	15:49	25:12	52:17	80:33	1:55:55	3:59:35	38
39	6:44	7:17	14:21	15:29	24:39	51:09	78:47	1:53:24	3:54:34	39
40	6:35	7:07	14:03	15:08	24:08	50:03	77:06	1:50:59	3:49:45	40
41	6:27	6:58	13:45	14:49	23:38	49:01	75:29	1:48:40	3:45:09	41
42	6:19	6:49	13:28	14:31	23:09	48:01	73:56	1:46:27	3:40:43	42
43	6:11	6:41	13:11	14:13	22:41	47:04	72:27	1:44:20	3:36:28	43
44	6:03	6:32	12:55	13:56	22:15	46:09	71:02	1:42:17	3:32:23	44
45	5:56	6:25	12:40	13:40	21:50	45:16	69:40	1:40:20	3:28:26	45
46	5:49	6:17	12:26	13:25	21:25	44:25	68:22	1:38:27	3:24:39	46
47	5:42	6:10	12:12	13:10	21:02	43:36	67:06	1:36:38	3:21:00	47
48	5:36	6:03	11:58	12:55	20:39	42:50	65:53	1:34:53	3:17:29	48
49	5:30	5:56	11:45	12:41	20:18	42:04	64:44	1:33:12	3:14:06	49
50	5:24	5:50	11:33	12:28	19:57	41:21	63:36	1:31:35	3:10:49	50
51	5:18	5:44	11:21	12:15	19:36	40:39	62:31	1:30:02	3:07:39	51
52	5:13	5:38	11:09	12:02	19:17	39:59	61:29	1:28:31	3:04:36	52
53	5:07	5:32	10:58	11:50	18:58	39:20	60:28	1:27:04	3:01:39	53
54	5:02	5:27	10:47	11:39	18:40	38:42	59:30	1:25:40	2:58:47	54
55	4:57	5:21	10:37	11:28	18:22	38:06	58:33	1:24:18	2:56:01	55
56	4:53	5:16	10:27	11:17	18:05	37:31	57:39	1:23:00	2:53:20	56
57	4:48	5:11	10:17	11:06	17:49	36:57	56:46	1:21:43	2:50:45	57
58	4:44	5:06	10:08	10:56	17:33	36:24	55:55	1:20:30	2:48:14	58
59	4:39	5:02	9:58	10:46	17:17	35:52	55:06	1:19:18	2:45:47	59
60	4:35	4:57	9:50	10:37	17:03	35:22	54:18	1:18:09	2:43:25	60
61	4:31	4:53	9:41	10:27	16:48	34:52	53:32	1:17:02	2:41:08	61
62	4:27	4:49	9:33	10:18	16:34	34:23	52:47	1:15:57	2:38:54	62
63	4:24	4:45	9:25	10:10	16:20	33:55	52:03	1:14:54	2:36:44	63
64	4:20	4:41	9:17	10:01	16:07	33:28	51:21	1:13:53	2:34:38	64
65	4:16	4:37	9:09	9:53	15:54	33:01	50:40	1:12:53	2:32:35	65
66	4:13	4:33	9:02	9:45	15:42	32:35	50:00	1:11:56	2:30:36	66
67	4:10	4:30	8:55	9:37	15:29	32:11	49:22	1:11:00	2:28:40	67
68	4:06	4:26	8:48	9:30	15:18	31:46	48:44	1:10:05	2:26:47	68
69	4:03	4:23	8:41	9:23	15:06	31:23	48:08	1:09:12	2:24:57	69
70	4:00	4:19	8:34	9:16	14:55	31:00	47:32	1:08:21	2:23:10	70
71	3:57	4:16	8:28	9:09	14:44	30:38	46:58	1:07:31	2:21:26	71
72	3:54	4:13	8:22	9:02	14:33	30:16	46:24	1:06:42	2:19:44	72
73	3:52	4:10	8:16	8:55	14:23	29:55	45:51	1:05:54	2:18:05	73
74	3:49	4:07	8:10	8:49	14:13	29:34	45:19	1:05:08	2:16:29	74
75	3:46	4:04	8:04	8:43	14:03	29:14	44:48	1:04:23	2:14:55	75
76	3:44	4:02	7:58	8:37	13:54	28:55	44:18	1:03:39	2:13:23	76
77	3:41+	3:58+	7:53	8:31	13:44	28:36	43:49	1:02:56	2:11:54	77
78	3:38.8	3:56.2	7:48	8:25	13:35	28:17	43:20	1:02:15	2:10:27	78
79	3:36.5	3:53.7	7:43	8:20	13:26	27:59	42:52	1:01:34	2:09:02	79
80	3:34.2	3:51.2	7:37.5	8:14.2	13:17.8	27:41	42:25	1:00:54	2:07:38	80
81	3:31.9	3:48.7	7:32.5	8:08.9	13:09.3	27:24	41:58	1:00:15	2:06:17	81
82	3:29.7	3:46.4	7:27.7	8:03.7	13:01.1	27:07	41:32	:59:38	2:04:57	82
83	3:27.6	3:44.0	7:23.0	7:58.6	12:53.0	26:51	41:06	:59:01	2:03:40	83
84	3:25.5	3:41.8	7:18.5	7:53.6	12:45.2	26:34	40:42	:58:25	2:02:24	84
85	3:23.5	3:39.6	7:14.0	7:48.8	12:37.4	26:19	40:17	:57:50	2:01:10	85

Prédiction de la performance en course à pied : demi-fond et fond

Tableau 13. Intensités d'entraînement basées sur la valeur VDOT actuelle, d'après Daniel's [84]

VDOT	E/L (endurance)		M (allure marathon)		T (seuil)			I (intervalles)				R (répétitions)					VDOT
	Km	Mile	Km	Mile	400 m	Km	Mile	400 m	Km	1 200 m	1 600 m	200 m	300 m	400 m	600 m	800 m	
30	7:27-8:14	12:00-13:16	7:03	11:21	2:33	6:24	10:18	2:22	—	—	—	67	1:41	—	—	—	30
31	7:16-8:02	11:41-12:57	6:52	11:02	2:30	6:14	10:02	2:18	—	—	—	65	98	—	—	—	31
32	7:05-7:52	11:24-12:39	6:40	10:44	2:26	6:05	9:47	2:14	—	—	—	63	95	—	—	—	32
33	6:55-7:41	11:07-12:21	6:30	10:27	2:23	5:56	9:33	2:11	—	—	—	61	92	—	—	—	33
34	6:45-7:31	10:52-12:05	6:20	10:11	2:19	5:48	9:20	2:08	—	—	—	60	90	2:00	—	—	34
35	6:36-7:21	10:37-11:49	6:10	9:56	2:16	5:40	9:07	2:05	—	—	—	58	87	1:57	—	—	35
36	6:27-7:11	10:23-11:34	6:01	9:41	2:13	5:33	8:55	2:02	—	—	—	57	85	1:54	—	—	36
37	6:19-7:02	10:09-11:20	5:53	9:28	2:10	5:26	8:44	1:59	5:00	—	—	55	83	1:51	—	—	37
38	6:11-6:54	9:56-11:06	5:45	9:15	2:07	5:19	8:33	1:56	4:54	—	—	54	81	1:48	—	—	38
39	6:03-6:46	9:44-10:53	5:37	9:02	2:05	5:12	8:22	1:54	4:48	—	—	53	80	1:46	—	—	39
40	5:56-6:38	9:32-10:41	5:29	8:50	2:02	5:06	8:12	1:52	4:42	—	—	52	78	1:44	—	—	40
41	5:49-6:31	9:21-10:28	5:22	8:39	2:00	5:00	8:02	1:50	4:36	—	—	51	77	1:42	—	—	41
42	5:42-6:23	9:10-10:17	5:16	8:28	1:57	4:54	7:52	1:48	4:31	—	—	50	75	1:40	—	—	42
43	5:35-6:16	9:00-10:05	5:09	8:17	1:55	4:49	7:42	1:46	4:26	—	—	49	74	98	—	—	43
44	5:29-6:10	8:50-9:55	5:03	8:07	1:53	4:43	7:33	1:44	4:21	—	—	48	72	96	—	—	44
45	5:23-6:03	8:40-9:44	4:57	7:58	1:51	4:38	7:25	1:42	4:16	—	—	47	71	94	—	—	45
46	5:17-5:57	8:31-9:34	4:51	7:49	1:49	4:33	7:17	1:40	4:12	5:00	—	46	69	92	—	—	46
47	5:12-5:51	8:22-9:25	4:46	7:40	1:47	4:29	7:09	98	4:07	4:54	—	45	68	90	—	—	47
48	5:07-5:45	8:13-9:15	4:41	7:32	1:45	4:24	7:02	96	4:03	4:49	—	44	67	89	—	—	48
49	5:01-5:40	8:05-9:06	4:36	7:24	1:43	4:20	6:56	95	3:59	4:45	—	44	66	88	—	—	49
50	4:56-5:34	7:57-8:58	4:31	7:17	1:41	4:15	6:50	93	3:55	4:40	—	43	65	87	—	—	50
51	4:52-5:29	7:49-8:49	4:27	7:09	1:40	4:11	6:44	92	3:51	4:36	—	43	64	86	—	—	51
52	4:47-5:24	7:42-8:41	4:22	7:02	98	4:07	6:38	91	3:48	4:32	—	42	64	85	—	—	52
53	4:43-5:19	7:35-8:33	4:18	6:56	97	4:04	6:32	90	3:44	4:29	—	42	63	84	—	—	53
54	4:38-5:14	7:28-8:26	4:14	6:49	95	4:00	6:26	88	3:41	4:25	—	41	62	82	—	—	54
55	4:34-5:10	7:21-8:18	4:10	6:43	94	3:56	6:20	87	3:37	4:21	—	40	61	81	—	—	55
56	4:30-5:05	7:15-8:11	4:06	6:37	93	3:53	6:15	86	3:34	4:18	—	40	60	80	2:00	—	56
57	4:26-5:01	7:08-8:04	4:03	6:31	91	3:50	6:09	85	3:31	4:14	—	39	59	79	1:57	—	57
58	4:22-4:57	7:02-7:58	3:59	6:25	90	3:46	6:04	83	3:28	4:10	—	38	58	77	1:55	—	58
59	4:19-4:53	6:56-7:51	3:56	6:19	89	3:43	5:59	82	3:25	4:07	—	38	57	76	1:54	—	59
60	4:15-4:49	6:50-7:45	3:52	6:14	88	3:40	5:54	81	3:23	4:03	—	37	56	75	1:52	—	60
61	4:11-4:45	6:45-7:39	3:49	6:09	86	3:37	5:50	80	3:20	4:00	—	37	55	74	1:51	—	61
62	4:08-4:41	6:39-7:33	3:46	6:04	85	3:34	5:45	79	3:17	3:57	—	36	54	73	1:49	—	62

Toutefois, si certains outils ou modèles à un paramètre, comme la formule de Riegel ou le modèle VDOT, ont pour avantage de n'utiliser qu'une seule distance de course pour prédire une performance, ces derniers peuvent également présenter quelques limites en ne se basant que sur une seule performance, et générer par conséquent des marges d'erreur relativement importantes [261]. Scudamore et al. [264] ont notamment étudié la précision du modèle VDOT chez de jeunes athlètes ainsi que chez des coureurs amateurs, et ont indiqué que le « *VDOT Running calculator* » sous-estimait significativement le $\dot{V}O_{2\max}$.

À la différence des modèles à un paramètre, les modèles ou outils, incluant davantage de distances et/ou de paramètres (e.g., physiologiques, anthropométriques, d'entraînement), tels que les nomogrammes [24-26, 265], la loi puissance [17, 30-32, 266, 267], les modèles de type hyperbolique [101, 108, 114, 268], exponentiel [153] et logarithmique [112], ou encore les algorithmes d'IA [23, 256, 257] permettent généralement des prédictions avec une précision accrue [23, 26, 31, 261]. Cependant, au regard des nombreux paramètres (e.g., $\dot{V}O_{2\max}$, capacité d'endurance aérobie, IMC) pouvant intégrer les modèles, il en ressort que la plupart de ces approches sont généralement plus complexes à utiliser que les modèles à un paramètre, notamment lorsque le nombre de paramètres s'accroît [31, 261, 269]. Si ces différentes approches ont fait l'objet d'études dans la littérature afin d'être validées et d'être comparées sur différentes distances de course à pied (e.g., demi-fond et fond), il apparaît toutefois que la plupart des travaux menés se sont majoritairement intéressés à des athlètes de niveau élite et de sexe masculin. De ce fait, en lien avec les travaux menés durant ce doctorat, certaines de ces approches sont présentées plus explicitement dans la section suivante.

2. Présentation des modèles/outils testés pour la prédiction de la performance en demi-fond

a. Nomogrammes

Depuis le début du XX^{ème} siècle, de nombreux auteurs ont modélisé les records du monde en course à pied afin de prédire les performances des athlètes. Les analyses de ces données ont suggéré que la prédiction de la performance sur une distance donnée à partir de performances récentes sur deux autres distances différentes était possible, mais souvent au travers de méthodes mathématiques complexes. Pour faciliter l'accès de la prédiction au grand public (e.g., athlètes, entraîneurs), des outils relativement simples d'utilisation ne nécessitant aucun calcul tels que les « nomogrammes » de Mercier et al. [26] et de Vandewalle [27] ont été développés à partir de connaissances empiriques ou à partir d'un modèle logarithmique, respectivement.

i. De Mercier et al. [26]

Le nomogramme de Mercier et al. [26] est un outil empirique (i.e., développé à partir de connaissances de terrain) qui présente de nombreux avantages puisqu'il peut être utilisé pour déterminer la spécialité d'un coureur, estimer le $\dot{V}O_{2max}$, évaluer la capacité d'endurance aérobie et prédire les performances en course à pied [26]. Le nomogramme, à la différence de formules complexes ou de pages de tables, est un outil sous forme de figure permettant de prédire une performance en course de demi-fond et de fond sur toute une série de distances (allant du 3 000 m au marathon) à partir d'autres performances au moyen d'une simple règle (Figures 27 et 28). Les prédictions obtenues, par interpolation ou extrapolation, illustrent notamment le potentiel actuel de l'athlète. L'interpolation est la prédiction d'une

performance sur une distance située entre deux autres performances de distances différentes ; alors que l'extrapolation est la prédiction d'une performance sur une distance plus grande ou plus petite que les deux autres performances de distances différentes (Figure 28).

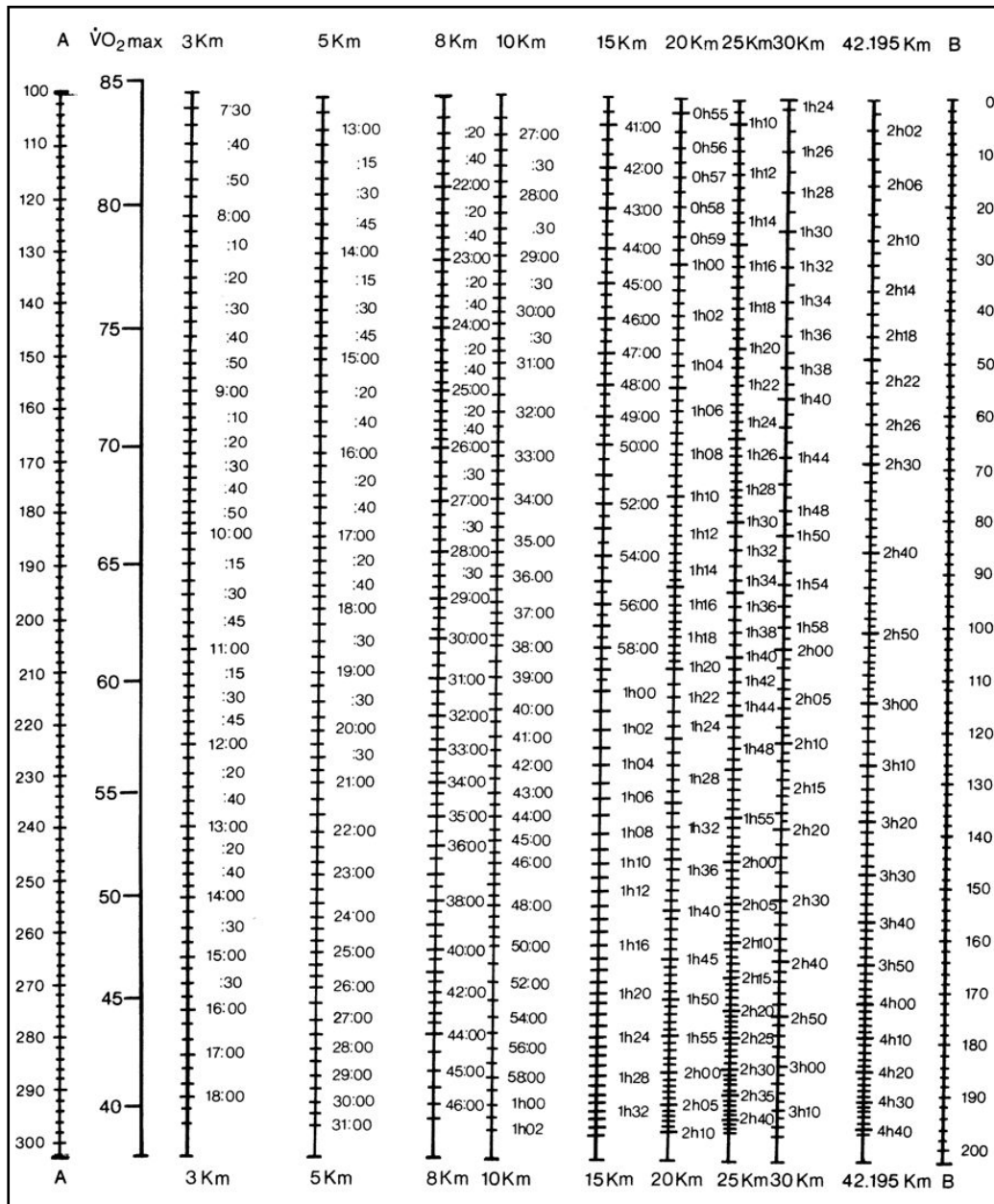


Figure 27. Nomogramme de Mercier et al. [26]

Légende : La droite reliant deux performances permet de prédire une performance sur une troisième distance ainsi qu'un indice de la capacité d'endurance aérobie (en soustrayant la valeur obtenue sur l'échelle B à celle relevée sur l'échelle A). Le $\dot{V}O_{2max}$ s'obtient par la droite horizontale passant par la performance prédite (ou réelle) au 3 000 m.

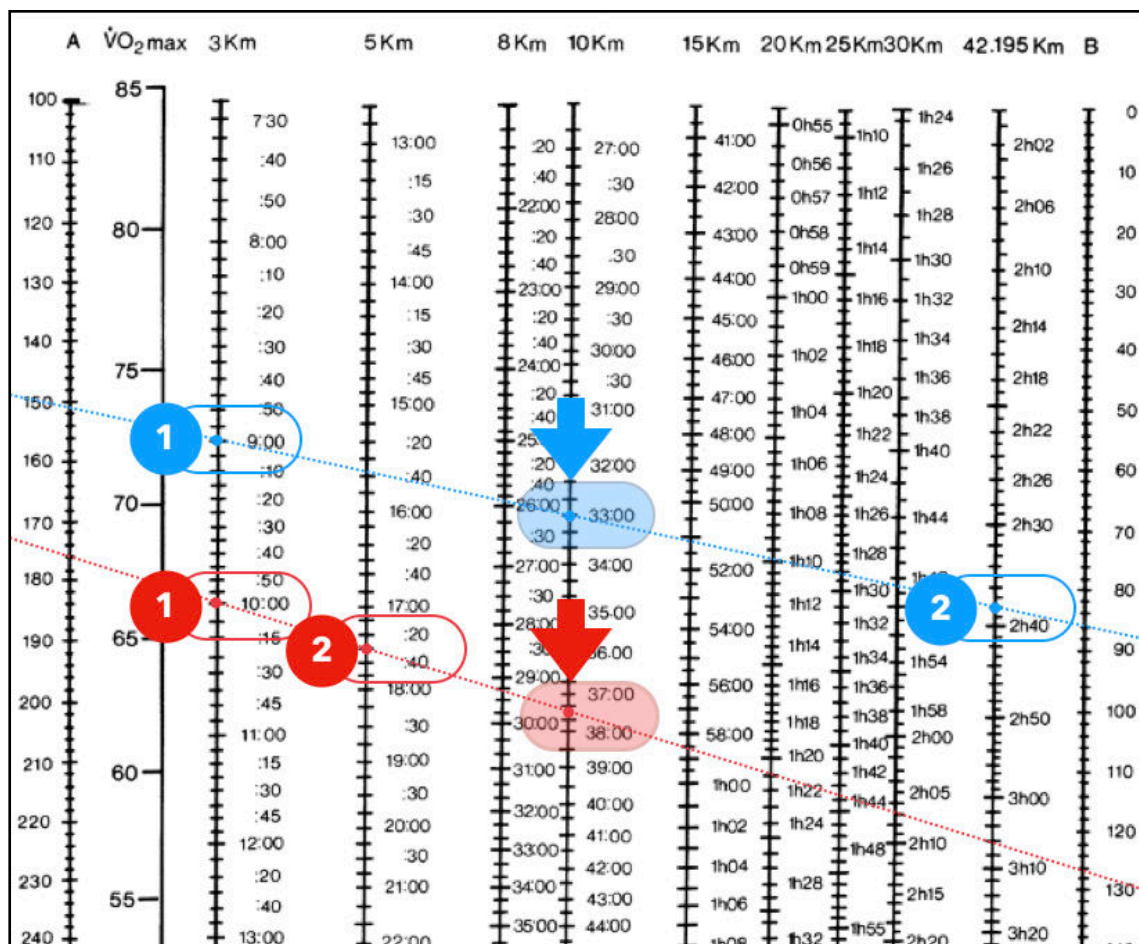


Figure 28. Exemples de prédiction de la performance par interpolation (en bleu) et extrapolation (en rouge) avec le nomogramme de Mercier et al. [26]

Note : Les chiffres 1 et 2 faisant référence aux deux performances de référence.

Pour prédire une performance en course à pied sur 10 000 m à partir du nomogramme de Mercier et al. [26], par interpolation ou extrapolation, il suffit pour chaque méthode de placer un repère sur l'échelle des distances correspondantes (*i.e.*, deux performances différant de la performance à prédire), de relier les deux références chronométriques (*i.e.*, points de repère) par un trait au moyen d'une règle, et enfin de reporter la valeur indiquée par l'échelle de la distance prédite. À titre d'exemples, par interpolation, si un athlète parcourt le 3 000 m en 9 min et le marathon en 2 h 38 min, il devrait être capable de courir le 10 000 m en 33 min ; par extrapolation, si un athlète réalise 10 min sur le 3 000 m et 17 min 30 s sur le 5 000 m, son potentiel de performance sur 10 000 m serait de 37 min 25 s (Figure 28) [26].

Cette approche prédictive a fait l'objet de travaux sur différentes distances, en fonction du sexe et/ou du niveau de performance [24, 25, 265, 270]. Si, les résultats ont montré que le nomogramme de Mercier et al. [26] fournit des prédictions valides et précises pour des performances de course à pied sur route (*i.e.*, 10 km, 20 km et marathon) [265], plusieurs études ont également confirmé ces résultats pour les performances sur piste pour les courses de demi-fond et de fond (*i.e.*, 3 000 m, 5 000 m et 10 000 m) chez les coureurs masculins ou féminins [270, 24, 25].

Toutefois, il est important de souligner le fait que le degré de précision des prédictions peut être dépendant de plusieurs facteurs dont : le positionnement des repères sur les échelles pour tracer le trait (*i.e.*, erreur humaine), les conditions de réalisation des deux performances de référence pour la prédiction (*i.e.*, laps de temps entre les deux performances réalisées, conditions environnementales), la spécialité athlétique du coureur, la distance séparant deux performances utilisées pour la prédiction, le type de méthode utilisée (*i.e.*, interpolation ou extrapolation) [24-26, 265, 270]. En effet, la comparaison, par exemple, de l'ensemble des résultats de prédiction, a démontré une précision accrue lorsque les performances sont prédites par interpolation plutôt que par extrapolation [24-26, 265, 270]. Ce degré de précision des prédictions pourrait notamment être expliqué physiologiquement par le fait que la prédiction de performances par interpolation plutôt que par extrapolation de courses de distance intermédiaire (*i.e.*, délimitée par des distances plus courte et plus longue) permet de prendre davantage en considération certains paramètres tels que le $\dot{V}O_{2max}$ (l'un des principaux prédicteurs de performances en demi-fond court comme le 3 000 m), l'économie de course et la capacité d'endurance aérobie (importance accrue avec l'augmentation la distance comme sur le 10 000 m, par exemple) [11, 251].

Enfin, bien que cela n'ait pas été exploité dans le cadre des travaux de ce doctorat, le nomogramme de Mercier et al. [26] permet également la prédiction du $\dot{V}O_{2\max}$ et de la capacité d'endurance aérobie des athlètes. Ces informations peuvent être intéressantes pour les athlètes et les entraîneurs, puisque les connaître permet de déterminer plus facilement les priorités d'entraînement et de construire la périodisation de la saison en conséquence.

ii. De Vandewalle [27]

Les nomogrammes de Vandewalle [27] ont été développés sur le modèle logarithmique de Péronnet et Thibault [112]. Ce modèle de Péronnet et Thibault [112] prend en considération les contributions des métabolismes aérobie et anaérobie pour la production totale d'énergie en fonction de la durée de la course. En partant du principe qu'un coureur n'est pas capable de maintenir une puissance aérobie sur un laps de temps infini, Péronnet et Thibault [112] ont supposé que la puissance aérobie maximale pouvait être maintenue à hauteur de 7 min (*i.e.*, 420 s). Ces derniers ont alors proposé un modèle exprimant la vitesse maximale (S) pouvant être maintenue, et diminuant avec le logarithme du temps total d'effort (t) :

$$S = mas - x \cdot \ln \left(\frac{t}{420} \right) \quad (2)$$

où mas est la VMA correspondant à la vitesse maximale pouvant être maintenue pendant un effort complet de 7 min, et x est un indice de la capacité d'endurance. Certains chercheurs ont tenté de modéliser cette capacité d'endurance [251, 271, 272], soit la capacité d'utiliser une fraction importante de $\dot{V}O_{2\max}$ (*i.e.*, F en % de $\dot{V}O_{2\max}$), plutôt que d'envisager la diminution de la vitesse de course en fonction de la durée et ainsi calculer un indice de la capacité d'endurance aérobie [251].

Toutefois, le modèle fournit une relation vitesse-temps qui ne peut pas être exprimée en fonction de la distance. Il est, par conséquent, impossible de prédire

la valeur de t correspondant à une distance donnée sans avoir recours à des programmes de calcul (e.g., algorithme), d'où le développement et l'utilisation de nomogrammes (présentés ci-après, Figures 29 à 31).

Les nomogrammes de Vandewalle [27] sont des diagrammes bidimensionnels conçus pour obtenir graphiquement une fonction sans utiliser de formule mathématique complexe (Figures 29 à 31). Ils peuvent être utilisés pour estimer la VMA, la capacité d'endurance aérobie et prédire les performances individuelles en demi-fond et fond chez des coureurs de niveau élite [27]. Cependant, bien que les nomogrammes de Vandewalle [27] soient construits de manière similaire, le nombre de distances et le nombre de graduations utilisées pour prédire les performances (pour une même distance) varient selon le nomogramme utilisé (Figure 32).

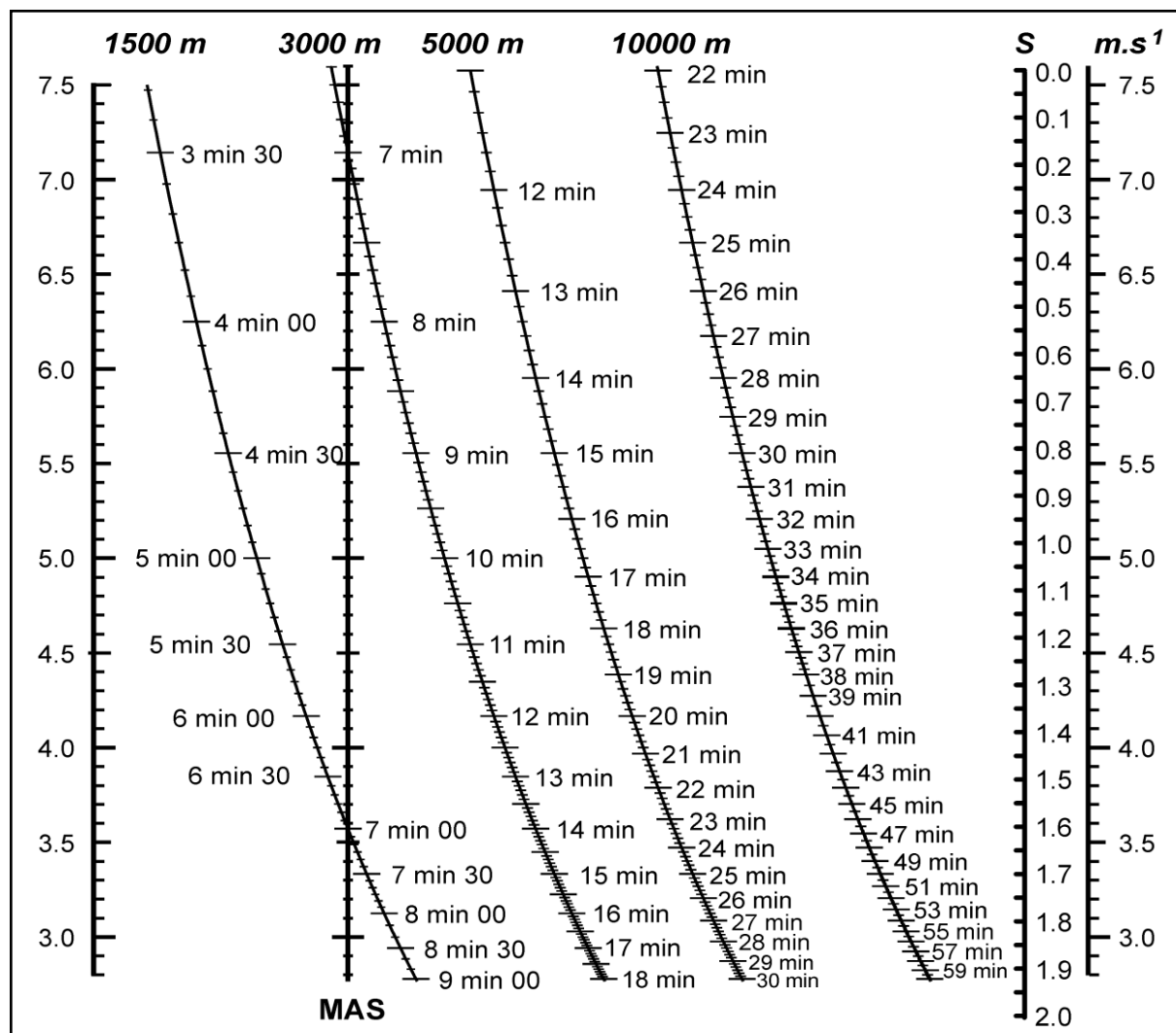


Figure 29. Premier nomogramme de Vandewalle [27]

Note : Nomogramme des performances correspondant aux distances de 1 500 m, 3 000 m, 5 000 m et 10 000 m, selon le modèle de Péronnet et Thibault [112].

La ligne verticale correspond à *tlim*, égal à 420 s (7 min), et les deux échelles (de 2,8 à 7,5 m.s⁻¹) permettent l'estimation de « MAS » (i.e., VMA). L'échelle « S » nommée « Slope » (de 0 à 2,0) à droite du graphique estime la capacité d'endurance (E) [27].

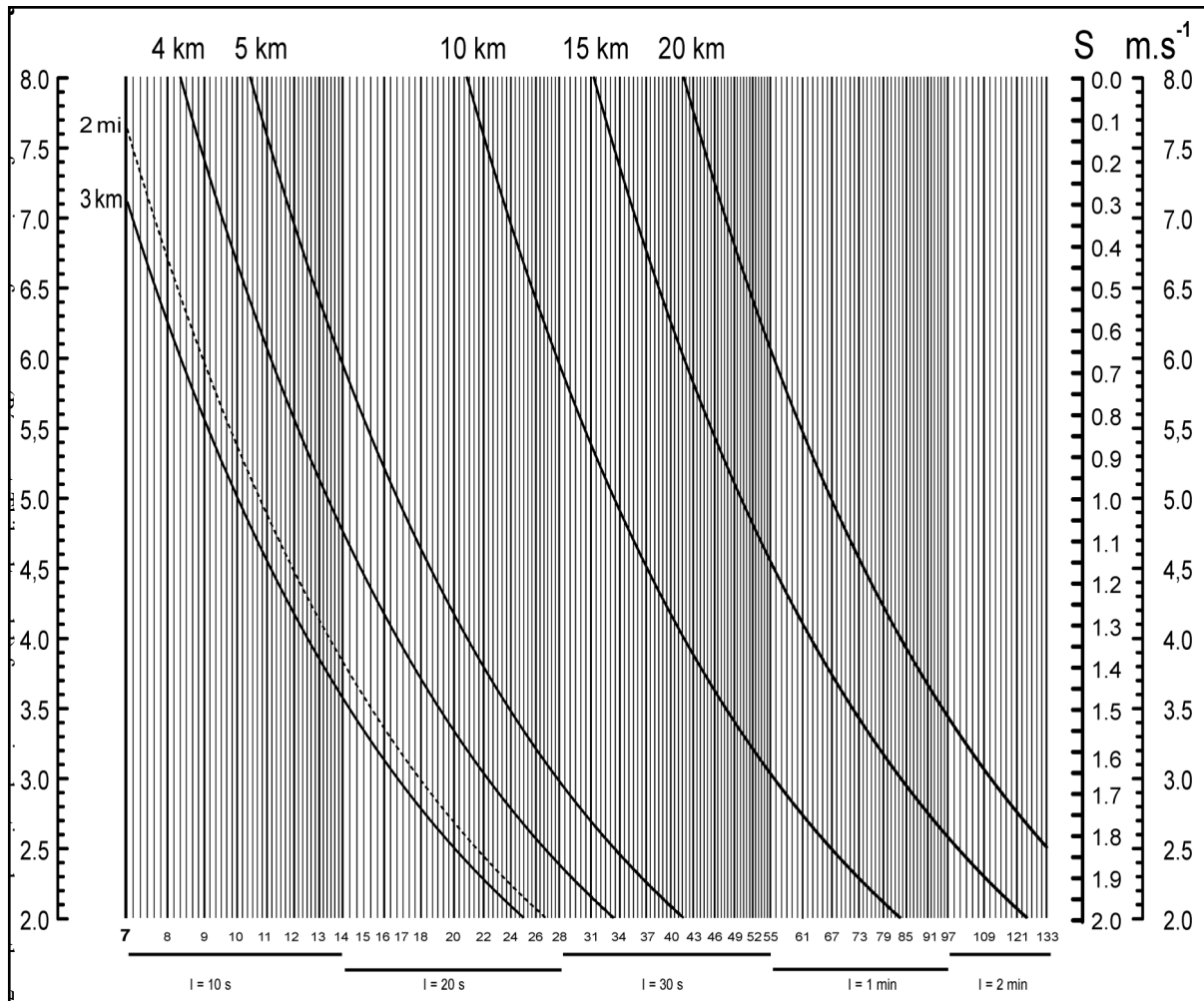


Figure 30. Second nomogramme de Vandewalle [27]

Note : Nomogramme des performances de course selon le modèle de Péronnet et Thibault [112] pour des durées d'exercice supérieures à 7 min. l = intervalles entre les lignes verticales (de 5 s à 2 min). La courbe en pointillés correspond aux performances sur 2 miles.

Prédiction de la performance en course à pied : demi-fond et fond

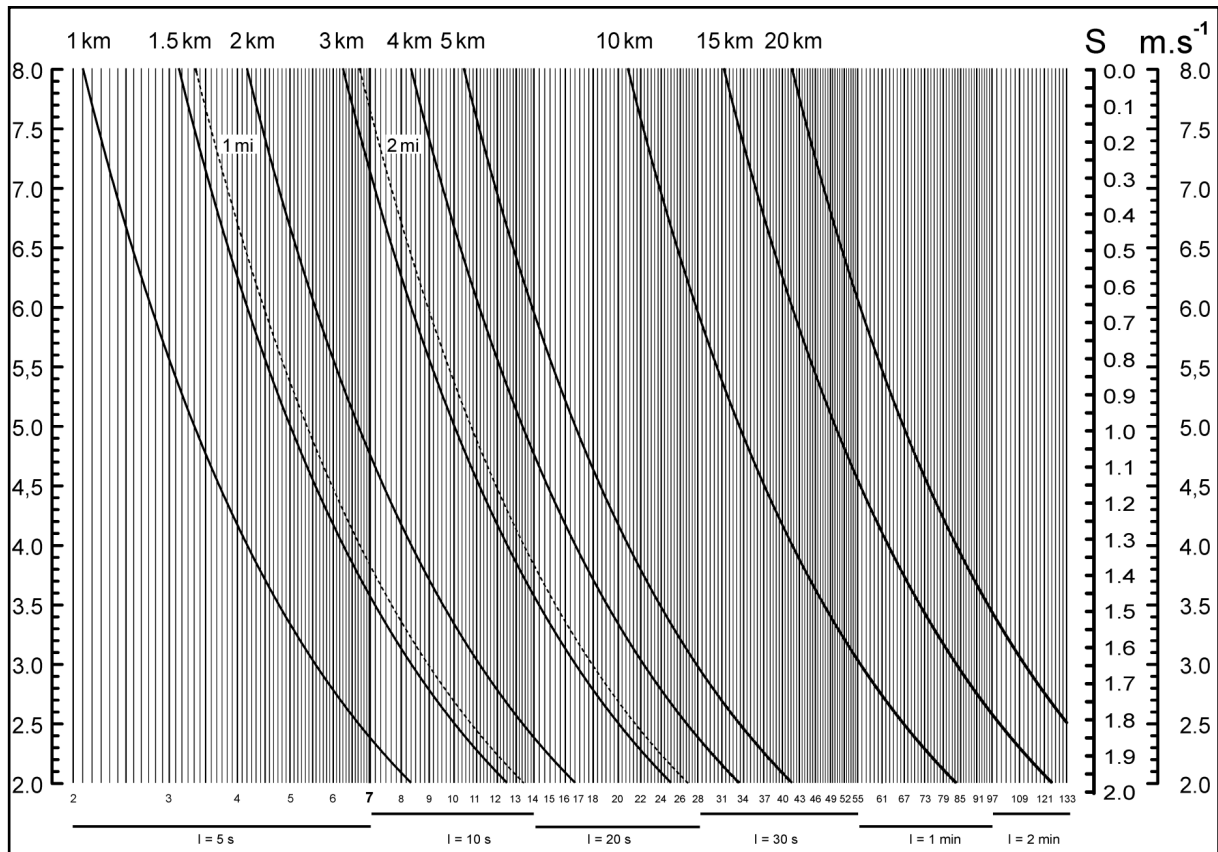


Figure 31. Troisième nomogramme de Vandewalle [27]

Note : Nomogramme des performances de course selon le modèle de Péronnet et Thibault [112] pour des distances allant de 1 000 m à 20 000 m. l = intervalles entre les lignes verticales (de 5 s à 2 min) Les courbes en pointillés correspondent aux performances sur 1 et 2 miles.

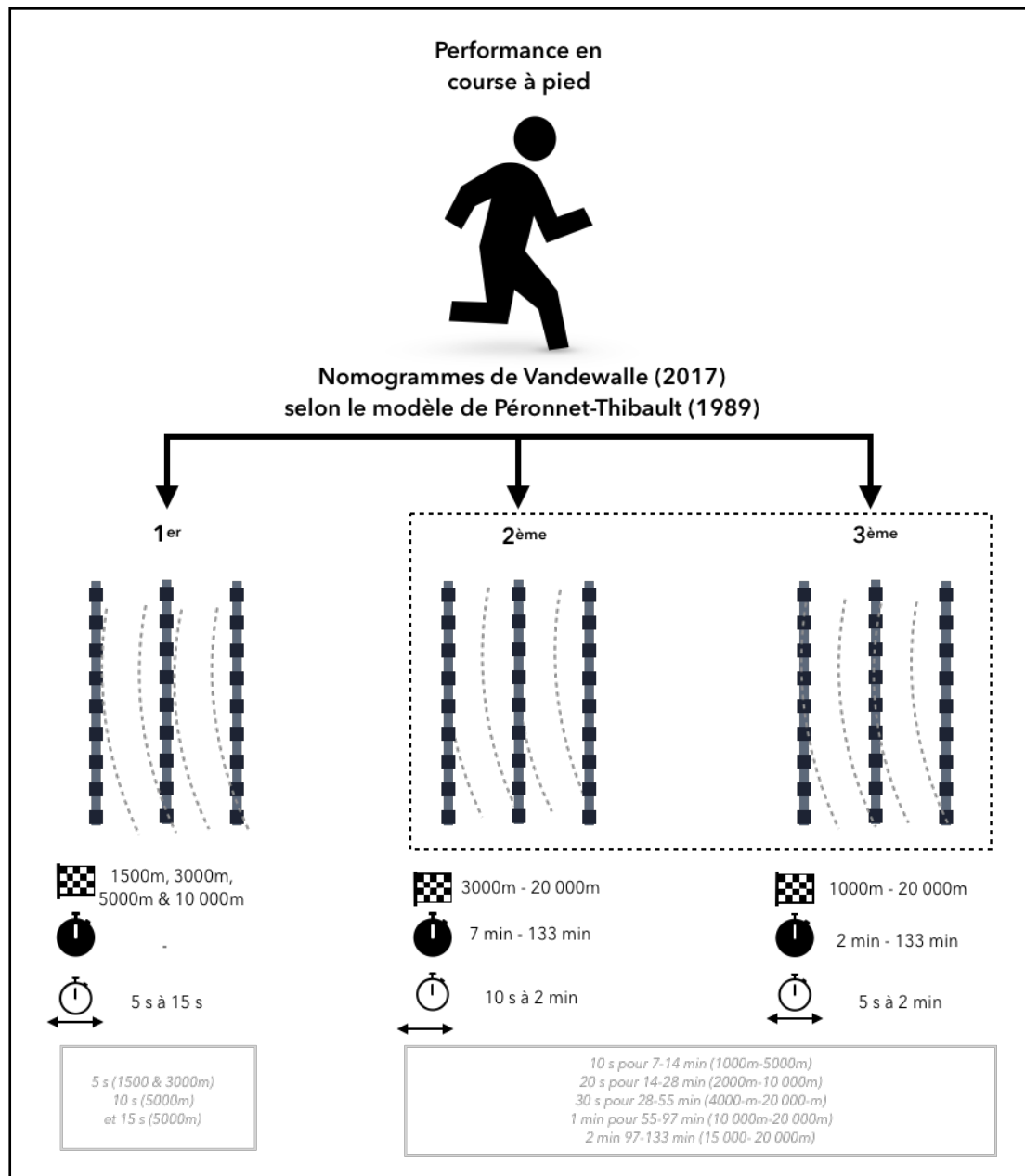


Figure 32. Caractéristiques des nomogrammes de Vandewalle [27]

Le premier nomogramme peut être utilisé pour prédire des performances sur piste sur quatre distances seulement dont le 1 500 m, 3 000 m, 5 000 m et 10 000 m (Figure 29) ; alors que le second et le troisième nomogramme (Figures 30 et 31), identiques dans leur échelle de temps et en graphisme, permettent de réaliser des estimations sur un plus grand nombre de distances, 3 000 m-20 000 m et 1 000 m-20 000 m, pour les second et troisième nomogrammes, respectivement (Figures 30 et 31). Toutefois, au-delà du nombre de distances incluses (*i.e.*, panel de distances)

et du graphisme qui diffèrent entre le premier type (*i.e.*, premier nomogramme) et le second type de nomogrammes (*i.e.*, deuxième et troisième nomogrammes), les échelles de temps utilisées pour prédire les performances pour une même distance varient (*i.e.*, nombre de graduations pour un temps ou une distance donnée) (Figures 30 à 32).

À titre d'exemples, le premier nomogramme de Vandewalle [27] a été conçu à partir de graduations de 5 s, 10 s et 15 s, respectivement, pour les distances de 3 000 m, 5 000 m et 10 000 m ; alors que le second et le troisième nomogramme ont été développés à partir de graduations variables de 10 s (lorsque la performance est comprise entre 7 et 14 min) ou de 20 s (lorsque la performance est comprise entre 14 et 28 min) pour les distances 3 000 m et 5 000 m ; de graduations de 30 s (lorsque la performance est comprise entre 28 et 55 min) pour 5 000 m et 10 000 m ; et de graduations de 1 min sur 10 000 m (lorsque la performance est comprise entre 55 et 97 min). À noter qu'au regard, des distances qui définissent chaque nomogramme et du fait que les échelles pour les second et troisième nomogrammes soient identiques, seuls les premier et deuxième nomogrammes ont fait l'objet d'analyses dans les travaux menés durant ce doctorat (Figures 29 et 30).

Comme pour le nomogramme Mercier et al. [26], pour prédire une performance par interpolation ou extrapolation il suffit de reporter les points correspondant aux performances sur les deux distances différentes ayant déjà été réalisées, et de lier ces deux repères par une ligne au moyen d'une règle (Figures 33 et 34).

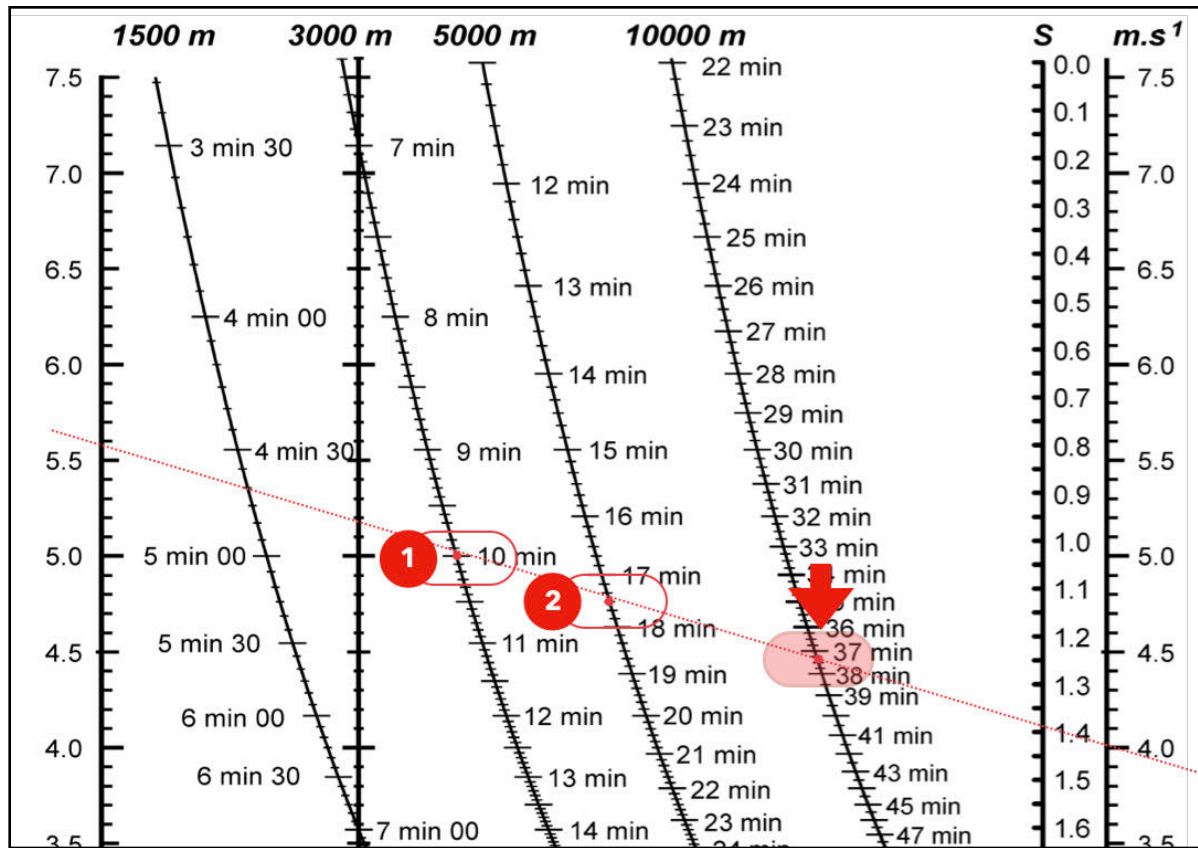


Figure 33. Exemples de prédiction de la performance sur 10 000 m par extrapolation avec le premier nomogramme de Vandewalle [27]

Note : Les chiffres 1 et 2 faisant référence aux deux performances de référence.

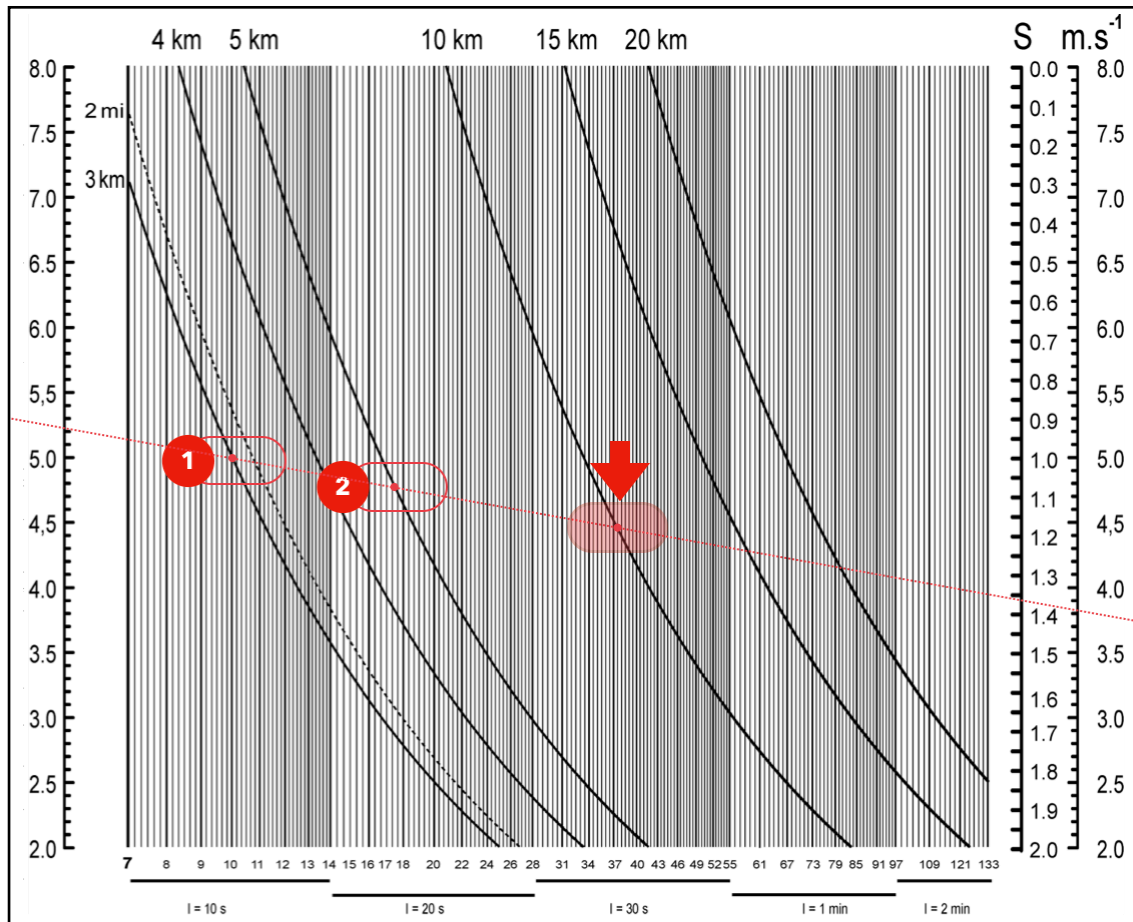


Figure 34. Exemples de prédiction de la performance sur 10 000 m par extrapolation avec le second nomogramme de Vandewalle [27]

Note : Les chiffres 1 et 2 faisant référence aux deux performances de référence.

En reprenant le même exemple de prédiction par extrapolation que pour le nomogramme de Mercier et al. [26], si un athlète réalise 10 min sur le 3 000 m et 17 min 30 s sur le 5 000 m, son potentiel de performance sur 10 000 m est estimé à 37 min 25 s selon le premier nomogramme de Vandewalle [27] (Figure 33), et estimé à 37 min 30 s selon le second nomogramme [27] (Figure 34).

En outre, pour les mêmes raisons qu'évoquées précédemment pour le nomogramme de Mercier et al. [26], les prédictions réalisées avec les nomogrammes de Vandewalle [27] peuvent être limitées par l'erreur humaine, les conditions de réalisation des deux performances support de prédiction, la discipline

du coureur, la distance séparant les deux performances utilisées pour la prédiction, et le type de méthode utilisée (*i.e.*, interpolation ou extrapolation).

Enfin, bien que Vandewalle ait confirmé la validité et la précision de ces nomogrammes, cette approche de prédiction n'a fait l'objet, à notre connaissance que d'une seule étude [27]. Ces nomogrammes n'ont été testés qu'à partir de performances de coureurs masculins de niveau mondial. Ainsi, il apparaît que leur validité reste à confirmer pour les performances sur piste de coureurs de niveau plus modeste indépendamment du sexe.

b. Concept de vitesse critique

Comme évoqué précédemment, la modélisation de la relation entre le temps et la distance (ou la vitesse) de course est un moyen courant et relativement facile pour prescrire des intensités d'entraînement, déterminer des stratégies de course ou encore prédire la performance en course à pied [31, 107, 270, 273].

Le concept de vitesse critique (ou puissance critique), appliqué sur la base des records du monde (en course à pied, natation, cyclisme et ski de fond), est la pente de la droite exprimant l'évolution du temps limite en fonction de la distance limite [109]. En d'autres termes, cette pente de la relation linéaire, entre le temps record et la distance de compétition, n'est autre que la représentation du record d'un athlète sur un panel de distances de courses.

Si plusieurs modèles, basés sur le concept de vitesse critique, ont été proposés *via* l'application des records du monde en course à pied notamment [29, 101, 114, 273], le modèle linéaire distance-temps original (MLDT), est représenté par l'équation suivante [109]:

$$D_{lim} = b \cdot T_{lim} + a \quad (3)$$

où D_{lim} est la distance maximale (en m) qui pourrait être parcourue en utilisant à la fois les réserves en oxygène (*i.e.*, métabolisme aérobie, soit la voie de phosphorylation oxydative) et l'énergie fournie par le métabolisme anaérobie (*i.e.*, voie glycolytique et voie des phosphagènes) ; b est la vitesse critique (V_{crit}), soit la vitesse qui peut en théorie être maintenue indéfiniment et à laquelle le taux maximal de synthèse des réserves par le métabolisme aérobie est atteint ; a est la distance maximale (en m) qui peut être parcourue avec le métabolisme anaérobie (*i.e.*, capacité de distance en anaérobie (CDA), qui correspond à l'ordonnée à l'origine du modèle) ; et T_{lim} est le temps jusqu'à épuisement (en min) pour une distance donnée (D_{lim}).

De cette première équation (3), l'équation hyperbolique finale obtenue pour la prédiction de performance, est la suivante :

$$T_{lim} = \frac{(D_{lim}-CDA)}{V_{crit}} \quad (4)$$

Au même titre que pour les nomogrammes, ce MLDT peut être utilisé pour prédire une performance sur une distance de course à pied à partir de deux autres performances. Par exemple, si un athlète court le 3 000 m en 10 min et le 5 000 m en 17 min 30 s, les valeurs de V_{crit} et de CDA sont de 267 m.min⁻¹ (*i.e.*, $V_{crit} = \frac{5\,000-3\,000}{17,5-10}$) et 333 m, respectivement. Par conséquent, en appliquant l'équation (4), la performance prédite sur 10 000 m est de 36 min 12 s (*i.e.*, $T_{lim} = \frac{10\,000-333}{267} = 36,21$ min).

Les modélisations de la relation linéaire distance-temps ont fait l'objet de nombreux travaux et s'il a été démontré qu'elles permettent de prédire des performances relativement précises en course à pied dans les épreuves de courtes (*e.g.*, 800 m) [274] voire de moyennes distances (*e.g.*, 3 000, 5 000 m voire 10 000 m) [270] ; ces modélisations peuvent s'avérer limitées dans la prédiction des performances de longues distances (ou durées) en course à pied (*e.g.*, 1 h) [28].

En effet, le MLDT, modélisé par l'équation (3) qui n'est autre qu'une fonction affine (*i.e.*, $y = a.x + b$) [29], affiche une relation linéaire distance-temps lorsque les valeurs de temps limites sont comprises entre 3,5 et 35 min environ [114, 153, 275]. En deçà et au-delà de cette plage de temps (où a et b en sont dépendants [109, 276, 277]), la pente de la relation distance-temps n'est, par conséquent, plus linéaire [109, 275]. La prédiction de performances sur de longues distances à l'aide de ce modèle peut alors engendrer une moindre précision des estimations en course à pied [31, 32, 270, 278, 279].

Toutefois, il faut souligner le fait que cette relation linéaire et le concept de vitesse (ou puissance) critique ont été réétudiés ultérieurement [28, 268, 280, 281]. De nouvelles équations plus complexes (*e.g.*, modèle hyperbolique à trois paramètres) ont ainsi été proposées et utilisées dans la modélisation des performances en course à pied [28, 261, 268, 281]. Si les études de Vandewalle et al. [31] et Gamelin et al. [28] ayant comparé des modèles hyperboliques à deux (*e.g.*, Hill and Scherrer) et trois paramètres (*e.g.*, Morton) ont montré une plus grande précision dans les valeurs de prédictions de performances en course à pied chez des coureurs élites ou chez des coureurs bien entraînés, respectivement, en appliquant le modèle le plus complexe ; les adaptations de ce modèle à trois paramètres n'ont pas fait l'objet de ce travail de recherche puisqu'il est nécessaire de connaître la vitesse maximale en *sprint*.

c. Loi puissance

La relation distance-temps peut également être modélisée à partir de lois de puissance individuelle. En effet, en 1906, Kennelly [30] a proposé une loi de puissance basée sur la relation entre la vitesse, la distance et le temps limite en prenant en considération les records du monde de différentes disciplines dont la course à pied. En d'autres termes, l'expression générale du modèle décrit les

vitesse les plus rapides qui peuvent être atteintes en course à pied. Un athlète est censé parcourir une distance de course avec une vitesse moyenne qui dépend de la distance totale de la course et de deux paramètres spécifiques au coureur, k et g . La relation entre la distance et le temps est telle que l'une est proportionnelle à une puissance fixe de l'autre :

$$D_{lim} = k \cdot T_{lim}^g \quad (5)$$

où k est une constante qui devrait, en théorie, être corrélée à la vitesse de course la plus élevée atteinte lors d'un *sprint* maximal [31, 32, 278, 279]. L'exposant g est un paramètre sans dimension qui peut être considéré comme un indice de capacité d'endurance (c'est-à-dire que g est proche de 1 lorsque les coureurs d'endurance ont une bonne capacité d'endurance et 0 lorsque la capacité est faible ; et si g est égal à 1, la relation de T_{lim} - D_{lim} serait parfaitement linéaire) [31, 32, 278, 279]. Lorsque ces deux paramètres sont inconnus, mais que des ensembles de données sont disponibles pour les variables T_{lim} et D_{lim} une analyse non linéaire des moindres carrés peut être effectuée pour déterminer les valeurs de k et g en utilisant des itérations répétées à travers un algorithme qui trouve les paramètres les mieux adaptés pour les données disponibles. Cela signifie que les coefficients sont fixés de telle sorte que la somme des différences au carré, entre les logarithmes observés des vitesses de course (*i.e.*, entre le logarithme de T_{lim} et le logarithme de D_{lim}) et ceux prédits, soit minimisée.

Par conséquent, si $D_{lim} = k \cdot T_{lim}^g$, le logarithme de D_{lim} est égal à :

$$\ln(D_{lim}) = \ln(k) + g \ln(T_{lim}) \quad (6)$$

$$\text{où } T_{lim} = e^{\ln(T_{lim})}$$

$$T_{lim} = e^{\frac{\ln(\frac{D_{lim}}{k})}{g}} \quad (7)$$

Comme pour les autres modèles abordés précédemment, cette loi puissance peut être utilisée pour prédire une performance sur une distance de course à pied à partir de deux autres performances ayant déjà été réalisées. Les valeurs de g et k sont

alors calculées à partir des valeurs de D_{lim} correspondant à T_{lim} des première et deuxième performances, sans tenir compte des valeurs de T_{lim} et D_{lim} de la troisième performance de course à pied.

Toutefois, comme l'illustre la Figure 35 pour les records du monde actuels en course à pied, plusieurs études s'accordent sur le fait que la loi de puissance a un point de rupture, en lien avec la transition entre l'effort en anaérobie (*i.e.*, exercice de haute intensité de courte durée) et l'effort en aérobie (*i.e.*, exercice léger ou modéré de plus longue durée) [8, 60, 105], différenciant les épreuves en fonction de la distance [113, 266]. Par conséquent, les coefficients de la loi puissance diffèrent. En d'autres termes, selon le panel de distances analysé, l'indice de la capacité d'endurance (*i.e.*, g) est modifié [17, 30, 113] en lien avec la prépondérance du métabolisme énergétique produisant l'énergie lors de l'effort.

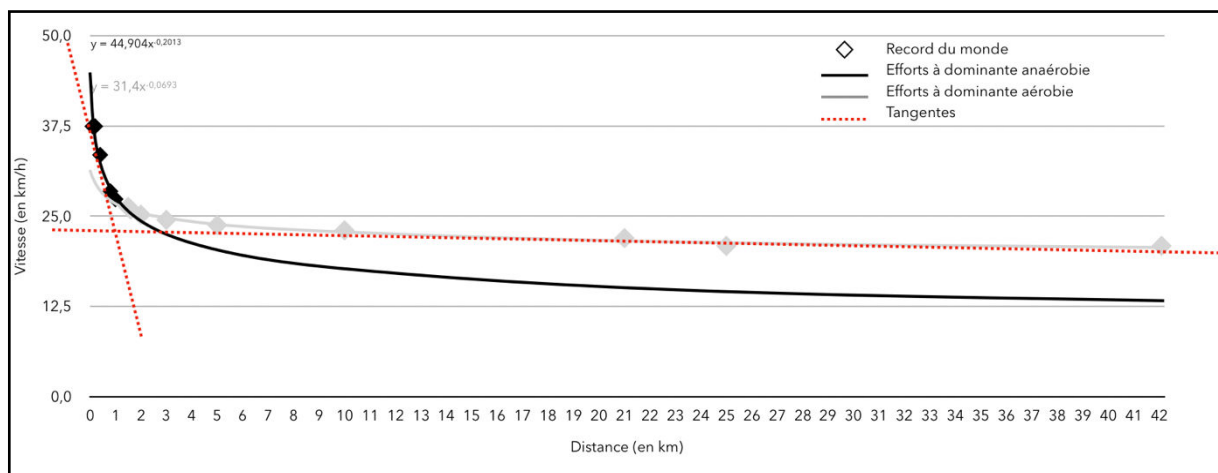


Figure 35. Évolutions des vitesses de records du monde actuels en course à pied en fonction de la distance selon deux lois de puissance différentes dépendantes de la filière énergétique dominante

Comme pour le modèle hyperbolique, l'augmentation du nombre de paramètres intégrés au modèle peut complexifier la modélisation de la loi puissance et influencer sur la précision des résultats. Par exemple, le modèle utilisé par Mulligan et al. [266] nécessite l'intégration de trois paramètres (*i.e.*, pente de la courbe en aérobie,

penne de la courbe en anaérobie ; point de rupture entre les deux métabolismes énergétiques) ; alors que, comme évoqué précédemment, le modèle de Riegel est relativement simple étant donné que l'exposant de la loi puissance est fixe (*i.e.*, indice de capacité d'endurance fixe) [113]. Toutefois, les modèles les plus simples [113, 255] apparaissent davantage limités (*i.e.*, biais plus important) dans la précision des prédictions des performances en course à pied en comparaison aux modèles plus complexes [266].

Enfin, si les modèles de loi puissance permettent de décrire les relations entre T_{lim} et D_{lim} (ou la vitesse), en particulier chez les coureurs d'endurance de haut niveau, et ont pour avantage d'être indépendants de la plage des temps d'épuisement [30], leur validité pour prédire les performances en course à pied chez les athlètes non-élites semble plus limitée [278, 279].

d. Intelligence artificielle (IA)

La dernière méthode présentée dans cette section repose sur l'IA. Ces dernières années, les algorithmes d'apprentissage automatique ont suscité un intérêt croissant dans le monde de la recherche. L'apprentissage supervisé est notamment l'une des méthodologies intelligentes qui a permis l'obtention de résultats encourageants (*i.e.*, prédiction de variables) dans de multiples domaines, dont le sport [282] qui est en pleine expansion [256, 282-284]. Cependant, bien que des modèles de régression par apprentissage supervisé tels que les réseaux de neurones artificiels (*i.e.*, abrégé ANN pour « *Artificial Neural Networks* ») [282, 284-286] ou les k-plus proches voisins (*i.e.*, abrégé KNN pour « *K-Nearest Neighbors* ») [23] aient été utilisés pour prédire les performances dans certaines activités sportives [287-291] ; ces modèles restent encore relativement méconnus et peu étudiés notamment en course à pied.

L'ANN est un puissant algorithme d'apprentissage supervisé de type boîte noire, capable de produire une cartographie entrée-sortie non linéaire [288, 292, 293]. Le modèle se compose d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une couche de sortie. Les composants interconnectés (*i.e.*, les neurones) transforment un ensemble d'entrées en une sortie souhaitée [292, 293]. La précision de ce type de modèle est généralement améliorée par l'utilisation de données supplémentaires (*i.e.*, les poids associés aux composants interconnectés changent continuellement) pendant le processus de formation du réseau ANN [282, 292].

Le modèle KNN, quant à lui, utilise l'un des types les plus simples d'algorithmes d'apprentissage automatique supervisé basé sur l'apprentissage par analogie, c'est-à-dire en comparant un exemple test donné avec des exemples d'apprentissage qui lui sont similaires [286, 292]. L'algorithme KNN de base comporte deux étapes. La première trouve les *k* exemples d'apprentissage qui sont les plus proches (la « proximité » est définie en fonction d'une métrique de distance, telle que la distance euclidienne) de l'exemple non vu, et la seconde étape prend la moyenne de ces *k* valeurs d'étiquette [288, 292]. Ce modèle d'apprentissage automatique est également connu pour ne pas nécessiter d'apprentissage (c'est-à-dire que le calcul de l'algorithme a lieu pendant l'exécution) car il mémorise l'ensemble de données d'apprentissage [288, 289, 292].

À notre connaissance, l'utilisation de ces algorithmes d'apprentissage, ANN et KNN, pour prédire la performance en course à pied (quelles que soient les disciplines étudiées, sur piste ou sur route), semble très limitée. Selon la littérature, il apparaît que les études réalisées se soient principalement intéressées à l'utilisation d'algorithmes d'apprentissage automatique (*e.g.*, *bagging*, réalisation de matrice) afin de prévoir des variations d'allure (*e.g.*, ralentissements pendant la course) ou de prédire des performances sur marathon [23, 256, 257]. De ces principaux

constats en course à pied et du réel intérêt que peut avoir l'IA sur la performance sportive en général (e.g., prévention des blessures, prédiction des performances) [282, 283, 285, 287, 290, 294], ces algorithmes supervisés constituent un champ de recherche particulièrement intéressant à exploiter.

3. Conclusion

De nombreux auteurs ont cherché à modéliser la performance sportive et à examiner la progression des performances humaines, en comparant les sexes et/ou les distances des épreuves de course à pied. À l'heure actuelle, la recherche et le développement de stratégies d'entraînement, de compétitions et d'outils permettant d'atteindre et de prédire les performances des coureurs, représentent un véritable intérêt pour les athlètes et les entraîneurs. La prédiction de la performance en course à pied, et donc de la performance potentielle, peut-être utile pour prescrire des allures d'entraînement et en compétition afin d'optimiser la performance en compétition, mais aussi détecter les futurs athlètes au fort potentiel. Différentes approches modélisant la relation entre le temps et la distance (ou la vitesse) de course ont été développées pour prédire les performances en course à pied, sur route ou sur piste. Chaque méthode de prédiction (interpolant et/ou extrapolant des performances), dépendante des données disponibles et/ou des paramètres qu'il est possible d'intégrer (e.g., données physiologiques, performances antérieures), présente des avantages et des inconvénients. Cependant, de manière générale, les travaux menés sur la prédiction de la performance en course à pied se sont principalement intéressés aux athlètes de sexe masculin et de haut niveau. En l'occurrence la compréhension et la prédiction des performances chez des athlètes, indépendamment du sexe, de niveau plus modeste en course à pied demeurent limitées. Ce champ, riche d'opportunités, constitue ainsi tout l'enjeu de la seconde partie de ce manuscrit à savoir tester et

comparer différents modèles de prédiction de la performance en course à pied pour tout athlète ou entraîneur désireux de l'appréhender et l'optimiser.

DEUXIÈME PARTIE : CADRE EXPÉRIMENTAL

Le cadre théorique et la revue de littérature, présentés dans les chapitres précédents, ont mis en avant de multiples paramètres à prendre en considération pour appréhender et comprendre la performance en course de demi-fond et de fond. Afin de modéliser et/ou prédire les performances, de nombreuses approches mathématiques considérant les données empiriques, biomécaniques ou physiologiques ont tenté de comprendre ce qui constitue la performance en course à pied. Certaines de ces approches, abordées précédemment, ont fait l'objet de comparaisons selon différents critères (e.g., distance de course, sexe, niveau de performance, nombre de performances, méthode d'interpolation ou d'extrapolation, type et nombre de paramètres intégrés). De ces analyses comparatives, il ressort des résultats de prédiction qui fluctuent selon les modèles en termes de justesse et de précision par rapport aux temps réels de performance [31, 32, 261, 270]. Si, de manière générale, les modèles basés sur le concept de vitesse critique (e.g., MLDT, modèle hyperbolique à deux paramètres) ont régulièrement fait l'objet de travaux sur la prédiction de la performance du fait de leur praticité d'application, ces derniers semblent être moins précis (*i.e.*, biais et variance élevés) que d'autres approches de prédiction telles que le nomogramme de Mercier et al. [26][270] ou le concept de loi puissance [31, 32, 261]. Dans la littérature, sur la base de comparaison de différents modèles (e.g., logarithmique, exponentiel, hyperbolique), ceux de type loi puissance se sont, en effet, généralement montrés davantage appropriés pour la prédiction des performances de course à pied, à condition que le paramètre g (*i.e.*, exposant pouvant être considéré comme l'indice de capacité d'endurance) soit correctement défini [17, 30, 113]. Plusieurs travaux ont effectivement démontré que ces modèles s'adaptent relativement bien aux références chronométriques de course à pied que constituent les records mondiaux [31, 32, 261, 266]. Il est notable également que les modèles

complexes, caractérisés par l'intégration de multiples paramètres, puissent accroître la précision des prédictions des performances en course à pied (e.g., modèle hyperbolique de vitesse critique à 3 paramètres) et ainsi améliorer la modélisation de la performance en comparaison à de simples modèles (e.g., modèle de vitesse critique linéaire) [31]. Toutefois, la difficulté d'application en lien avec le nombre de données à intégrer et à ajuster constitue vraisemblablement l'une des principales limites de ces modèles complexes [31].

Ainsi, modéliser et prédire la performance de course à pied n'est pas sans se questionner sur les intérêts et limites des approches développées, ni sans poser de problèmes méthodologiques (*i.e.*, choix de l'approche pour l'obtention de résultats valides) et épistémologiques (*i.e.*, connaissances développées dans le domaine de la course à pied). En d'autres termes, il apparaît que la prédiction de la performance s'établit sur l'identification des « bons » paramètres permettant de créer les meilleurs outils (*i.e.*, valides et précis) ; soit des outils capables de s'appliquer sur différents échantillons, tels que des échantillons de coureurs incluant différents niveaux d'expertise et diverses distances, par exemple.

Au regard du nombre d'athlètes pratiquant la course à pied et de l'intérêt de pouvoir modéliser et prédire la performance en demi-fond et fond, il semble essentiel de pouvoir tester la validité et comparer la précision de différents modèles de prédiction en course à pied. En outre, il apparaît que la plupart des modèles développés aient ciblé un profil d'athlète en particulier (e.g., haut niveau, sexe masculin). La grande variabilité dans les conditions de réalisation des performances des coureurs amateurs (e.g., courir sans préparation spécifique, mauvaise gestion de l'effort), pouvant conduire à des prédictions de performance aberrantes, pourrait être l'une des explications plausibles de ce choix d'athlètes de haut niveau dans de nombreux travaux [261]. Cependant, la majorité des modèles étant conçu sur les bases de records du monde, il semble essentiel de pouvoir tester leur validité ainsi que leur justesse et leur précision sur des échantillons d'athlètes de profil

hétérogène (e.g., niveau, sexe) [279], du fait de l'intérêt que peut avoir la performance en course à pied à l'heure actuelle.

La seconde partie de ce manuscrit est destinée à présenter les résultats de l'ensemble des études rétrospectives réalisées durant ce doctorat, contribuant à documenter et alimenter les connaissances des athlètes et entraîneurs de course à pied. Ces études, issues d'un partenariat avec la FFA, s'appuient sur le système d'information fédéral de la fédération FFA. Les bases de données de la FFA ou encore de *World Athletics* répertorient notamment l'ensemble des résultats athlétiques (e.g., enregistrements des performances, des compétitions, des records), les bilans ou encore le nombre de licenciés. La validité, la justesse et la précision des différentes approches analysées (i.e., capacité à prédire les performances) ont ainsi été testées sur des performances individuelles réelles d'athlètes de différents niveaux, de catégorie « *senior* » principalement, collectées directement *via* ces bases de données. En effet, comme évoqué précédemment, bien que la pratique puisse être pour le loisir, cette catégorie reste l'une des plus compétitives. Les athlètes *seniors*, en théorie plus matures physiologiquement que les catégories jeunes pour les courses de demi-fond et de fond notamment [13, 93-95], se préparent aux échéances de compétitions selon des calendriers dédiés afin de pouvoir performer.

Le chapitre I de cette deuxième partie expose notamment l'évolution des performances françaises des courses de demi-fond et de fond chez les femmes. Les chapitres II à V, portent sur la validation et la comparaison de méthodes de prédiction des performances des courses de demi-fond et de fond, sur piste et sur route, chez les hommes et/ou les femmes. Enfin le chapitre VI met en avant une analyse du taux d'amélioration des performances de demi-fond et de fond précédant la réalisation de records personnels chez les hommes et chez les femmes. Un index de performance à visée pratique, est également proposé, afin d'évaluer

l'évolution des performances et permettre une éventuelle détection et orientation des athlètes au fort potentiel.

CHAPITRE I. ÉVOLUTION DES PERFORMANCES/CARACTÉRISTIQUES DES COUREUSES FRANÇAISES SUR LES 12 DERNIÈRES ANNÉES

- ❖ Publication scientifique (cf. Annexe 2)
 - **Lerebourg L**, Coquart JB. *Changes in performances/characteristics of French female runners over the last 12 years. Research in Sports Medicine*. 2021, 29 (2), 185-195.

- ❖ Communication orale avec acte
 - **Lerebourg L**, Coquart JB. Évolution des performances de moyennes et de longues distances sur piste et des caractéristiques des coureuses françaises lors des trois derniers cycles des Jeux Olympiques. 18^{ème} Congrès de l'ACAPS. 29-31 octobre 2019. Paris, France.

1. Introduction

Malgré les nombreuses recherches sur les performances en course à pied sur route de longue distance comme le semi-marathon ou le marathon [17, 127, 157, 295-299], les études portant sur les courses de plus courtes distances sur piste (e.g., 10 000 m) apparaissent relativement limitées et plus particulièrement chez les athlètes de niveau modeste.

Plusieurs auteurs ont étudié la progression des performances afin d'identifier les variables qui peuvent être modifiées avec l'entraînement ainsi que d'autres facteurs hors de contrôle de l'athlète qui pourraient influencer les performances de longue distance [157, 201, 295-297] telles que les meilleures performances mondiales sur marathon [5, 7, 201]. Par exemple, le pic de performance sur marathon est atteint en moyenne vers 25-27 ans pour les athlètes féminines de niveau mondial [13],

probablement en raison des limites physiologiques liées à l'âge. De plus, un IMC de 18,2 kg.m⁻² semble être optimal pour les marathoniennes figurant dans le Top 100 mondial et il a également été démontré, selon Marc et al. [7] que ces féminines disputent généralement cette épreuve sur les périodes du mois d'avril (27,6%) ou d'octobre (21,1%). Si ces différentes études, analysant les performances et/ou les caractéristiques des coureuses internationales, peuvent aider à mieux comprendre les performances de haut niveau ainsi que leur évolution, ces dernières ne concernent toutefois qu'un petit nombre d'athlètes européennes sous-représentées dans le Top 100 mondial. En effet, Marc et al. [7] ont indiqué que 52% des 100 meilleures marathoniennes en 2011 étaient africaines et que seulement 24% étaient européennes (sans aucune Française dans le Top 10 européen). Plus récemment, Marc et al. [300] ont confirmé cette supériorité des athlètes féminines d'origine africaine sur marathon, mais également sur piste. Cependant, la majeure partie des athlètes pratiquant la course à pied n'est pas de niveau international, et la plupart des études réalisées à ce sujet n'ont pas analysé les performances chez les athlètes de niveau plus modeste (e.g., national, interrégional, régional, départemental) [13, 73, 93, 158, 300]. Par conséquent, les résultats obtenus peuvent ne pas être très représentatifs du niveau de pratique de la plupart des athlètes de nationalité française. Une étude plus étroitement axée sur les performances des athlètes européennes, comme les coureuses françaises, semble donc intéressante pour fournir de nouvelles données aux athlètes, aux entraîneurs et aux fédérations préparant les futures compétitions internationales, comme les JO ou les championnats à différents niveaux.

L'objectif de cette étude était donc d'examiner l'évolution des performances en course à pied sur cinq distances sur piste (en demi-fond et fond) et les caractéristiques des coureuses françaises au cours des trois derniers cycles olympiques.

2. Méthodes

Les bilans officiels de la FFA pour les athlètes féminines âgées de 23 à 34 ans (*i.e.*, tranche d'âge correspondant à la catégorie « *senior* » en France) disputant les épreuves de 800 m, 1 500 m, 3 000 m, 5 000 m et 10 000 m du 1^{er} janvier 2005 au 31 décembre 2016 ont été étudiés. Les performances sur route comme le 10 km ont été exclues, et seules les performances sur piste ont été conservées afin de contrôler les éventuels effets du profil de la course sur la performance. Pour chaque distance et chaque année, seule la meilleure performance de l'athlète a été retenue. Au total, 22 839 performances ont été collectées. Pour chaque performance, le temps de course, l'indication du record personnel et la date de naissance de l'athlète ont été extraits.

De plus, en accord avec le barème officiel de la FFA, toutes les athlètes retenues ont été classées selon cinq groupes de niveau (*i.e.*, départemental, régional, interrégional, national et international ; Tableau 14).

Tableau 14. Barème FFA « Femme » du niveau de performance en course de demi-fond et de fond au 1^{er} janvier 2010

Distance	International	National	Interrégional	Régional	Départemental
800 m	2'01"84	2'17"00	2'28"00	2'46"00	3'08"00
1 500 m	4'11"00	4'41"00	5'05"00	5'40"00	6'50"00
3 000 m	9'05"00	10'05"00	11'00"00	12'15"00	14'40"00
5 000 m	15'30"00	17'40"00	19'00"00	21'00"00	26'00"00
10 000 m	32'40	37'00	41'00	47'00	1 h 00'00

Cette étude, comme les suivantes qui font partie d'un même *work package*, a été approuvée par le CERSTAPS (2019-22-02-31). Par ailleurs, ce *work package* a fait l'objet d'une déclaration légale auprès de la déléguée à la protection des données de l'URN dans le cadre du règlement général sur la protection des données (inscrit

sous le n°2020.034), conformément au règlement général européen sur la protection des données.

Les données ont été présentées sous forme d'effectif et de pourcentage pour les variables nominales, ou sous forme de moyenne \pm écart-type (*i.e.*, abrégé SD pour « *Standard Deviation* ») pour les données numériques.

Pour chaque année et chaque distance, le nombre d'athlètes, la meilleure performance annuelle de chaque athlète, les temps de course, le nombre de records personnels, l'âge et le nombre d'athlètes pour chaque niveau de performance ont été analysés. De plus, chaque performance a été classée à l'aide du barème présenté dans le Tableau 14.

Cette analyse statistique descriptive a ensuite été réitérée pour les athlètes figurant dans le Top 10 français afin d'obtenir des informations sur les meilleures coureuses nationales.

De plus, ces analyses descriptives ont été complétées par des analyses statistiques inférentielles. Pour ce faire, les performances ont été regroupées par période de 4 ans correspondant aux cycles des JO. Ainsi, la période de 2005 à 2008 correspondait à Athènes-Pékin ($P_{2005-2008}$), la période allant de 2009 à 2012 correspondait à Pékin-Londres ($P_{2009-2012}$), et la période allant de 2013 à 2016 correspondait à Londres-Rio ($P_{2013-2016}$).

Les données numériques ont été comparées entre les cycles olympiques à l'aide d'une analyse de la variance (*i.e.*, abrégé ANOVA pour « *Analysis of Variance* ») à trois facteurs (**groupes** : totalité des athlètes vs Top 10 \times **périodes** : $P_{2005-2008}$ vs $P_{2009-2012}$ vs $P_{2013-2016}$ \times **distances** : 800 m vs 1 500 m vs 3 000 m vs 5 000 m vs 10 000 m).

Lorsque des différences significatives ont été observées, un test *post-hoc* de Tukey a été effectué. L'amplitude de la différence a été évaluée par la taille de l'effet (*i.e.*, abrégé ES pour « *Effect Size* »). L'échelle proposée par Cohen [301] a été utilisée pour l'interprétation. L'amplitude de la différence a été considérée comme insignifiante ($ES < 0,2$), petite ($0,2 \leq ES < 0,5$), modérée ($0,5 \leq ES < 0,8$) ou importante ($ES \geq 0,8$).

Une analyse statistique descriptive a également été utilisée pour comparer les données en fonction des catégories étudiées.

Le seuil de significativité statistique a été fixé à $p < 0,05$ et toutes les analyses ont été effectuées avec le logiciel *Statistical Package for the Social Sciences* (version 20.0, Chicago, IL, USA).

3. Résultats

Le nombre retenu de performances était de 7400 sur la période allant de 2005 à 2008, 7293 sur la période allant de 2009 à 2012, et 8146 sur la période allant de 2013 à 2016. Le nombre le plus élevé de performances a été enregistré lors de la période 2013-2016 pour toutes les distances, à l'exception du 10 000 m où 90 performances ont été enregistrées contre 113 entre 2005 et 2008 et 102 entre 2009 et 2012.

Les performances moyennes, pour les deux groupes, sur le 5 000 m ont été significativement améliorées entre 2013 et 2016 par rapport à 2005-2008 ($p = 0,01$; avec $ES = 0,20$ pour le Top 10 et $ES = 0,15$ pour la totalité des athlètes ; Tableau 15). De plus, sur le 10 000 m, les performances moyennes des deux groupes étaient significativement plus faibles entre 2005 et 2008 par rapport aux autres périodes ($p < 0,001$). Bien qu'un effet négligeable ait été observé entre les périodes 2005-2008

et 2009-2012 pour le Top 10 (ES = 0,08), l'ES était plus élevé pour l'échantillon regroupant la totalité des coureuses (ES = 0,27). De plus, ont été constatés, un ES similaire pour l'ensemble des athlètes (ES = 0,33) et un effet modéré pour le Top 10 (ES = 0,73) entre les périodes 2005-2008 et 2013-2016. Enfin, les performances sur le 10 000 m étaient également significativement plus élevées, quel que soit le groupe étudié, lors de la période 2013-2016 par rapport à 2009-2012 ($p < 0,001$; ES = 0,60 pour Top 10 et ES = 0,10 pour toutes les coureuses). Aucune différence significative n'a été montrée sur les épreuves de 800 m, 1 500 m et 3 000 m quels que soient le groupe et les périodes prises en considération (Tableau 15).

Tableau 15. Évolution des performances des athlètes françaises entre 2005 et 2016

Distance	Groupe	Période		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800 m	Meilleure performance	1'59"	1'59"	1'58"
	Top 10	2'03"± 2"	2'04"± 2"	2'04"± 2"
	Totalité des athlètes	2'36"± 17"	2'36"± 16"	2'35"± 16"
1 500 m	Meilleure performance	4'00"	4'00"	4'03"
	Top 10	4'15"± 7"	4'15"± 7"	4'16"± 4"
	Totalité des athlètes	5'22"± 30"	5'22"± 30"	5'20"± 30"
3 000 m	Meilleure performance	8'35"	8'56"	8'52"
	Top 10	9'17"± 12"	9'30"± 13"	9'25"± 10"
	Totalité des athlètes	11'52"± 1'15"	11'52"± 1'13"	11'52"± 1'16"
5 000 m	Meilleure performance	14'59"	15'21"	15'08"
	Top 10	16'22"± 28"	16'26"± 28"	16'17"± 26" ^a
	Totalité des athlètes	19'47"± 2'04"	19'38"± 2'02"	19'28"± 2'00" ^a
10 000 m	Meilleure performance	31'57"	31'36"	31'53"
	Top 10	38'03"± 3'19"	37'47"± 3'22" ^{a,c}	36'02"± 2'03" ^{a,b}
	Totalité des athlètes	44'07"± 5'47"	42'40"± 5'05" ^{a,c}	42'04"± 6'39" ^{a,b}

Légende : ^a Différence significative avec P₂₀₀₅₋₂₀₀₈ ($p < 0,01$), ^b Différence significative avec P₂₀₀₉₋₂₀₁₂ ($p < 0,001$), ^c Différence significative avec P₂₀₁₃₋₂₀₁₆ ($p < 0,001$)

Pour l'ensemble des athlètes, le nombre de records personnels sur le 800 m était plus élevé lors de la période 2005-2008, et sur le 1 500 m le nombre était légèrement inférieur lors de la période 2009-2012 (Tableau 16).

Dans le Top 10, le nombre de records personnels sur le 800 m et sur le 1 500 m était plus élevé lors de la période 2013-2016 (Tableau 16).

Sur les épreuves de 3 000 m, 5 000 m ou 10 000 m, quel que soit le groupe étudié, le nombre de records personnels était relativement proche.

Tableau 16. Évolution des pourcentages de records personnels chez les performeuses françaises entre 2005 et 2016

Distance	Groupe	Période		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800 m	Meilleure performance	1 (100,0)	1 (100,0)	1 (100,0)
	Top 10	9 (22,5)	8 (20,0)	20 (50,0)
	Totalité des athlètes	936 (41,6)	770 (34,9)	908 (36,7)
1 500 m	Meilleure performance	0 (0,0)	1 (100,0)	1 (100,0)
	Top 10	9 (22,5)	7 (17,5)	21 (52,5)
	Totalité des athlètes	981 (42,1)	822 (35,6)	1022 (40,7)
3 000 m	Meilleure performance	1 (100,0)	1 (100,0)	1 (100,0)
	Top 10	18 (45,0)	15 (37,5)	16 (40,0)
	Totalité des athlètes	1035 (47,5)	991 (45,5)	1178 (47,7)
5 000 m	Meilleure performance	0 (0,0)	0 (0,0)	1 (100,0)
	Top 10	10 (25,0)	14 (35,0)	19 (47,5)
	Totalité des athlètes	273 (52,0)	255 (51,6)	319 (53,0)
10 000 m	Meilleure performance	0 (0,0)	1 (100,0)	1 (100,0)
	Top 10	26 (65,0)	25 (62,5)	25 (62,5)
	Totalité des athlètes	79 (69,9)	71 (69,6)	65 (72,2)

Légende : Les nombres représentent le nombre d'athlètes (%) ayant réalisé un record personnel.

Plus les distances étaient longues, plus les athlètes étaient âgées (quel que soit le groupe ; $p < 0,001$), excepté sur les épreuves de 3 000 m et 5 000 m ($p = 0,103$; Tableau 17).

Les athlètes du Top 10 étaient significativement plus jeunes que le groupe rassemblant la totalité des athlètes lors de la période 2013-2016 ($p < 0,05$; ES =

0,40 pour le 800 m ; ES = 0,44 pour le 1 500 m ; ES = 0,38 pour le 3 000 m ; ES = 0,25 pour le 5 000 m ; et ES = 0,21 pour le 10 000 m ; Tableau 17). Des différences significatives ont été démontrées entre chaque distance, excepté entre le 3 000 m et le 5 000 m, et aucune différence significative n'a été constatée sur une même distance, quelle que soit la période ($p > 0,05$).

Tableau 17. Évolution de l'âge des athlètes françaises entre 2005 et 2016

Distance	Groupe	Période		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800 m	Meilleure performance	23	31	26
	Top 10	28,7 ± 3,7	28,0 ± 3,6	26,7 ± 2,7
	Totalité des athlètes	28,7 ± 4,7	28,7 ± 4,8	28,6 ± 4,8 ^a
1 500 m	Meilleure performance	26	31	23
	Top 10	29,6 ± 3,9	29,6 ± 3,7	27,7 ± 3,6
	Totalité des athlètes	30,1 ± 5,0	29,9 ± 5,0	29,9 ± 5,0 ^a
3 000 m	Meilleure performance	26	27	29
	Top 10	30,0 ± 3,6	30,0 ± 4,3	29,3 ± 4,5
	Totalité des athlètes	31,4 ± 4,9	31,5 ± 5,1	31,2 ± 5,0 ^a
5 000 m	Meilleure performance	31	38	24
	Top 10	31,4 ± 2,7	30,1 ± 5,2	29,7 ± 4,7
	Totalité des athlètes	31,4 ± 4,8	30,8 ± 5,1	30,9 ± 4,8 ^a
10 000 m	Meilleure performance	34	38	24
	Top 10	32,7 ± 3,6	33,4 ± 4,7	31,4 ± 5,4
	Totalité des athlètes	32,8 ± 4,4	33,4 ± 5,1	32,5 ± 5,2 ^a

Légende : ^a Différence significative avec le Top 10 ($p < 0,05$).

Note : Différences significatives obtenues entre les distances, excepté entre les épreuves de 3 000 m et de 5 000 m ($p < 0,05$).

Sur l'épreuve de 10 000 m, le pourcentage de performeuses françaises internationales était plus élevé (*i.e.*, 3,9% à 6,2%), quelle que soit la période, en comparaison aux épreuves de 800 m, 1 500 m, 3 000 m et 5 000 m (*i.e.*, 0,1% à 0,6%) (Tableau 18).

Tableau 18. Évolution du nombre et du pourcentage d'athlètes françaises selon chaque niveau de performance entre 2005 et 2016

Distance	Niveau	Période		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800 m	International	13 (0,6)	8 (0,4)	11 (0,4)
	National	212 (9,4)	222 (10,1)	241 (9,7)
	Interrégional	588 (26,1)	522 (23,6)	607 (24,6)
	Régional	896 (39,8)	905 (41,0)	1040 (42,1)
	Départemental	542 (24,1)	551 (25,0)	573 (23,2)
1 500 m	International	12 (0,5)	9 (0,4)	8 (0,3)
	National	182 (7,8)	192 (8,3)	258 (10,3)
	Interrégional	502 (21,5)	516 (22,3)	581 (23,1)
	Régional	973 (41,7)	921 (39,9)	1013 (40,3)
	Départemental	662 (28,4)	672 (29,1)	653 (26,0)
3 000 m	International	8 (0,4)	2 (0,1)	3 (0,1)
	National	115 (5,3)	110 (5,0)	136 (5,5)
	Interrégional	433 (19,9)	403 (18,5)	531 (21,5)
	Régional	893 (41,0)	930 (42,7)	919 (37,2)
	Départemental	731 (33,5)	734 (33,7)	880 (35,6)
5 000 m	International	3 (0,6)	3 (0,6)	2 (0,3)
	National	84 (16,0)	97 (19,6)	115 (19,1)
	Interrégional	109 (20,8)	94 (19,0)	154 (25,6)
	Régional	182 (34,7)	181 (36,6)	186 (30,9)
	Départemental	147 (28,0)	119 (24,1)	145 (24,1)
10 000 m	International	7 (6,2)	4 (3,9)	5 (5,6)
	National	8 (7,1)	11 (10,8)	22 (24,4)
	Interrégional	13 (11,5)	22 (21,6)	18 (20,0)
	Régional	52 (46,0)	40 (39,2)	20 (22,2)
	Départemental	33 (29,2)	25 (24,5)	25 (27,8)

4. Discussion

L'objectif de cette étude était d'analyser l'évolution des performances des athlètes françaises *seniors* sur cinq distances de course à pied sur piste au cours de trois cycles olympiques.

Les résultats ont montré que le nombre de performances, sur les épreuves de 800 m, 1 500 m, 3 000 m et 5 000 m, a augmenté lors de la période 2013-2016. Cet accroissement pourrait s'expliquer par l'augmentation du nombre de clubs d'athlétisme lors de cette période (+6,5%), passant de 2252 en 2013 à 2399 en 2016 [302], ainsi que par l'augmentation du nombre de licenciées au sein de la fédération. En effet, chez les femmes *seniors* (entre 23 et 34 ans), la FFA a connu une forte croissance à la fin de chaque cycle olympique, avec 7622 athlètes en 2004, 9965 en 2008 (+30,7%), 12 969 en 2012 (+30,1%) et 17 932 en 2016 (+38,3%). Cette augmentation du nombre de licenciées peut également s'expliquer par la mise en place d'un plan de développement [302]. Par exemple, lors de la période 2013-2016, la FFA a fait l'objet d'une réorganisation fédérale et depuis cette dernière s'organise en fonction de l'émergence de nouveaux métiers (avec un budget de 22,6 M€ en 2015 contre 17,7 M€ en 2012) et de la mise en œuvre de nouveaux moyens pour répondre à des objectifs précis (e.g., développement territorial et restructuration des clubs d'athlétisme, rénovation du site internet de la FFA, communication pour promouvoir l'athlétisme avec une image plus attractive...) [302].

Bien que le nombre de performances, sur les épreuves de 800 m, 1 500 m, 3 000 m et 5 000 m, ait augmenté lors de la période 2013-2016 chez les athlètes françaises, il a diminué sur le 10 000 m. Cette baisse pourrait s'expliquer par la réorientation des athlètes vers une même distance de course sous un format sur route plus

populaire (e.g., 10 km route à la place du 10 000 m piste). En effet, si le nombre de courses sur route comme le 10 km (compétition officielle) ou le *trail-running* a considérablement augmenté ces dernières années, le nombre de courses de 10 000 m piste est relativement limité en France [91]. À noter que dans d'autres pays tels que les États-Unis ou le Japon, les courses de 10 000 m piste sont organisées beaucoup plus régulièrement et offrent ainsi une densité de course potentiellement plus importante pour la réalisation de performances. En France, si les championnats nationaux de 10 000 m sont l'une des principales opportunités pour courir sur la distance, dans l'ensemble les occasions pour réaliser des performances sur 10 000 m restent relativement très limitées sur le territoire national.

Les résultats ont également montré que les performances des 2 groupes ont été améliorées sur le 5 000 m et le 10 000 m lors de la période 2013-2016 (Tableau 15). Cette amélioration est probablement liée à de multiples facteurs tels que les progrès technologiques (e.g., chaussures de course) ou encore le plan de développement mis en place par la FFA (e.g., directives du développement des courses sur route) [89, 302, 303]. Avec l'évolution de la technologie, certains demi-fondeurs ou fondeurs de niveau régional voire national, ont fait le choix de s'équiper de chaussures de *running* aux propriétés élastiques favorisant l'absorption des chocs (moins traumatisantes afin de diminuer le risque de blessures par rapport aux chaussures à pointes) ou encore favorisant la restitution de l'énergie avec des plaques de carbone pour optimiser leurs performances (e.g., le premier marathon en moins de 2 h a été couru avec ce type de chaussures) [174, 179, 181, 228]. De plus, parmi les directives de la FFA, l'un des objectifs était d'augmenter la popularité et le nombre de courses sur route *via* le développement du secteur « hors stade » afin de donner aux athlètes la possibilité de courir avec davantage de densité. À noter également que, la FFA a permis et permet encore la qualification sur 10 000 m aux championnats de France *via* la réalisation d'une performance sur 10 km route

(i.e., la seule et unique distance sur route où il est possible de se qualifier sur la distance équivalente à des championnats nationaux sur piste).

Les résultats ont révélé une augmentation du nombre de records personnels sur les épreuves de 800 m et de 1 500 m pour le Top 10 (Tableau 16). Une des raisons de cette augmentation pourrait être la diminution progressive de l'âge moyen de ces athlètes du Top 10. En effet, l'âge moyen des femmes du Top 10 sur 800 m était de $28,7 \pm 3,7$ ans lors de la période 2005-2008, de $28,0 \pm 3,6$ ans lors de la période 2009-2012 et seulement de $26,7 \pm 2,7$ ans lors de la période 2013-2016 (Tableau 17). La même tendance a été observée sur le 1 500 m (i.e., 27,7 ans entre 2013 et 2016 vs 29,6 ans sur les deux autres périodes étudiées). Les études menées à ce sujet ont généralement observé un âge moyen du pic de performance à hauteur de 25-27 ans chez les athlètes féminines de niveau mondial [13, 73, 93, 159]. Ainsi, il semble possible d'atteindre le pic de performance à un âge plus avancé sur les distances les plus longues (i.e., courses de fond), nécessitant une capacité d'endurance aérobie plus élevée (e.g., 5 000 m et 10 000 m) que sur les courses de moyenne distance (i.e., courses de demi-fond) caractérisées par une puissance aérobie plus élevée (e.g., 800 m et 1 500 m). Cette différence, en accord avec plusieurs études [158, 159], pourrait être expliquée par les paramètres physiologiques caractérisant les capacités humaines comme la capacité à produire de la puissance aérobie ou à développer une capacité d'endurance aérobie, mais aussi les connaissances ou l'expérience de course évoluant avec la maturité de l'athlète [93-95, 127].

Le nombre élevé de records personnels sur 800 m et 1 500 m pour le Top 10 entre 2013-2016 peut également être lié au nombre de compétitions, plus important, durant la même période [304]. À l'inverse, un nombre plus limité de compétitions (i.e., opportunités plus restreintes) dans le programme du circuit national sur les

épreuves de 3 000 m, 5 000 m et 10 000 m, lors des périodes 2009-2012 et 2013-2016 (avec respectivement 13 et 8 compétitions sur 3 000 m, 10 et 14 sur 5 000 m, et 0 sur 10 000 m), pourrait expliquer le fait qu'il y ait eu moins de records personnels établis pour les athlètes figurant dans le Top 10.

Pour l'ensemble des performances analysées (comprenant le Top 10), le nombre de records personnels sur le 800 m était plus élevé entre 2005 et 2008 comparé aux deux autres périodes (*i.e.*, 41,6% de records établis dans la période 2005-2008 vs 34,9 et 36,7 pour les périodes 2009-2012 et 2013-2016, respectivement) (Tableau 16). Là encore, les athlètes françaises ont pu bénéficier d'un plus grand nombre de compétitions. De plus, il est possible que les modifications du règlement officiel de la FFA en 2012, permettant de se qualifier aux championnats nationaux (*i.e.*, abandon des temps de qualification vs introduction de la place au bilan), aient exercé une légère influence sur les performances en incitant notamment les athlètes à battre leur record personnel afin de figurer dans le quota des places qualificatives. À noter également qu'à l'échelle internationale, l'annonce de minima qualificatifs relativement élevés (pouvant être revus à la hausse par chaque fédération) permettant l'accès aux championnats d'Europe ou du monde, par exemple, a potentiellement de façon indirecte incité les meilleurs athlètes à améliorer leurs performances afin de pouvoir espérer se qualifier.

Enfin, de manière générale, un nombre de records personnels plus faible lors de la période 2009-2012 pourrait être expliqué par un manque relatif de compétitions ou être en lien avec des conditions environnementales moins favorables. Il est bien admis que les conditions météorologiques telles que les précipitations, le vent et la température de l'air influencent les performances de course [5, 7, 201]. Sur les courses de demi-fond comme le 1 500 m, il a été montré que les meilleures performances sont réalisées à 20-25°C car cette plage de température semble

favoriser les capacités physiologiques humaines optimales [6]. Si l'on s'intéresse par exemple à quelques bulletins officiels météorologiques (français) des années 2011, ces derniers ont notamment indiqué que le printemps 2011 fut le plus chaud jamais enregistré et que l'automne 2011 fut le deuxième automne le plus chaud depuis l'année 1900. Ces conditions environnementales inhabituelles pourraient, par conséquent, avoir eu une influence potentielle sur le nombre de records personnels lors de la période 2009-2012.

L'étude a aussi montré que plus les distances étaient longues, plus les athlètes étaient âgés, quel que soit le groupe, excepté sur les épreuves de 3 000 m et de 5 000 m, où il n'y avait pas de différence significative entre les âges (Tableau 17). Selon Longo et al. [159], cette observation suggère que l'âge du pic de performance en course à pied augmente avec la longueur de la course.

Notre étude a confirmé cette tendance, en soulignant le fait, par exemple, que les femmes du Top 10 ont atteint leur performance maximale à un plus jeune âge sur le 800 m que leurs homologues alignées sur l'épreuve du 10 000 m (*i.e.*, $26,7 \pm 2,7$ vs $31,4 \pm 5,4$; respectivement). Cependant, l'âge du pic de performance dans le Top 10 français semble être légèrement plus élevé que chez les meilleurs athlètes de niveau mondial (*i.e.*, $25,8 \pm 4,8$ ans pour la course de demi-fond vs $28,9 \pm 4,7$ ans pour la course de fond) [159]. L'âge inférieur moyen du Top 10 pourrait s'expliquer par la politique de la FFA (*e.g.*, « *Kid Stadium* » en 1998 ou encore « Planètes Athlé Jeunes » en 2008) visant à : (a) faire en sorte que les enfants/adolescents puissent découvrir les activités sauter, courir, lancer et développer des compétences athlétiques en privilégiant le plaisir afin qu'ils puissent continuer à pratiquer l'athlétisme après l'âge de 16 ans, en se spécialisant selon leurs envies et leurs qualités, et (b) développer des espaces de travail et des équipements adaptés aux enfants, facilement implantables sur différents sites pour former et détecter les jeunes talents, afin de les orienter vers une spécialité athlétique vers l'âge de 16 ans.

En effet, les athlètes commençant un entraînement spécialisé à un jeune âge peuvent être plus susceptibles d'atteindre leur performance maximale à un âge plus précoce que leurs homologues se spécialisant un peu plus tard (*i.e.*, à partir de 16 ans, la catégorie U18 se caractérise par un accent accru sur la performance et le développement de compétences spécifiques) [9, 305, 306].

Les athlètes figurant dans le Top 10 étaient en moyenne plus jeunes que l'ensemble des athlètes analysées, quelle que soit l'épreuve. Par conséquent, l'âge semble être un facteur limitant les performances maximales en raison des modifications physiologiques qui y sont liées. Par exemple, bien que la performance maximale des coureurs de fond soit dépendante de multiples facteurs physiologiques, l'un des principaux déterminants de la performance est le $\dot{V}O_{2max}$ [108]. En Physiologie de l'effort, il est reconnu que le $\dot{V}O_{2max}$ détermine la puissance du système aérobie et permet notamment d'évaluer l'efficacité des systèmes respiratoires, cardiovasculaires et musculaires de l'athlète. Le $\dot{V}O_{2max}$ peut être limité et régulé par des facteurs liés au débit sanguin maximal de l'athlète ainsi qu'à la quantité maximale d'oxygène que le muscle peut utiliser. L'équation de Fick (*i.e.*, $\dot{V}O_2 = \dot{Q}c \times [Ca_{O_2} - Cv_{O_2}]$) permet notamment de faire le lien avec le $\dot{Q}c$ (*i.e.*, $FC \times VES$) et l'utilisation périphérique de l'oxygène (*i.e.*, $Ca_{O_2} - Cv_{O_2}$). Le $\dot{V}O_{2max}$ est, par conséquent, limité par les valeurs maximales de $\dot{Q}c$ et $Ca_{O_2} - Cv_{O_2}$ [134]. En outre, il est largement reconnu que la FC_{max} diminue avec l'âge [157]. Ainsi, il ne semble pas illogique que les athlètes du Top 10, possédant généralement un $\dot{V}O_{2max}$ supérieur aux autres coureurs, soient plus jeunes.

Le pourcentage d'athlètes féminines, selon chaque niveau de performance, était relativement constant d'une période à l'autre (Tableau 18). Cependant, l'analyse descriptive de ces données montre un pourcentage plus élevé de concurrentes de niveau international sur le 10 000 m par rapport aux autres distances, quelle que soit la période (*i.e.*, 3,9-6,2% pour le 10 000 m vs 0,1-0,6% pour toutes les autres

distances et pour toutes les périodes) (Tableau 18). Ce pourcentage plus élevé pourrait être expliqué par un meilleur niveau concernant les athlètes spécialisées dans cette discipline ou également par des procédures de qualification relativement moins élevées pour atteindre les championnats. En 2016, par exemple, les minima imposés par la FFA pour le niveau international (Tableau 14) correspondaient à 94,6%, 93,8%, 92,3%, 91,7% et 89,6% de la meilleure performance mondiale annuelle sur les épreuves de 800 m, 1 500 m, 3 000 m, 5 000 m et 10 000 m, respectivement, suggérant ainsi qu'il était plus facile d'atteindre un haut niveau de performance avec l'allongement de la distance de course. Par conséquent, au regard des minima fixés, certaines athlètes françaises auraient pu se réorienter en s'entraînant spécifiquement sur des distances plus longues telles que le 10 000 m, dans l'espoir de se qualifier plus facilement en équipe de France.

5. Conclusion

Les principaux résultats ont montré que le nombre de performances sur piste a augmenté ces dernières années (à l'exception du 10 000 m) et que les performances se sont nettement améliorées sur les plus longues distances (5 000 m et 10 000 m). Ces résultats peuvent être expliqués par le plan de développement mis en place par la FFA. De manière générale, les distances les plus longues étaient généralement effectuées par des athlètes féminines plus âgées. De plus, un pourcentage plus élevé d'athlètes de niveau international a été constaté sur l'épreuve du 10 000 m (certainement en raison de minima qualificatifs moins élevés). Par conséquent, le plan de développement mis en place par la FFA [302] semble avoir contribué à encourager les athlètes françaises à participer aux courses sur piste et à améliorer leurs performances sur les plus longues distances.

CHAPITRE II. PRÉDICTION DES PERFORMANCES EN COURSE DE DEMI-FOND ET DE FOND À L'AIDE DE NOMOGRAMMES

❖ Publication scientifique (cf. Annexe 3)

- **Lerebourg L**, Guignard B, Racil G, Jlid MC, Coquart JB. *Prediction of middle-distance running performances of female runners using nomograms. International Journal of Sports Medicine*. 2021,19 (10). Epub ahead of print.

1. Introduction

Les relations, entre la distance (ou la vitesse) de course et le temps, ont longtemps été utilisées pour prédire la performance en course à pied [107, 109, 114]. Si plusieurs études se sont intéressées à la prédiction des performances de course sur piste, ces dernières ont principalement été réalisées avec des coureurs masculins [23, 270, 25, 32], notamment les hommes de niveau élite [31, 32, 32] en utilisant par exemple les concepts de loi puissance [23, 31, 32] ou de vitesse critique [23, 28, 31, 32, 270], mais aussi des nomogrammes [24, 25, 27, 270]. Cependant, il apparaît que les études sur la prédiction des performances en course à pied chez les athlètes féminines, et plus particulièrement chez des coureuses non-élites, soient très limitées.

Les analyses empiriques et expérimentales d'études antérieures suggèrent que la prédiction de la performance en course à pied sur une distance à partir de performances récentes sur deux autres distances plus courtes et/ou plus longues est réalisable en utilisant des méthodes pratiques simples. Les nomogrammes de Mercier et al. [26] et de Vandewalle [27] en sont des exemples notables.

Coquart et al. [265] ont montré que le nomogramme de Mercier et al. [26] fournit des prédictions valides des performances de course de fond sur route (*i.e.*, 10 km, 20 km et marathon). Les corrélations obtenues entre les performances réelles et prédites étaient de $r = 0,89$ sur 10 km et sur le marathon, et de $r = 0,97$ sur le 20 km [265]. Les performances prédites étaient surestimées en moyenne de 13 s et de 1 min 25 s, sur 10 km et marathon, respectivement, et sous-estimées à hauteur de 27 s sur 20 km [265]. D'autres études ont également confirmé la validité des prédictions pour les performances sur piste en demi-fond et en fond (*i.e.*, 3 000 m, 5 000 m et 10 000 m) [24, 25, 270]. Par exemple, selon Coquart et Bosquet [270], sur 3 000 m, 5 000 m et 10 000 m, les performances prédites étaient surestimées en moyenne de 5 s et 1 s sur 3 000 m et 10 000 m, respectivement, et étaient sous-estimées de 1 s sur 5 000 m. En outre, les résultats de Coquart et al. [24] ont également confirmé la validité de ce nomogramme de Mercier et al. [26] pour prédire des performances sur ces mêmes distances de piste chez les coureurs de sexes masculin et féminin, quelle que soit leur spécialité de référence (*i.e.*, coureur de 3 000 m, 5 000 m ou 10 000 m). Le nomogramme de Mercier et al. [26] est donc considéré comme un outil valide pour prédire avec précision et justesse les performances de la course sur piste chez les coureurs masculins et féminins.

Plus récemment, Vandewalle [27] a développé des nomogrammes basés sur le modèle logarithmique de Péronnet et Thibault [112], au regard du fait que cette relation logarithmique (*i.e.*, régression linéaire entre la vitesse de course et le logarithme $\frac{t_{lim}}{420}$) ne permet pas de prédire les performances sur d'autres distances. Vandewalle [27] a démontré la validité des premier et troisième nomogrammes (le second et le troisième nomogrammes étant identiques dans leur échelle de temps et graphiquement pour des distances allant de 3 000 m à 20 000 m, le second nomogramme s'avère par conséquent également valide) en obtenant des estimations de VMA et de capacité d'endurance aérobie non différentes

significativement ($p > 0,05$) des valeurs calculées à partir de la régression logarithmique de Péronnet et Thibault [112] chez des coureurs élités sur piste (e.g., 1 500 m, 3 000 m, 5 000 m et 10 000 m). Vandewalle a également confirmé graphiquement la validité du troisième nomogramme par l'obtention d'une droite reliant les différentes performances individuelles étudiées pour chaque coureur élite. Cependant, bien que Vandewalle ait confirmé la validité de ces nomogrammes pour estimer la VMA, la capacité d'endurance et la performance, ces nomogrammes n'ont été testés qu'à partir de performances masculines de niveau mondial [27]. Ainsi, leur validité [27] reste à confirmer pour les performances sur piste de niveau plus modeste et chez les athlètes féminines également.

L'objectif de la présente étude était donc de tester la validité et de comparer la justesse (*via* le biais, c'est-à-dire, la proximité de la performance réelle par rapport à la performance prédite, et donc une erreur absolue moyenne ou un biais plus faible, signifie ainsi une plus grande justesse) et la précision (*via* les limites d'agrément : LoA ; c'est-à-dire la proximité des performances prédites, et donc une plus petite distance entre les LoA, signifie une plus grande précision) de trois nomogrammes (*i.e.*, le nomogramme de Mercier et al. [26] et deux nomogrammes de Vandewalle [27]) afin de prédire les performances en course de demi-fond et de fond chez des athlètes françaises au cours de ces dernières années.

2. Méthodes

Tous les bilans officiels de la FFA pour les coureuses âgées de 23 à 34 ans ayant des références chronométriques sur les épreuves de 3 000 m, 5 000 m et 10 000 m piste ont été analysés du 1^{er} janvier 2005 au 31 décembre 2019. Toutes les performances ont été collectées, quel que soit le niveau de performance (*i.e.*, départemental, régional, interrégional, national et international). Les performances des athlètes

ayant couru sur les trois distances au cours d'une période de 12 mois ont été retenues. Les performances sur route, dans des épreuves telles que le 10 km, ont en revanche été exclues et seules les performances sur piste ont été conservées afin de contrôler tout effet du profil de la course (e.g., plat ou vallonné) sur les performances. Une fois les athlètes identifiées, l'allure de course pour chaque distance a été calculée individuellement, et lorsque la vitesse était plus élevée sur une plus longue distance, l'athlète était exclue. En effet, même s'il existe certaines preuves que les coureurs, notamment ceux de longue distance comme le marathon, se heurtent parfois à un « mur » en subissant un ralentissement relatif ou des variations d'allure [307], il ne semble pas courant de courir plus lentement sur une distance plus courte. Cette procédure permettant l'exclusion de données aberrantes a notamment été utilisée dans des travaux antérieurs [24, 25, 227]. Un tel résultat (*i.e.*, vitesse supérieure sur une plus longue distance) peut être la conséquence d'une mauvaise stratégie (ou gestion de l'allure) de course, d'une préparation non spécifique, de mauvaises conditions météorologiques lors de la compétition, ou encore d'une diminution des capacités physiques liée à une potentielle blessure.

Cette étude, comme les autres, a été approuvée par les autorités compétentes (CERSTAPS 2019-22-02-31).

Les performances de 158 femmes ont été prises en considération. Les performances moyennes sur 3 000 m, 5 000 m et 10 000 m étaient respectivement de 11 min 17 s \pm 1 min 20 s (8 min 53 s-15 min 56 s), 19 min 29 s \pm 2 min 20 s (15 min 8 s-27 min 53 s) et 41 min 18 s \pm 5 min 7 s (31 min 53 s-59 min 14 s) (Tableau 20). Ces performances sur 3 000 m, 5 000 m et 10 000 m représentaient, respectivement, 60 \pm 17%, 62 \pm 17% et 59 \pm 18% des records du monde actuels en plein air (*i.e.*, 8 min 6 s par Junxia Wang en 1993, 14 min 11 s par Tirunesh Dibaba en 2008 et 29 min 17 s par Almaz Ayana en 2016).

Pour l'ensemble des courses de demi-fond et de fond étudiées (*i.e.*, 3 000 m, 5 000 m et 10 000 m), les performances ont été prédites individuellement en utilisant le nomogramme de Mercier et al. [26] et les deux nomogrammes de Vandewalle [27].

Pour ce faire, une procédure numérique a été mise en œuvre à l'aide d'un programme MATLAB personnalisé (MATLAB R2016b, *The MathWorks Inc.*, Natick, MA, USA) afin de prédire automatiquement les performances requises à partir de chaque nomogramme et limiter l'erreur humaine. Les coefficients de corrélation intra-classe (ICC) et l'intervalle de confiance à 95% (IC 95%) ont été évalués pour confirmer la cohérence interne de cette méthode [308]. Les résultats sont présentés dans le Tableau 19 et permettent de conclure que cette méthode automatisée est valide.

Tableau 19. Coefficient de corrélation intra-classe (ICC) entre le programme MATLAB et les méthodes manuelles

Distance (m)	Coefficients de corrélation intra-classe (ICC)	Interprétation de la fiabilité [308]	Intervalle de confiance à 95%		
			Limite inférieure	Limite supérieure	
3 000	Préd. <i>via</i> le nomogramme de Mercier et al.	0,991	Très élevée	0,987	0,993
	Préd. <i>via</i> le 1 ^{er} nomogramme de Vandewalle	0,998	Très élevée	0,997	0,998
	Préd. <i>via</i> le 2 ^{ème} nomogramme de Vandewalle	0,876	Élevée	0,834	0,908
5 000	Préd. <i>via</i> le nomogramme de Mercier et al.	0,944	Très élevée	0,925	0,959
	Préd. <i>via</i> le 1 ^{er} nomogramme de Vandewalle	0,750	Élevée	0,673	0,811
	Préd. <i>via</i> le 2 ^{ème} nomogramme de Vandewalle	0,726	Modérée	0,643	0,792
10 000	Préd. <i>via</i> le nomogramme de Mercier et al.	0,998	Très élevée	0,997	0,999
	Préd. <i>via</i> le 1 ^{er} nomogramme de Vandewalle	0,820	Élevée	0,761	0,865
	Préd. <i>via</i> le 2 ^{ème} nomogramme de Vandewalle	0,866	Élevée	0,821	0,901

Note : Calculs réalisés à partir du nomogramme de Mercier et al. [26] et des nomogrammes de Vandewalle [27] chez 158 coureuses sur les épreuves de 3 000 m, 5 000 m et 10 000 m.

Par exemple, pour le nomogramme de Mercier et al. [26], le programme MATLAB charge une matrice à deux colonnes pour chaque prédiction avec les performances réelles des athlètes et numérise le nomogramme classique, en ajoutant un système de coordonnées dans le coin inférieur gauche de l'image (cf. Étape 1, Figure 36, par exemple). Ensuite, l'utilisateur clique sur plusieurs points d'ancrage sur l'axe vertical pour toutes les distances choisies. Cette procédure n'est effectuée qu'une seule fois, et les résultats (*i.e.*, coordonnées des points et temps correspondants sur le nomogramme) sont stockés pour toutes les prédictions ultérieures. Les performances de course réelles sont ensuite transformées en coordonnées précises qui sont calculées à partir des points d'ancrage les plus proches (cf. Étape 2, Figure 36). Les deux coordonnées sont ensuite utilisées pour calculer la pente (a) et l'ordonnée à l'origine (b) de la droite passant par les performances de course réelles (cf. Étape 3, Figure 36). Comme la coordonnée x est connue pour tous les points d'ancrage, la coordonnée y de la performance de course prédite est calculée comme suit : $y = a \cdot x + b$. Cette coordonnée finale correspond précisément au point d'intersection entre la droite passant par la performance de course réelle et l'axe vertical de la distance souhaitée. Le point d'intersection (exprimé en pourcentage de la distance séparant les coordonnées des deux points d'ancrage les plus proches) est finalement transformé en performance de course prédite. Ce processus a été automatisé puisque le programme MATLAB a implémenté chaque ligne de la matrice à deux colonnes avec les 158 performances réelles. L'automatisation permet notamment d'éviter les erreurs possibles lors du processus manuel, comme évoqué dans les travaux de Coquart et al. [24]. Une procédure similaire a été appliquée pour les nomogrammes de Vandewalle [27], à l'exception du fait que les points d'ancrage avaient des coordonnées x différentes pour chaque distance de course.

Prédiction de la performance en course à pied : demi-fond et fond

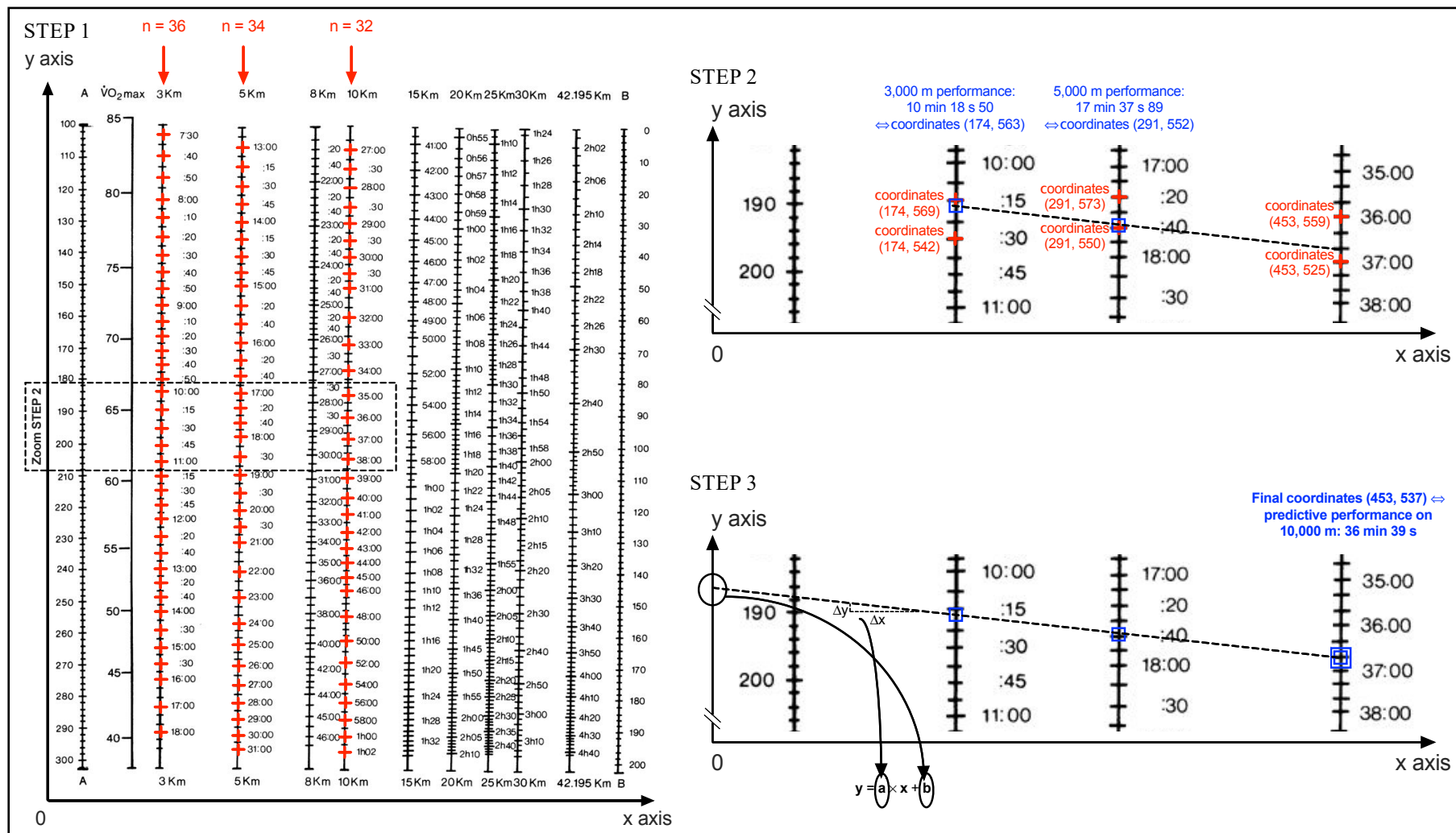


Figure 36. Exemple du processus automatique utilisé pour prédire la performance en course à pied (illustration avec le nomogramme de Mercier et al. [26])

Note : Le processus automatique de prédiction de la performance sur 10 000 m est basé sur la numérisation du nomogramme de Mercier et al. [26].

Étape 1 : Numérisation du nomogramme, adjonction d'un système de coordonnées et représentation de tous les points d'ancrage.

Étape 2 : Vue rapprochée de la partie concernée du nomogramme : intégration des performances de l'athlète transformées en coordonnées en utilisant les points d'ancrage les plus proches.

Étape 3 : Calcul de la pente et de l'intersection de la droite passant par la performance. La performance restante (ici pour le 10 000 m) est prédite à l'aide de l'équation $y = a \cdot x + b$ en transformant les coordonnées finales en performance de course correspondante.

Des méthodes statistiques standards ont été utilisées pour calculer les moyennes et les SD des performances. La distribution normale gaussienne a été vérifiée par le test de Shapiro-Wilk. Étant donné que les données n'ont pas passé le test de normalité, un test de Wilcoxon a été utilisé pour comparer les performances réelles et prédites sur 3 000 m, 5 000 m et 10 000 m. L'amplitude des différences a été évaluée par l'ES et interprétée selon l'échelle de Cohen [301]. L'ES était considéré comme : insignifiant pour $ES < 0,2$; faible pour $0,2 \leq ES < 0,5$; modéré pour $0,5 \leq ES < 0,8$, et important pour $ES \geq 0,8$. La relation entre les performances de course réelles et prédites a été testée avec le test de corrélation non-paramétrique de Spearman. Le coefficient de corrélation a été considéré comme très élevé si $r \geq 0,90$; élevé si $0,70 \leq r \leq 0,89$; modéré si $0,50 \leq r \leq 0,69$; et faible si $0,26 \leq r \leq 0,49$ [309].

Le biais, évaluant la justesse de chaque nomogramme afin de prédire les performances de course réelles, et les LoA à 95% (95% LoA), évaluant la précision des trois nomogrammes [310], ont été calculés selon la méthode de Bland et Altman [311, 312]. Cette méthode s'appuie sur le calcul de la différence moyenne entre deux méthodes de mesure (*i.e.*, le biais) et $\pm 1,96 \times SD$ (*i.e.*, LoA). Le modèle a été considéré comme valide si l'amplitude des différences, évaluée par l'ES, était acceptable, et si le biais était juste et les LoA étaient précises. Le meilleur modèle a été sélectionné sur la base de la plus grande justesse et de la plus grande précision évaluées à partir de la méthode de Bland et Altman [311, 312].

Le seuil de significativité statistique a été fixé à $p < 0,05$, et toutes les analyses ont été effectuées avec le logiciel *Statistical Package for the Social Sciences* (version 21.0, Chicago, USA).

3. Résultats

Les moyennes et les SD des performances de course réelles et prédites pour les 158 coureuses selon chaque nomogramme sont présentés dans le Tableau 20.

Aucune différence significative n'a été constatée entre les performances de course réelles et prédites, quel que soit le nomogramme ($p > 0,05$, Tableau 20). Toutes les performances prédites étaient significativement corrélées avec les performances réelles, avec un coefficient de corrélation très élevé ($p < 0,001$, $r \geq 0,94$; Tableau 20, Figures 37 à 39).

Le biais et les 95% LoA des nomogrammes de Mercier et al. [26] et des nomogrammes de Vandewalle [27] sont présentés dans le Tableau 20 et les Figures 37 à 39.

Tableau 20. Résultats des prédictions à partir du nomogramme de Mercier et al. [26] et des deux nomogrammes de Vandewalle [27] chez 158 coureuses

Distance (m)		Performance Moyenne (min) ± SD	p	Corrélation		Biais (min) ± 95% LoA	Biais (%) ± 95% LoA	Amplitude du biais	
				r	Interprétation			ES	Interprétation
3 000	Perf. Réelle	11,28 ± 1,33	-	-	-	-	-	-	-
	Préd. via le nomogramme de Mercier et al.	11,28 ± 1,37	0,850	0,962*	Très élevée	-0,002 ± 0,699	-0,0 ± 6,2	-0,002	Insignifiante
	Préd. via le 1 ^{er} nomogramme de Vandewalle	11,29 ± 1,37	0,888	0,967*	Très élevée	0,009 ± 0,677	0,1 ± 6,0	0,007	Insignifiante
	Préd. via le 2 ^{ème} nomogramme de Vandewalle	11,28 ± 1,37	0,897	0,966*	Très élevée	0,007 ± 0,678	0,1 ± 6,0	0,005	Insignifiante
5 000	Perf. Réelle	19,49 ± 2,34	-	-	-	-	-	-	-
	Préd. via le nomogramme de Mercier et al.	19,50 ± 2,32	0,996	0,987*	Très élevée	0,005 ± 0,712	0,0 ± 3,7	0,002	Insignifiante
	Préd. via le 1 ^{er} nomogramme de Vandewalle	19,47 ± 2,32	0,387	0,987*	Très élevée	-0,020 ± 0,708	-0,1 ± 3,6	-0,009	Insignifiante
	Préd. via le 2 ^{ème} nomogramme de Vandewalle	19,49 ± 2,32	0,951	0,987*	Très élevée	-0,000 ± 0,716	0,0 ± 3,7	-0,000	Insignifiante
10 000	Perf. Réelle	41,03 ± 5,12	-	-	-	-	-	-	-
	Préd. via le nomogramme de Mercier et al.	41,01 ± 5,25	0,837	0,946*	Très élevée	-0,026 ± 3,634	-0,1 ± 8,9	-0,005	Insignifiante
	Préd. via le 1 ^{er} nomogramme de Vandewalle	41,10 ± 5,34	0,706	0,942*	Très élevée	0,070 ± 3,842	0,2 ± 9,3	0,013	Insignifiante
	Préd. via le 2 ^{ème} nomogramme de Vandewalle	41,06 ± 5,30	0,895	0,941*	Très élevée	0,026 ± 3,828	0,1 ± 9,3	0,005	Insignifiante

Légende : Valeurs moyennes et SD des performances réelles et prédites à partir du nomogramme de Mercier et al. [26] et des deux nomogrammes de Vandewalle [27] chez 158 coureuses ; différence entre les performances réelles et prédites (p), corrélation de Spearman avec les performances réelles (r), biais et SD (min), biais et LoA 95% (min et %), et amplitude du biais pour les épreuves sur piste de 3 000 m, 5 000 m et 10 000 m

Note : Biais = différence entre les performances réelles et prédites ; r = coefficient de corrélation ; 95% LoA = limites d'agrément à 95% ; SD = écart-type ; ES = taille de l'effet. *Corrélation significative à $p < 0,001$

Prédiction de la performance en course à pied : demi-fond et fond

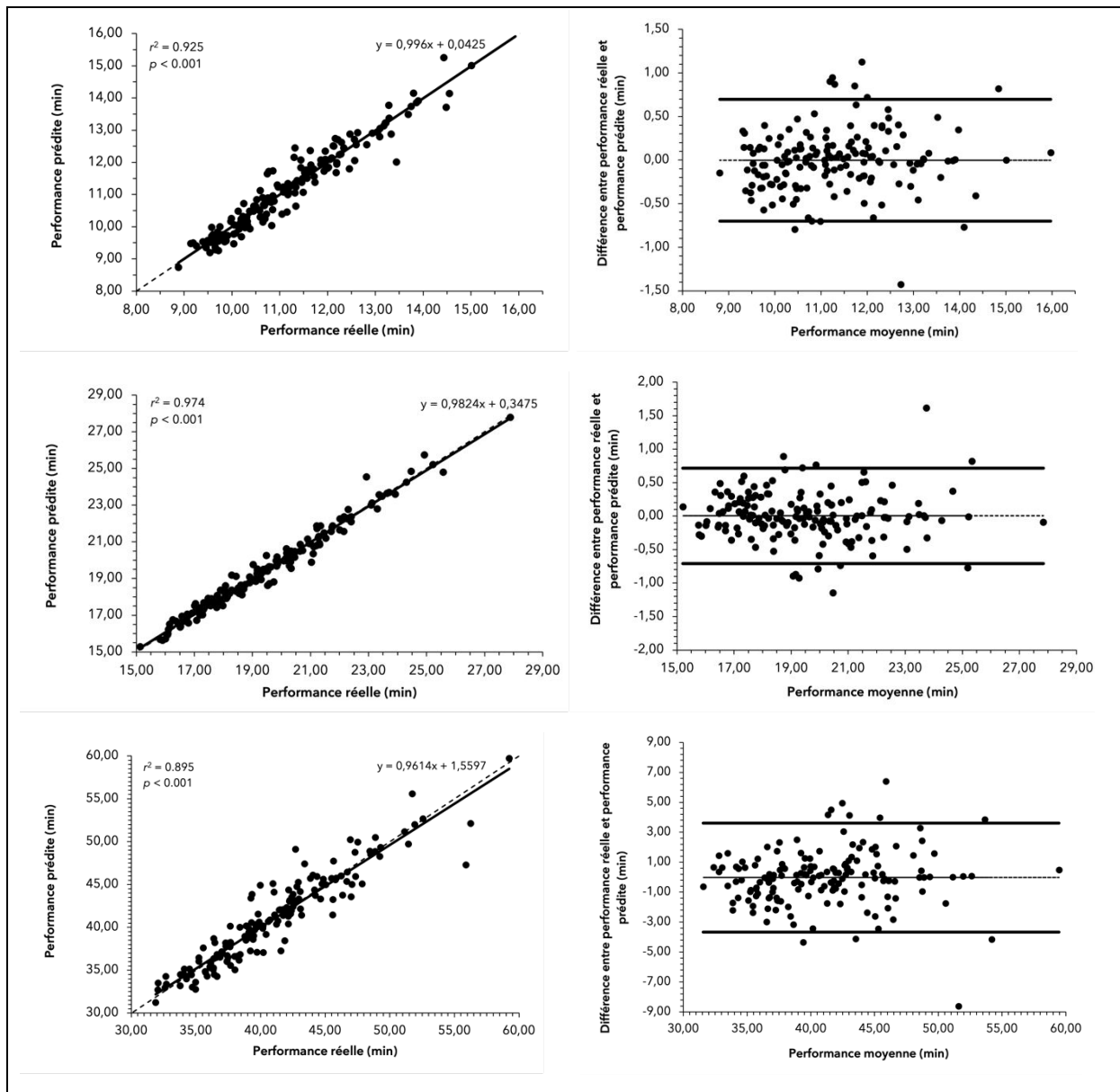


Figure 37. Validité des mesures avec le nomogramme de Mercier et al. [26] pour prédire la performance

Panneau de gauche : relation entre les performances réelles et prédites à partir du nomogramme de Mercier et al. [26] sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La droite en continu est la régression linéaire. r^2 est le coefficient de détermination.

Panneau de droite : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La ligne en pointillés représente le biais et les lignes pleines sont les limites d'agrément à 95%.

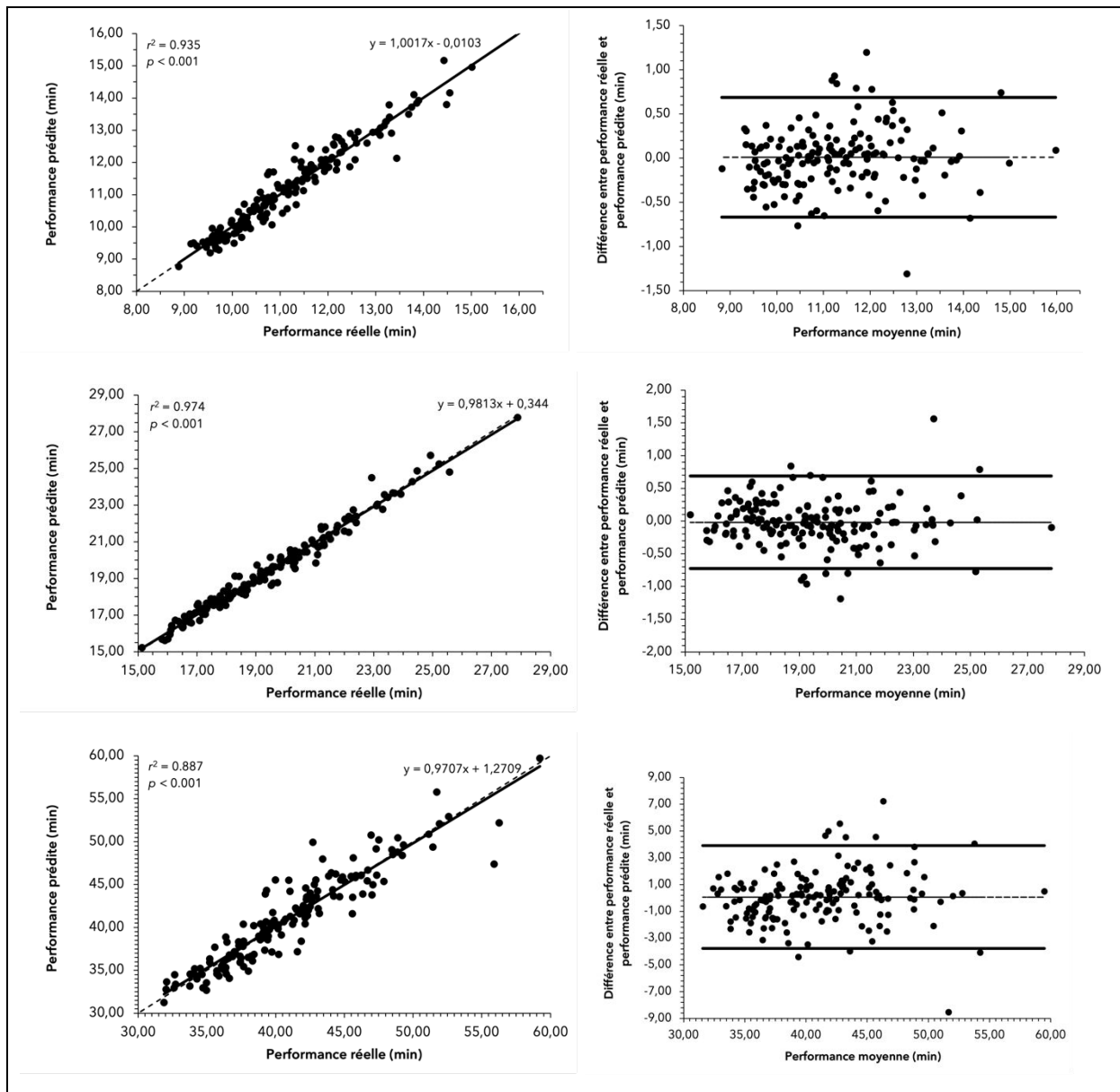


Figure 38. Validité des mesures avec le premier nomogramme de Vandewalle [27] pour prédire la performance

Panneau de gauche : relation entre les performances réelles et prédites à partir du premier nomogramme de Vandewalle [27] basé sur le modèle logarithmique de Péronnet et Thibault [112] sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La droite en continu est la régression linéaire. r^2 est le coefficient de détermination.

Panneau de droite : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La ligne en pointillés représente le biais et les lignes pleines sont les limites d'agrément à 95%.

Prédiction de la performance en course à pied : demi-fond et fond

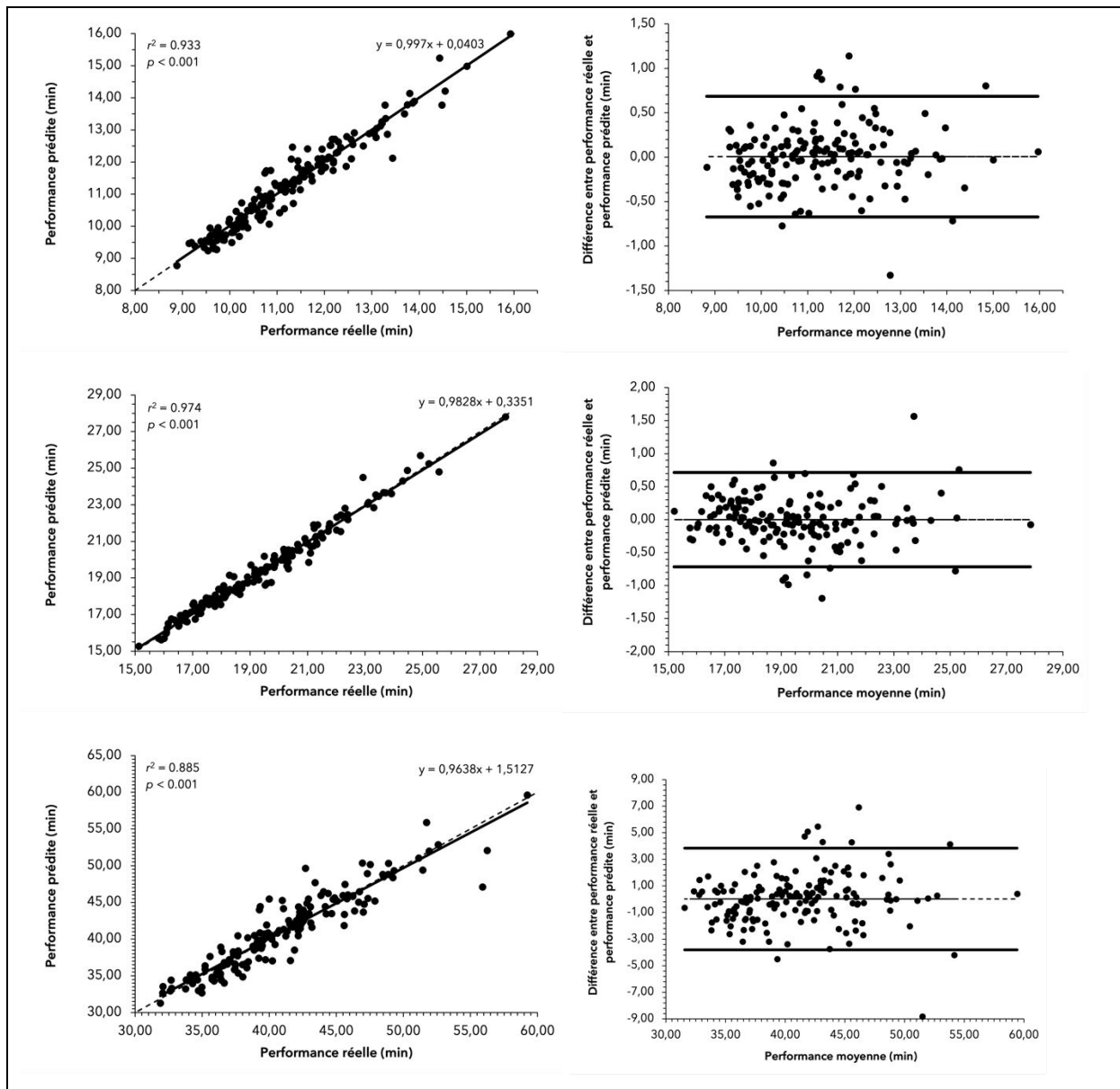


Figure 39. Validité des mesures avec le second nomogramme de Vandewalle [27] pour prédire la performance

Panneau de gauche : relation entre les performances réelles et prédites à partir du second nomogramme de Vandewalle [27] basé sur le modèle logarithmique de Péronnet et Thibault [112] sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La droite en continu est la régression linéaire. r^2 est le coefficient de détermination.

Panneau de droite : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La ligne en pointillés représente le biais et les lignes pleines sont les limites d'agrément à 95%.

Le biais et les LoA 95% obtenus étaient très légèrement inférieurs avec le nomogramme de Mercier et al. [26] par rapport aux nomogrammes de Vandewalle [27] sur le 10 000 m, alors que le premier nomogramme de Vandewalle [27] était très légèrement plus précis (*i.e.*, écart entre les LoA plus faible) par rapport aux autres nomogrammes de prédiction des performances concernant les épreuves du 3 000 m et du 5 000 m (Tableau 20). De plus, une précision de prédiction plus élevée (observée à partir d'un écart réduit entre les 95% LoA) a été obtenue sur le 5 000 m, quel que soit le nomogramme utilisé (Tableau 20). L'amplitude du biais pour les performances prédites était systématiquement insignifiante ($ES \leq 0,013$; Tableau 20).

4. Discussion

L'objectif de cette étude était de tester la validité et de comparer la justesse (*via* le biais) et la précision (*via* les LoA) des prédictions de performances de course de demi-fond et de fond chez des athlètes féminines en utilisant le nomogramme de Mercier et al. [26] et les deux nomogrammes de Vandewalle [27]. Notre principale conclusion est que ces trois nomogrammes peuvent être considérés comme valides, justes et précis pour prédire les performances des athlètes féminines. En effet, conformément aux études précédentes [24, 25, 27, 270], tous les résultats de la présente étude confirment la validité des trois nomogrammes pour prédire les performances en demi-fond et fond sur une distance donnée à partir de performances sur deux autres distances, mais auprès de femmes de tous niveaux. Les résultats ont révélé que dans 95% des prédictions, les différences entre les performances réelles et les performances prédites par le nomogramme de Mercier et al. [26], par le premier nomogramme de Vandewalle [27] et par le second nomogramme de Vandewalle [27], seraient respectivement inférieures ou égales à

$-0,0 \pm 6,2\%$, $0,1 \pm 6,0\%$ et $0,1 \pm 6,0\%$ pour le 3 000 m, $0,0 \pm 3,7\%$, $-0,1 \pm 3,6\%$ et $0,0 \pm 3,7\%$ pour le 5 000 m, et $-0,1 \pm 8,9\%$, $0,2 \pm 9,3\%$ et $0,1 \pm 9,3\%$ pour le 10 000 m.

Cependant, bien que les mesures de précision (*i.e.*, les valeurs de 95% LoA) des prédictions soient proches entre les trois nomogrammes, il a été constaté de légères différences concernant les 95% LoA. En effet, il semble que le premier nomogramme de Vandewalle [27] soit légèrement plus précis (*i.e.*, espacement plus faible entre les 95% LoA) par rapport aux deux autres nomogrammes pour prédire les performances sur le 3 000 m et le 5 000 m, alors que pour l'épreuve du 10 000 m, le nomogramme de Mercier et al. [26] était plus précis. Ces résultats pourraient être expliqués par les différences de conception (*i.e.*, la méthode de construction des modèles) entre les nomogrammes. Le nomogramme de Mercier et al. [26] a été développé à partir de connaissances empiriques des auteurs, alors que les nomogrammes de Vandewalle [27] sont basés sur un modèle logarithmique. De plus, bien que les nomogrammes de Vandewalle [27] aient été construits de manière similaire et basés tous deux sur le modèle de Péronnet et Thibault [112], les distances et les graduations permettant de prédire les performances diffèrent. Par exemple, le premier nomogramme de Vandewalle [27] se focalise sur les distances sur piste qui sont les plus souvent réalisées dans les compétitions internationales (*i.e.*, 1 500 m, 3 000 m, 5 000 m et 10 000 m), alors que le second nomogramme [27] permet de réaliser des prédictions sur un plus large éventail de distances (*i.e.*, 1 000 m-20 000 m). En outre, les graduations sont de 5 s, 10 s et 15 s sur 3 000 m, 5 000 m et 10 000 m, respectivement, pour le premier nomogramme [27], alors qu'elles sont de 10 s, 20 s, 30 s et 1 min pour les efforts de 7-14 min (1 000 m-5 000 m), 14-28 min (2 000 m-10 000 m), 28-55 min (4 000 m-20 000 m) et 55-97 min (10 000 m-20 000 m), respectivement, pour le second nomogramme [27]. Ces variations pourraient expliquer certaines différences de précision de prédiction observées, mais non significatives. Dans la même optique et compte tenu des

moindres différences obtenues sur 10 000 m, il est probable que l'échelle de prédiction des performances définie par Mercier et al. [26] (*i.e.*, échelle variant avec le temps pour une distance donnée avec, par exemple, une graduation de 15 s de 27 min à 31 min sur 10 000 m ; une graduation de 20 s de 31 min à 38 min...) soit plus proche des performances réelles de l'épreuve du 10 000 m que celles prédites par les deux nomogrammes de Vandewalle [27].

D'autre part, en raison de corrélations supérieures (*i.e.*, $r \geq 0,987$) et de LoA 95% inférieures sur le 5 000 m (*i.e.*, LoA 95% variant que de $\pm 3,6-3,7\%$ seulement), il apparaît que la prédiction est meilleure sur cette distance (quel que soit le nomogramme) par rapport au 3 000 m (*i.e.*, $0,962 \leq r \leq 0,967$; 95% LoA variant de $\pm 6,0-6,2\%$) et au 10 000 m (*i.e.*, $0,706 \leq r \leq 0,895$; 95% LoA variant de $\pm 8,9-9,3\%$). Ce résultat semble être en accord avec l'étude de Coquart et Bosquet [270], ayant montré que les LoA 95% variaient de $\pm 6,4\%$, $\pm 3,5\%$, et $\pm 10,4\%$ par rapport aux moyennes des performances réelles et prédites sur 3 000 m, 5 000 m, et 10 000 m, respectivement, avec le nomogramme de Mercier et al. [26]. Cette justesse et cette précision accrues dans les résultats de prédiction pourraient être liées au fait que la distance prédite a été obtenue par interpolation (*i.e.*, lorsque la distance est délimitée par des distances plus courte et plus longue ; performance prédite sur le 5 000 m à partir des performances réelles du 3 000 m et du 10 000 m dans l'étude actuelle), comme recommandé par Mercier et al. [26], plutôt que par extrapolation. En effet, d'un point de vue physiologique, les performances en course à pied sont déterminées par des paramètres tels que le $\dot{V}O_{2max}$ (*i.e.*, le facteur le plus discriminant de la performance dans les courses de demi-fond ou de fond lorsque la population est hétérogène), l'économie de course et la capacité d'endurance aérobie (qui, en tant que prédicteur, a une part de plus en plus importante au fur et à mesure que la distance de la course augmente, telle que la course de fond comme le 10 000 m) [108, 251]. En prédisant les performances des courses de distance

intermédiaire par interpolation (comme le 5 000 m) plutôt que par extrapolation, le $\dot{V}O_{2max}$ et la capacité d'endurance aérobie peuvent être davantage pris en considération, et permettre ainsi l'obtention de prédictions plus justes et plus précises. Par conséquent, ces résultats soutiennent cette recommandation, qui a déjà été confirmée dans d'autres études [24, 25, 27, 265, 270].

De plus, l'automatisation des données à l'aide d'un programme MATLAB semble avoir permis d'éviter de potentielles erreurs liées au processus manuel, comme décrit dans les travaux de Coquart et al. [24]. Les biais ($< 1\%$) et LoA à 95% ($\leq 9,3\%$) obtenus dans cette étude étaient légèrement meilleurs que ceux des études réalisées antérieurement [24, 25], suggérant ainsi que la méthode MATLAB automatisée ait pu limiter les erreurs humaines et ainsi accroître la justesse et la précision des prédictions.

Si l'on s'intéresse davantage aux travaux de Tabben et al. [25] et Coquart et al. [24], ces derniers ont également pris en considération le niveau de performance et/ou la spécialité athlétique, respectivement, afin de valider et de comparer la justesse et la précision des performances prédites par le nomogramme de Mercier et al. [26]. Ceci pourrait notamment expliquer certaines différences de résultats et d'interprétations, étant donné que Mercier et al. [26] ont suggéré de meilleures prédictions, par exemple, lorsque les coureurs étaient spécifiquement préparés sur les distances à partir desquelles les prédictions étaient réalisées. Cette hypothèse testée par Coquart et al. [24] n'a cependant pas été confirmée.

Si l'étude actuelle s'est focalisée sur un large échantillon de performances, exclusivement féminines (quel que soit leur niveau de performance), il est important de souligner que jusqu'à présent, ces dernières ont été très peu étudiées.

Les principales limites de cette étude sont possiblement liées au fait que les performances de course réelles n'aient pas été toutes réalisées de manière

exhaustive (e.g., course tactique) ou n'aient pas été réalisées dans des conditions environnementales similaires (e.g., mêmes conditions climatiques et altitude), dans des conditions physiques (e.g., $\dot{V}O_{2max}$, économie de course, capacité d'endurance aérobie...) et psychologiques (e.g., motivation, anxiété et stress...) identiques, et en utilisant le même terrain d'opposition et/ou la même stratégie de course. En effet, comme mentionné précédemment, l'analyse a été limitée aux performances sur piste afin de pouvoir contrôler l'effet du profil de la course et limiter les biais. Toutefois, il est probable que les variables susmentionnées aient pu affecter négativement la justesse et la précision des prédictions. Cependant, malgré cette limitation potentielle, le constat est tel que les performances prédites à l'aide des trois nomogrammes sont valides, justes et précises. Il est très intéressant de noter que ces nomogrammes, et plus particulièrement le nomogramme de Mercier et al. [26] conçu dans les années 1980, soient encore d'actualité pour prédire les performances réelles malgré l'évolution des performances constatée lors des dernières décennies [313].

L'un des objectifs de cette étude était d'analyser l'ensemble des performances des coureuses de demi-fond et de fond de nationalité française, quel que soit leur niveau de performance. S'il a été démontré que les nomogrammes de Vandewalle [27], permettent de prédire les performances des coureurs de niveau élite sur piste (1 500 m, 3 000 m, 5 000 m et 10 000 m), la présente étude confirme ces résultats chez des coureurs non élites. Les performances de la présente étude auraient pu être classées en quatre niveaux, par exemple : départemental, régional, national et international. Cependant, ce choix aurait limité la taille des échantillons, et certains auraient été trop faibles pour envisager des analyses plus avancées ($n = 1$, 1 et 4 pour les athlètes de niveau international, respectivement sur les épreuves de 3 000 m, 5 000 m et 10 000 m).

Enfin, en raison de la rareté des études sur les femmes et/ou les coureurs non-élites, la comparaison de performances prédites chez les femmes pourrait être effectuée entre les nomogrammes et d'autres méthodes de prédiction telles que les concepts de vitesse critique de loi de puissance, ou encore des méthodes basées sur de l'IA. De même, du point de vue de l'entraînement ou de la performance, des études futures pourraient également être menées sur la prédiction des performances sur route (e.g., 10 km, semi-marathon et marathon) chez les femmes, car il semble que seules les performances masculines aient été étudiées sur de longues distances [265] avec le nomogramme de Mercier et al. [26].

En outre, au-delà de la prédiction, les nomogrammes peuvent être utilisés pour estimer des indicateurs de la condition physique tels que le $\dot{V}O_{2max}$ [314], et la capacité d'endurance aérobie [26, 27], ce qui contribuerait théoriquement à optimiser et à adapter les programmes d'entraînement des athlètes et des entraîneurs.

5. Conclusion

Les résultats de la présente étude indiquent la validité du nomogramme de Mercier et al. [26] et des deux nomogrammes de Vandewalle [27] pour prédire avec une grande justesse et une plus grande précision les performances individuelles des coureuses sur piste sur les épreuves de 3 000 m, 5 000 m et 10 000 m à partir de performances sur deux autres distances. Les résultats ne montrent aucune différence significative entre les performances réelles et prédites chez les athlètes féminines, quel que soit le nomogramme utilisé. Toutefois, bien que les prédictions de ces nomogrammes soient similaires (quelle que soit la distance), le biais et les LoA à 95% légèrement inférieurs obtenus avec le nomogramme de Mercier et al. [26] entre les performances réelles et prédites sur 3 000 m notamment, lui confèrent un léger avantage pratique (non significatif) pour calibrer les sessions

Prédiction de la performance en course à pied : demi-fond et fond

d'entraînement et définir les limites potentielles des capacités d'un athlète afin d'optimiser la performance.

CHAPITRE III. PRÉDICTION DES PERFORMANCES EN COURSE DE DEMI-FOND ET DE FOND VIA UNE MODÉLISATION LINÉAIRE DISTANCE-TEMPS, UNE LOI PUISSANCE ET UN NOMOGRAMME

❖ Publication scientifique (cf. Annexe 4)

- **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. *Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram. Journal of Strength and Conditioning Research*. Soumis.

1. Introduction

Pour prédire la performance, en plus de certains nomogrammes [26, 27], plusieurs modèles théoriques comme le concept de vitesse critique [110] ou la loi de puissance [30] peuvent être utilisés.

En effet, la modélisation de la relation entre le temps (ou la vitesse) et la distance est un moyen courant et relativement facile pour prédire la performance [29, 31, 273, 315]. Plusieurs modèles basés sur le concept de vitesse critique ont été proposés [29, 101, 109, 114, 273], mais le modèle original est le MLDT [110], qui a été présenté antérieurement (cf. première partie, chapitre III).

La reformulation mathématique de ce modèle permet de prédire la performance à partir de la simple équation suivante :

$$T_{lim} = \frac{(D_{lim} - CDA)}{V_{crit}} \quad (4)$$

où T_{lim} est le temps jusqu'à épuisement (en min) pour une distance donnée (D_{lim}), D_{lim} est la distance limite (en m), CDA est la capacité de distance en anaérobie (en m) et V_{crit} est la vitesse critique (en $m \cdot s^{-1}$).

Cependant, cette équation n'est utilisable que pour des efforts compris entre 3,5 et 35 min [153, 275].

D'autre part, la relation distance-temps peut également être modélisée à partir de lois de puissance individuelles, dont l'équation est la suivante :

$$D_{lim} = k \cdot T_{lim}^g \quad (5)$$

où k est une constante et l'exposant g est un indice de capacité d'endurance [31, 32, 278, 279], qui peuvent être déterminés par une analyse non linéaire des moindres carrés. Le modèle de Kennelly [30] décrit avec précision les relations entre T_{lim} et D_{lim} en particulier chez les coureurs d'endurance de niveau élite [31, 32], et a pour avantage d'être indépendant de la plage de temps d'épuisement. Cependant, la validité de l'application du modèle de loi de puissance de Kennelly [30] permettant de prédire les performances de course à pied chez les coureurs non-élites n'a pas été vérifiée [278].

Comme montré dans l'étude précédente (*cf.* deuxième partie, chapitre II), l'automatisation à l'aide d'un programme MATLAB de la prédiction de la performance à l'aide du nomogramme de Mercier et al. [26] est une méthode valide chez les femmes de tous niveaux. Néanmoins, cette méthode prédictive basée sur les connaissances empiriques de ces concepteurs n'est peut-être pas la plus précise, et nécessite d'être confrontée à des modèles théoriques comme le concept de vitesse critique et la loi de puissance.

L'objectif de cette étude était donc de tester la validité et de comparer la justesse et la précision de trois méthodes, dont le MLDT, la loi de puissance et le nomogramme de Mercier et al. [26]. L'hypothèse émise était que les trois modèles seraient valides et que le modèle dépendant de la plage de temps d'épuisement (*i.e.*, MLDT) serait le moins exact et le moins précis des trois.

2. Méthodes

Pour cette étude, nous avons donc repris les données collectées lors l'étude précédente (cf. deuxième partie, chapitre II) portant sur la prédiction des performances en course de demi-fond et de fond *via* l'utilisation de trois nomogrammes (*i.e.*, le nomogramme de Mercier et al. [26] et deux nomogrammes de Vandewalle [27]).

Comme présenté dans l'étude précédente les performances (sur 3 000, 5 000 et 10 000 m) de 158 femmes ont été analysées.

Pour toutes les épreuves de course à pied ciblées (*i.e.*, 3 000 m, 5 000 m et 10 000 m), les performances ont été prédites individuellement en utilisant le MLDT [29], la loi de puissance [30, 31] et le nomogramme de Mercier et al. [26].

Pour la première méthode, la relation distance-temps a été modélisée à l'aide du MLDT [29]. Deux performances de course ont été sélectionnées et l'équation (4) a été utilisée pour déterminer la troisième performance.

Pour la deuxième méthode, la loi de puissance a été calculée par des régressions non linéaires des moindres carrés entre le logarithme de T_{lim} et le logarithme de D_{lim} .

Par conséquent, si $D_{lim} = k \cdot T_{lim}^g$, le logarithme de D_{lim} est égal à :

$$\ln(D_{lim}) = \ln(k) + g \ln(T_{lim}) \quad (6)$$

$$\text{où } T_{lim} = e^{\ln(T_{lim})}$$

$$T_{lim} = e^{\frac{\ln(\frac{D_{lim}}{k})}{g}} \quad (7)$$

Les valeurs de g et de k sont alors calculées à partir des valeurs de D_{lim} correspondant à T_{lim} de la première et de la deuxième performance, sans tenir compte des valeurs de T_{lim} et D_{lim} de la troisième performance de course à pied.

Les valeurs de g et de k ont été générées à l'aide de MATLAB (MATLAB R2016b, *The MathWorks Inc.*, Natick, MA, USA). Un modèle de série de puissance à un facteur tel que donné par $y = a \cdot x^b$ a été fourni par la boîte à outils « *Curve Fitting* », et une fonction *fit* pour ajuster les modèles de série de puissance aux données a été utilisée (*i.e.*, *fit(time,distance, 'power1')*). La boîte à outils a calculé des points de départ optimisés pour les modèles de série de puissance, sur la base de l'ensemble des données actuelles.

Enfin, pour prédire la performance avec le nomogramme de Mercier et al. [26], le processus est décrit dans l'étude (*cf.* deuxième partie, chapitre II).

Des méthodes statistiques standards ont été utilisées pour calculer les moyennes et les SD des performances. La distribution normale gaussienne a été vérifiée par le test de Shapiro-Wilk. Étant donné que les données n'ont pas passé le test de normalité, un test de Wilcoxon a été utilisé pour comparer les performances réelles et prédites sur 3 000 m, 5 000 m et 10 000 m. L'amplitude des différences a été évaluée par l'ES et interprétée selon l'échelle de Cohen [301]. La relation entre les performances de course réelles et prédites a été testée avec le test de corrélation non-paramétrique de Spearman, et interprétée par l'échelle de Munro [309].

Le biais (pour la justesse) et les 95% LoA (pour la précision) ont été calculés selon la méthode de Bland et Altman [311, 312], laquelle s'appuie sur le calcul de la différence moyenne entre deux méthodes de mesure (*i.e.*, le biais) et $\pm 1,96 \times SD$ (*i.e.*, LoA).

Le seuil de significativité statistique a été fixé à $p < 0,05$, et toutes les analyses ont été effectuées avec le logiciel *Jeffrey's Amazing Statistics Program* (JASP Team, version 0.14.1, *computer software*).

3. Résultats

Les IC 95% pour le biais des trois méthodes prédites sont présentés dans le Tableau 21. Par rapport au nomogramme de Mercier et al. [26] et à la loi de puissance, les prédictions avec la méthode du MLDT semblent moins précises, car l'IC est plus important (*i.e.*, prédictions plus éloignées de la réalité : marge d'erreur supérieure).

Les moyennes et les SD des performances de course réelles et prédites pour chaque méthode sont présentés dans le Tableau 21.

Aucune différence significative n'a été constatée entre les performances réelles et prédites avec la loi de puissance ou le nomogramme ($p > 0,05$, Tableau 21), à la différence du MLDT ($p < 0,01$, Tableau 21). Cependant, quelles que soient la méthode et la distance choisies, l'amplitude du biais était insignifiante ou faible ($ES \leq 0,293$, Tableau 21).

Toutes les performances prédites étaient significativement corrélées avec les performances réelles, avec des coefficients de corrélation élevés ou très élevés ($p < 0,001$, $r \geq 0,895$, Tableau 21 et Figures 37, 40 et 41).

Le biais $\pm 95\%$ LoA du MLDT, de la loi de puissance, et du nomogramme de Mercier et al. [26] sont présentés dans le Tableau 21 et les Figures 37, 40 et 41.

Dans l'ensemble, la justesse (observée *via* le biais) et la précision (observée *via* l'espacement entre les 95% LoA) étaient très légèrement supérieures avec le

nomogramme de Mercier et al. [26] par rapport à la loi de puissance, et étaient supérieures avec le nomogramme de Mercier et al. [26] par rapport au MLDT (Tableau 21). De plus, une meilleure précision de prédiction (observée à partir des LoA à 95% plus rapprochées) a été constatée sur le 5 000 m, quelle que soit la méthode (Tableau 21).

Tableau 21. Résultats des prédictions à partir du modèle linéaire distance-temps (MLDT), de la loi de puissance et du nomogramme de Mercier et al. [26] chez 158 coureuses

Distance (m)	Modèle	Performance moyenne (min) ± SD	p	Corrélation		Biais (min) ± 95% LoA	Biais (%) ± 95% LoA	IC 95% du biais (min)		Amplitude du biais	
				r	Interprétation			Limite inférieure	Limite supérieure	ES	Interprétation
3 000	Perf. réelle	11,28 ± 1,33									
	Pred. via MLDT	10,88 ± 1,40	< ,001	0,895*	Élevée	-0,399 ± 1,179	-3,7 ± 10,8	-0,445	-0,290	-0,293	Faible
	Pred. via loi de puissance	11,27 ± 1,37	0,782	0,962*	Très élevée	-0,007 ± 0,715	-0,1 ± 6,3	-0,060	0,450	-0,005	Insignifiante
	Pred. via nomogramme	11,28 ± 1,37	0,845	0,962*	Très élevée	-0,002 ± 0,699	0,0 ± 6,2	-0,550	0,045	-0,002	Insignifiante
5 000	Perf. réelle	19,49 ± 2,34									
	Pred. via MLDT	19,78 ± 2,38	< ,001	0,984*	Très élevée	0,285 ± 0,842	1,4 ± 4,3	0,210	0,320	0,121	Insignifiante
	Pred. via loi de puissance	19,50 ± 2,32	0,752	0,987*	Très élevée	0,010 ± 0,722	0,1 ± 3,7	-0,045	0,060	0,005	Insignifiante
	Pred. via nomogramme	19,50 ± 2,32	0,988	0,987*	Très élevée	0,005 ± 0,712	0,0 ± 3,7	-0,050	0,050	0,002	Insignifiante
10 000	Perf. réelle	41,03 ± 5,12									
	Pred. via MLDT	40,03 ± 4,94	< ,001	0,962*	Très élevée	-0,997 ± 2,948	-2,5 ± 7,4	-1,105	-0,720	-0,199	Insignifiante
	Pred. via loi de puissance	40,99 ± 5,26	0,783	0,946*	Très élevée	-0,040 ± 3,636	-0,1 ± 8,9	-0,280	0,210	-0,008	Insignifiante
	Pred. via nomogramme	41,01 ± 5,25	0,837	0,946*	Très élevée	-0,026 ± 3,634	-0,1 ± 8,9	-0,270	0,210	-0,005	Insignifiante

Légende : Valeurs moyennes et écart-type (SD) des performances réelles et prédites à partir du modèle linéaire distance-temps (MLDT), de la loi de puissance et du nomogramme de Mercier et al. [26] chez 158 coureuses ; différence entre les performances réelles et prédites (p), corrélation de Spearman avec les performances réelles (r), biais et LoA 95% (min et %), IC 95% du biais (min), et amplitude du biais (effet de taille) pour les épreuves sur piste sur 3 000 m, 5 000 m et 10 000 m

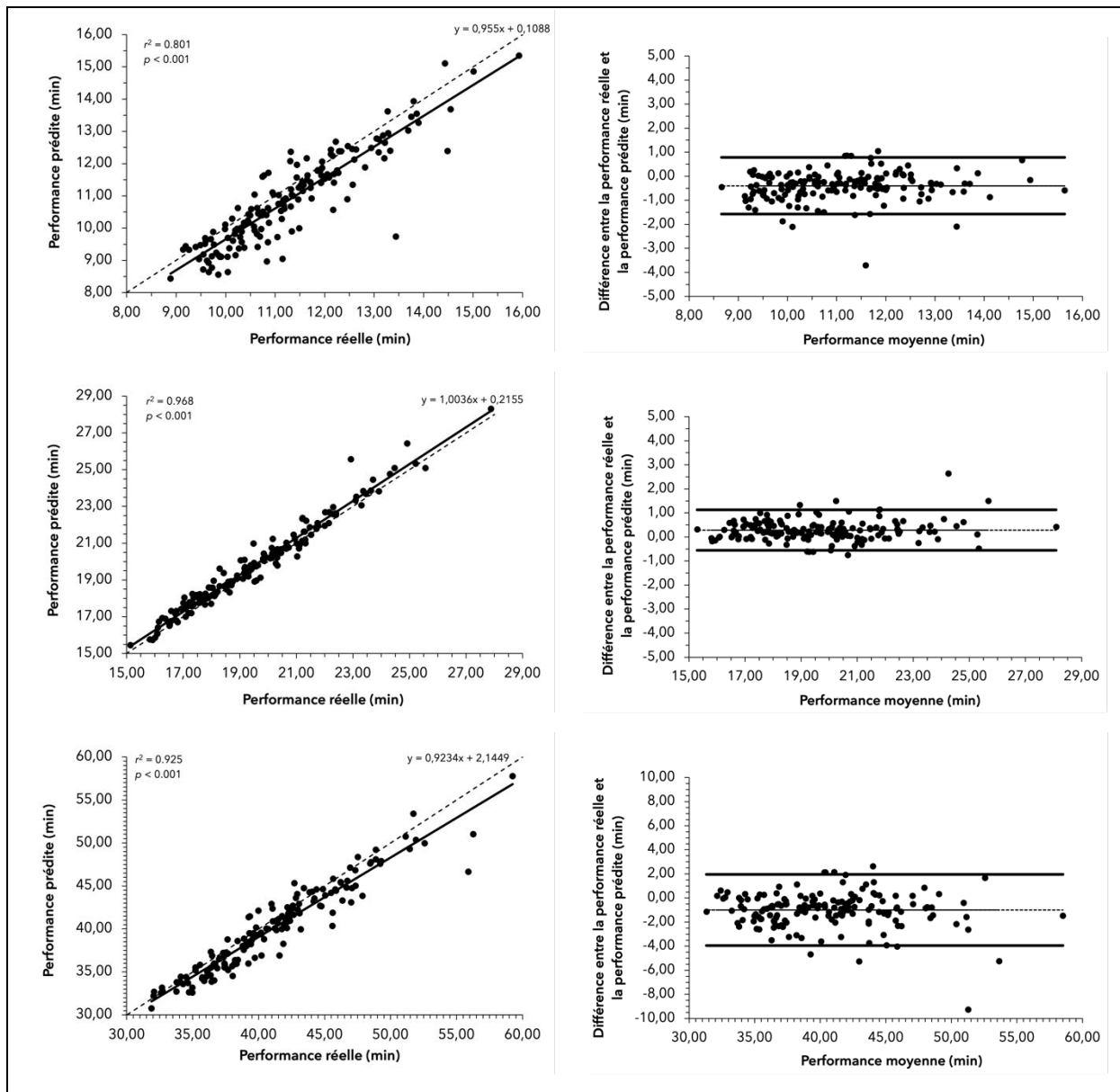


Figure 40. Validité des mesures avec le modèle linéaire distance-temps (MLDT) pour prédire la performance

Panneau de gauche : relation entre les performances réelles et prédites avec le MLDT sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La droite continue représente la régression linéaire. r^2 est le coefficient de détermination.

Panneau de droite : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. Les lignes en pointillés représentent le biais et les lignes pleines sont les limites d'agrément à 95%.

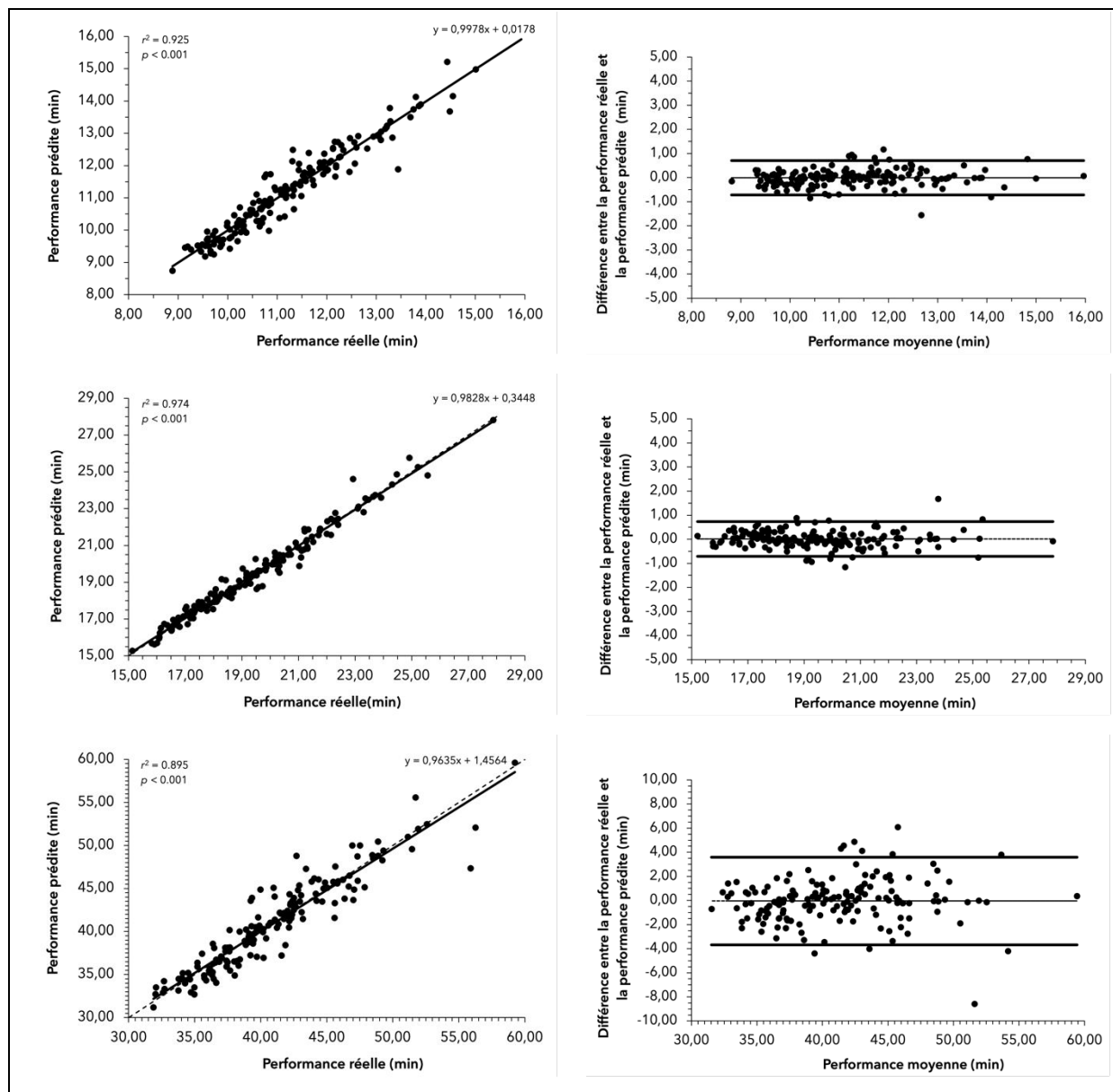


Figure 41. Validité des mesures avec la loi de puissance pour prédire la performance

Panneau de gauche : relation entre les performances réelles et prédites avec la loi de puissance sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. La droite continue représente la régression linéaire. r^2 est le coefficient de détermination.

Panneau de droite : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m (en haut), 5 000 m (au milieu) et 10 000 m (en bas) chez 158 coureuses. Les lignes en pointillés représentent le biais et les lignes pleines sont les limites d'agrément à 95%.

4. Discussion

L'objectif de cette présente étude était de tester la validité et de comparer la justesse et la précision des prédictions des performances en course de demi-fond et de fond chez des athlètes féminines en utilisant le MLDT, la loi de puissance et le nomogramme de Mercier et al. [26].

La principale conclusion est que ces trois méthodes peuvent être considérées comme valides pour prédire les performances des athlètes féminines. En effet, comme le montre le Tableau 21, l'amplitude des différences a été considérée comme faible, voire insignifiante. De plus, les résultats ont révélé que pour 95% des prédictions, les différences entre les performances réelles et prédites avec le MLDT, avec la loi de puissance et avec le nomogramme de Mercier et al. [26], respectivement, étaient inférieures ou égales à $-3,7 \pm 10,8\%$, $-0,1 \pm 6,3\%$ et $0,0 \pm 6,2\%$ sur le 3 000 m, $1,4 \pm 4,3\%$, $0,1 \pm 3,7\%$ et $0,0 \pm 3,7\%$ sur le 5 000 m, et $-2,5 \pm 7,4\%$, $-0,1 \pm 8,9\%$ et $-0,1 \pm 8,9\%$ sur le 10 000 m. La comparaison de ces différences peut laisser suggérer que l'outil le plus juste et le plus précis est le nomogramme de Mercier et al. [26] (*i.e.*, biais plus faible et écart réduit entre les 95% de LoA) suivi par la loi de puissance et enfin le MLDT, bien que ce dernier soit le plus cité dans la littérature [31]. En effet, malgré la proximité des prédictions du nomogramme de Mercier et al. [26] et de la loi de puissance, les différences sont plus importantes avec le MLDT. La valeur d'IC 95% pour le biais [316] a également renforcé le fait que le MLDT était le modèle le moins précis. En accord avec les études précédentes, les résultats ont confirmé la validité des trois méthodes [24, 25, 31, 32, 270] et la plus faible précision du MLDT pour prédire les performances en demi-fond et fond des athlètes féminines. En effet, bien qu'aucune étude n'ait comparé la loi de puissance et le nomogramme, des études réalisées antérieurement ont montré que la loi de puissance était plus précise que le MLDT pour prédire les performances [31, 32,

279], et que le nomogramme était également meilleur que le MLDT [270]. Ces résultats pourraient s'expliquer par des différences dans la méthode de construction des modèles. En effet, si le nomogramme de Mercier et al. [26] a été développé à partir de connaissances empiriques des auteurs, la loi de puissance a quant à elle été modélisée comme suit $y = a \cdot x^b$. Il s'agit d'une relation entre le logarithme de T_{lim} et le logarithme de D_{lim} calculée par des régressions non linéaires des moindres carrés [30, 31]. Le MLDT a été modélisé par $y = a \cdot x + b$ (fonction affine), soit une relation linéaire entre D_{lim} et T_{lim} [29], en particulier lorsque les valeurs de T_{lim} sont comprises entre 3,5 et 35 min [153, 275]. Au-delà de cet intervalle de T_{lim} , la relation entre T_{lim} et D_{lim} n'est pas linéaire. Par conséquent, la prédiction du temps peut être surestimée (*i.e.*, prédiction meilleure que la réalité) lorsque des distances hors intervalle et supérieures à 35 min d'effort par exemple sont utilisées, ce qui expliquerait la plus faible justesse et la plus faible précision des prédictions pour les performances sur 10 000 m dans cette étude. Les résultats (*i.e.*, validité, justesse et précision) pour prédire une performance à partir de deux autres performances, présentées ici semblent en accord avec les résultats des études antérieures [32, 278, 279]. Cependant, il est bien évident que la justesse et/ou la précision des performances prédites peuvent être amenées à varier selon le nombre de performances de course à pied prises en considération dans les modèles de prédiction [31, 261].

De plus, comme évoqué dans le chapitre précédent (*cf.* deuxième partie, chapitre II) pour le nomogramme de Mercier et al. [26], les prédictions étaient plus précises sur le 5 000 m (quelle que soit la méthode) avec des LoA à 95% moins espacées par rapport à celles du 3 000 m et du 10 000 m (Tableau 21). Cette observation, en accord avec des études antérieures [24, 25, 31, 265, 270], suggère qu'il est préférable de prédire par interpolation qu'extrapolation. En effet, comme expliqué précédemment (dans le chapitre II), la prédiction d'une performance sur une

distance intermédiaire à l'aide de deux distances éloignées permet de mieux tenir compte des déterminants physiologiques de la performance en course à pied.

Pour toutes les méthodes comparées, les prédictions ont été réalisées à l'aide de programmes MATLAB (pour le nomogramme de Mercier et al. [26]) et/ou de formules mathématiques assistés par ordinateur, éliminant ainsi le risque d'erreur humaine. Par conséquent, les trois méthodes testées ont pu être comparées puisqu'elles intégraient toutes un processus informatisé et qu'aucune ne dépendait d'une approche manuelle. Cette automatisation des données peut alors être considérée comme l'un des points forts de notre étude.

Au regard des autres études, certaines différences dans les résultats et interprétations pourraient être expliquées par les décisions prises quant à l'intégration de la spécialité et/ou du niveau de performance des athlètes [24, 25]. En effet, les études antérieures portant sur la prédiction de la performance en course *via* des concepts de loi de puissance et de vitesse critique ont principalement été validées chez les coureurs masculins de haut-niveau. À la différence, le nomogramme de Mercier et al. [26] n'a pas encore été testé chez des athlètes féminines de haut-niveau. D'autres études, notamment chez les femmes, sont nécessaires pour vérifier si les prédictions varient en fonction du niveau de performance (e.g., si la précision des prédictions est supérieure lorsque le niveau de performance est élevé). À l'exception de l'étude de Coquart et al. [24] utilisant le nomogramme de Mercier et al. [26], les performances féminines ont été relativement peu étudiées comme mentionné dans la première étude présentée (*cf.* deuxième partie, chapitre I) [313]. Une compréhension plus approfondie de leurs performances pourrait être d'un grand intérêt pour les athlètes et les entraîneurs au regard du fait que les stratégies d'allure peuvent différer entre les sexes lors des compétitions de course à pied [20, 205], par exemple.

Comme l'étude précédente, et celles à venir, les principales limites sont possiblement liées au fait que les performances de course réelles n'aient pas été réalisées de manière exhaustive pour une ou plusieurs distances donnée(s) ou n'aient pas été réalisées dans des conditions physiologiques, psychologiques ou environnementales comparables.

Cependant, malgré cette limitation potentielle, le constat est tel que les performances prédites à l'aide des trois approches sont valides, avec un très léger avantage pour le nomogramme de Mercier et al. [26].

Enfin, des comparaisons de prédiction de performances pourraient être faites en utilisant d'autres méthodes de prédiction basées sur l'IA. Par exemple, l'apprentissage supervisé a récemment montré des résultats prometteurs dans la prédiction de résultats sportifs [317], donc pourquoi ne pas envisager de la prédiction de la performance en course à pied.

5. Conclusion

Les résultats de la présente étude confirment la validité des trois approches (*i.e.*, MLDT, loi de puissance et nomogramme de Mercier et al. [26]) pour prédire les performances individuelles des coureuses sur les distances de 3 000 m, 5 000 m et 10 000 m à partir de performances sur deux autres distances. Le MLDT semble être le modèle le moins juste et le moins précis. Bien que les LoA à 95% du nomogramme et de la loi de puissance puissent être considérées comme similaires, le biais légèrement inférieur obtenu avec le nomogramme, entre les performances réelles et prédites quelle que soit la distance, lui confère un léger avantage pratique pour calibrer les sessions d'entraînement et définir les limites potentielles de la vitesse d'un athlète afin d'optimiser la performance.

CHAPITRE IV. PRÉDICTION DES PERFORMANCES SUR 3000 M STEEPLE À L'AIDE D'UNE SIMPLE ÉQUATION

❖ Publication scientifique (cf. Annexe 5)

- Barbry A, **Lerebourg L**, Racil G, Jlid MC, Coquart JB. *Prediction of the 3000-m steeplechase performance using a simple equation. Research Quarterly for Exercise and Sport*. Soumis.

1. Introduction

Les études précédentes ont montré qu'il était possible de prédire une performance en course à pied à l'aide d'au moins deux autres performances maximales. Cependant, la répétition d'efforts exhaustifs peut parfois (s'ils sont mal planifiés) perturber la préparation de l'athlète, et nuire à l'optimisation de sa performance lors de la compétition principale. Des méthodes prédictives incluant un nombre limité de tests exhaustifs peuvent donc être à privilégier. D'autre part, il apparaît que dans des épreuves de courses à pied caractérisées par le franchissement d'obstacles, telles que celles de *steeple-chase*, la prédiction de performance ne semble pas avoir été étudiée et reste, par conséquent, très limitée.

Le 3 000 m *steeple-chase* est une épreuve de demi-fond relativement technique au regard des 4 barrières et de la rivière (fosse avec de l'eau) à franchir qui sont réparties sur le tour de piste (*i.e.*, 35 obstacles au total) [77]. La principale différence réglementaire, d'ordre technique, entre les sexes est la hauteur des barrières qui est de 0,914 m et 0,762 m pour les hommes et les femmes, respectivement [77].

Les variables prédictives d'un 3 000 m *steeple* pourraient être les mêmes que pour les courses, sans obstacle, de demi-fond et de fond, telles que le $\dot{V}O_{2max}$, l'économie de course et la capacité d'endurance aérobie [14, 318], soit des facteurs physiologiques clés de la course à pied dépendant de caractéristiques individuelles

comme l'âge, le sexe, la taille et l'IMC [17, 319]. Dans la littérature, certaines études se sont intéressées à la relation entre la performance et l'âge chez les coureurs de *steeple* de niveau élite [13, 93]. Les auteurs ont notamment démontré que l'âge des athlètes de niveau élite sur 3 000 m *steeple-chase* se situe à hauteur de 25 ans pour les hommes et entre 25 et 27 ans pour les femmes, lorsque ces derniers réalisent leur meilleure performance de carrière sur la distance. Le sexe semble également être une variable à prendre en considération dans la prédiction de performance dans une course technique telle que le 3 000 m *steeple-chase*. En effet, au-delà du fait que les hommes obtiennent généralement de meilleures performances que les femmes dans les épreuves de demi-fond et de fond [10, 22], des différences dans la technique de course (e.g., longueur de la foulée et la position du corps lors de l'approche des obstacles) sont observables entre les sexes lors du franchissement [82]. Cette différence technique peut notamment être expliquée par la hauteur inférieure des barrières chez les femmes [77]. La vitesse d'approche et la taille des athlètes sont également des éléments pouvant influencer la technique de course [82]. Dans ce sens, Sánchez Muñoz et al. [320] ont récemment montré que les coureurs de *steeple* et de 5 000 m sont plus grands que les coureurs de demi-fond alignés sur de plus courtes distances comme le 800 m et le 1 500 m. En effet, le fait d'être plus grand, le coureur pourrait être avantagé dans le franchissement des obstacles (en lien avec une longueur supérieure des membres inférieurs). De plus, au-delà du possible avantage lié à la taille, il est bien établi qu'un IMC plus faible peut également avoir un impact positif sur les performances des coureurs de demi-fond et de fond [15]. Par conséquent, en raison de l'aspect technique et du coût énergétique lié au franchissement des obstacles [81], la taille et l'IMC pourraient être d'autres variables prédictives de la performance en *steeple-chase*.

L'objectif de la présente étude était de développer et de tester la validité, la justesse et la précision d'une équation relativement simple afin de prédire la performance sur 3 000 m *steeple* à partir de variables individuelles.

2. Méthodes

Tous les bilans officiels français de la FFA de l'année 2019, des épreuves de 3 000 m piste ($n = 5\,385$) et 3 000 m *steeple* ($n = 1\,333$) uniquement, ont été analysés rétrospectivement. À partir de ces bilans, tous les athlètes adultes (≥ 18 ans) ayant une référence sur ces deux distances ont été retenus ($n = 387$). Les coureurs n'ayant pas autodéclaré leur masse corporelle et leur taille ont été retirés de l'analyse ($n = 211$). À la suite de cette étape, 176 athlètes ont été inclus. Enfin, les coureurs ayant maintenu une vitesse supérieure sur l'épreuve du 3 000 m *steeple* en comparaison au 3 000 m plat ont également été éliminés ($n = 5$). Enfin, les 171 athlètes restants ont été séparés aléatoirement en deux échantillons de sujets à l'aide du logiciel SPSS (*Statistical Package for the Social Sciences*). Le premier échantillon était composé de 80% des coureurs ($n = 137$, dont 20 femmes) afin de développer une équation de prédiction simple (*i.e.*, le groupe de « développement »), tandis que le second, comprenant les 20% restant ($n = 34$, dont 5 femmes), était destiné à tester la validité, la justesse et la précision de l'équation de prédiction développée (*i.e.*, groupe de validation croisée).

Pour chaque athlète, le sexe (femme ou homme), la date de naissance (essentielle pour le calcul de l'âge), la taille et la masse corporelle (pour le calcul de l'IMC), ainsi que les temps de course sur 3 000 m *steeple* ($Perf_{Steeple}$) et 3 000 m plat ($Perf_{3000}$) ont été enregistrés.

Un avis favorable du CERSTAPS (n°2019-22-02-31), et une déclaration au règlement général sur la protection des données (RGPD) était préalablement réalisée.

Des méthodes statistiques standards ont été utilisées pour calculer les moyennes et les SD. La distribution normale gaussienne a été vérifiée par le test de Shapiro-Wilk.

Développement de l'équation de prédiction simple

Afin de déterminer si une équation pouvait être développée pour prédire la $Perf_{Steeple}$, une corrélation de Pearson a été utilisée pour évaluer la relation entre la variable dépendante (*i.e.*, $Perf_{Steeple}$) et toutes les variables indépendantes collectées (*i.e.*, $Perf_{3000}$, âge, taille, masse corporelle et IMC).

Une équation de prédiction simple a ensuite été développée à partir des variables significativement corrélées en utilisant une analyse de régression linéaire multiple.

Un indicateur de détection de multicolinéarité, soit le facteur d'inflation de la variance (*i.e.*, abrégé VIF pour « *Variance Inflation Factor* ») a été utilisé pour détecter la multicolinéarité entre les variables indépendantes dans le modèle de régression.

Le test de Fisher a été utilisé pour examiner la contribution de chaque variable au modèle, et la confirmation a été basée sur l'analyse des coefficients β normalisés.

De plus, la relation entre la $Perf_{Steeple}$ estimée par l'équation de prédiction simple et la $Perf_{Steeple}$ réelle a été analysée par la méthode de Bravais-Pearson et quantifiée par le coefficient de corrélation de Pearson.

L'autocorrélation des résidus a été examinée à l'aide du test de Durbin-Watson.

Validation croisée de l'équation de prédiction simple

Le test *t* des échantillons appariés de Student a été utilisé pour comparer les $Perf_{Steeple}$ réelles et prédites. L'amplitude des différences a été évaluée par l'ES.

La relation entre la $Perf_{Steeple}$ réelle et prédite a été testée à l'aide de la méthode de Bravais-Pearson et quantifiée par le coefficient de corrélation de Pearson.

Le biais (*i.e.*, différence entre la $Perf_{Steeple}$ réelle et la $Perf_{Steeple}$ prédite) et les 95% LoA (*i.e.*, $\pm 1,96$ SD) ont été calculés selon la méthode de Bland et Altman [311, 312].

Le seuil de significativité statistique a été fixé à $p < 0,05$ et toutes les analyses ont été effectuées avec le logiciel *Statistical Package for the Social Sciences* (version 20.0, Chicago, IL, USA).

3. Résultats

Les caractéristiques des athlètes des deux échantillons sont présentées dans le Tableau 22.

Tableau 22. Données descriptives et variables utilisées pour développer l'équation simple de prédiction (valeurs moyennes \pm SD)

	Échantillon « Développement » de l'équation de prédiction	Échantillon « Validation croisée » de l'équation de prédiction
Taille de l'échantillon	137	34
Hommes (%)	85,4	85,3
Âge (ans)	27,8 \pm 8,0	27,3 \pm 7,3
Masse corporelle (kg)	61,2 \pm 8,0	61,8 \pm 7,0
Taille (cm)	175 \pm 8	176 \pm 5
IMC (kg.m ⁻²)	19,8 \pm 1,6	19,8 \pm 1,5
<i>Perf_{Steeple}</i> (min and s)	10 min 26 s \pm 1 min 01 s	10 min 23 s \pm 1 min 02 s
<i>Perf₃₀₀₀</i> (min and s)	9 min 34 s \pm 47 s	9 min 29 s \pm 46 s

Légende : SD : écart-type ; IMC : indice de masse corporelle ; *Perf_{Steeple}* : performance sur 3 000 m *steeple* et *Perf₃₀₀₀* : performance sur 3000 m plat.

Développement de l'équation de prédiction simple

Des corrélations bivariées significatives ont été démontrées entre la *Perf₃₀₀₀* ($p < 0,001$, $r = 0,892$), l'âge ($p < 0,001$, $r = 0,428$), le sexe ($p = 0,021$, $r = -0,173$), la taille ($p = 0,050$, $r = -0,141$), l'IMC ($p = 0,003$, $r = 0,230$) et la *Perf_{Steeple}*.

Une analyse de régression linéaire multiple, utilisant la *Perf₃₀₀₀*, le sexe et l'âge comme variables indépendantes potentielles, a abouti à une équation de prédiction simple.

Cette équation était la suivante :

$$Perf_{Steeple} = -89,740 + 1,153 \times Perf_{3000} + 0,952 \times \text{âge} + 14,810 \times \text{sexe} \quad (8)$$

Où la *Perf_{Steeple}* et la *Perf₃₀₀₀* sont exprimées en s, les chiffres 1 et 2 sont attribués respectivement au sexe féminin et masculin, et l'âge est en années.

Les VIF étaient de 1,341, 1,241 et 1,233 pour la *Perf₃₀₀₀*, l'âge et le sexe, respectivement. Par conséquent, la multicolinéarité était très faible (car VIF < 5).

En ajoutant un deuxième (*i.e.*, l'âge) et un troisième prédicteurs (*i.e.*, le sexe) dans l'équation de prédiction, le r^2 a été significativement amélioré, avec $F(1,134) = 14,912$ ($p < 0,001$) et $F(1,133) = 4,465$ ($p = 0,036$), respectivement. De plus, le test de Fisher a révélé un $p < 0,001$. La performance estimée par l'équation de prédiction simple (incluant les trois variables indépendantes : $Perf_{3000}$, âge et sexe) était significativement corrélée avec la $Perf_{Steeple}$ réelle ($r^2 = 0,822$). Les coefficients β normalisés et les valeurs p du test t de Student étaient de 0,878 ($p < 0,001$), 0,125 ($p = 0,003$) et 0,086 ($p < 0,036$) pour la $Perf_{3000}$, l'âge et le sexe, respectivement. Le test de Durbin-Watson n'a indiqué aucune autocorrélation dans les résidus, avec une valeur de 1,737. De plus, selon l'interprétation graphique, les résidus normalisés suivaient une droite (et aucun résidu n'est apparu trop élevé). La prédiction était, par conséquent, valide pour toutes les performances.

Validation croisée de l'équation de prédiction simple

Aucune différence significative n'a été notée entre la $Perf_{Steeple}$ réelle et la $Perf_{Steeple}$ prédite ($p = 0,387$), et l'amplitude de la différence a été considérée comme faible ($ES = 0,36$) selon l'interprétation proposée par Cohen [301].

La $Perf_{Steeple}$ prédite était significativement corrélée à la $Perf_{Steeple}$ réelle ($p < 0,001$), et un coefficient de corrélation, très élevé selon l'échelle de Munro [309] ($r = 0,934$), a été démontré.

Le biais et les LoA à 95% étaient de -3 ± 22 s, soit $-0,5 \pm 7,1\%$ de la $Perf_{Steeple}$ moyenne (Figure 42).

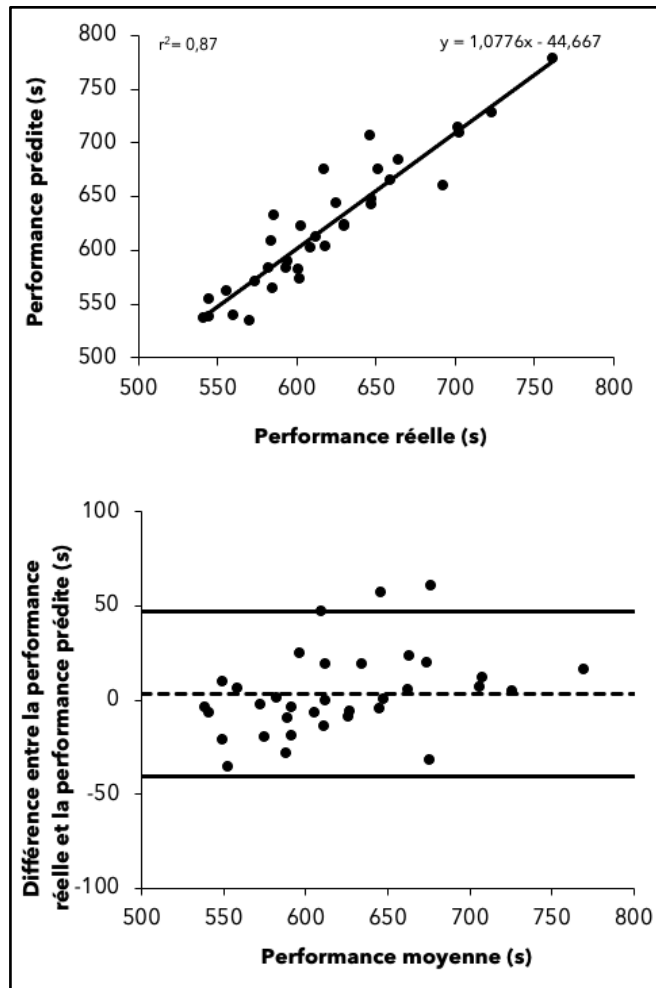


Figure 42. Validité des mesures avec l'équation simple pour prédire la performance

Partie supérieure : Relation entre les performances réelles et prédites à partir d'une équation simple permettant de prédire la performance sur 3 000 m *steeple*. La ligne continue est la régression linéaire. r^2 est le coefficient de détermination.

Panneau inférieur : Graphiques de Bland et Altman comparant les performances réelles et prédites sur 3 000 m *steeple*. La ligne pointillée représente le biais et les lignes en continu sont les limites d'accord à 95%.

4. Discussion

Le but de cette étude était de développer et de tester la validité, la justesse et la précision d'une équation simple pour prédire la $Perf_{Steeple}$ à partir de variables individuelles. Le principal résultat suggère qu'une équation simple incluant la $Perf_{3000}$, l'âge et le sexe (cf. équation 8) est une méthode valide, juste et précise pour prédire les performances des athlètes sur 3 000 m *steeple*. Les résultats ont révélé que pour 95% des prédictions, la différence entre les performances réelles et prédites serait $-0,5 \pm 7,1\%$ sur 3 000 m *steeple*. Bien qu'aucune étude dans la littérature n'ait porté sur la prédiction des performances en *steeple* à partir d'une équation simple, les résultats obtenus ici (*i.e.*, biais et LoA 95%) sont similaires à des études présentées dans les chapitres II et III (cf. deuxième partie) sur 3 000 m plat [24, 25, 270].

Les résultats suggèrent également que le principal prédicteur de la $Perf_{Steeple}$ est la $Perf_{3000}$. Ce résultat pourrait s'expliquer par le fait que l'intensité et la durée de l'effort de ces deux épreuves sont relativement proches. En effet, malgré la présence d'obstacles sur *steeple-chase*, allongeant la durée de l'effort à hauteur de 3-5% comparativement au 3 000 m plat, la distance de course de ces deux épreuves est identique, ce qui par conséquent a relativement peu d'impact sur l'importance des prédicteurs physiologiques clés de la performance, tels que le $\dot{V}O_{2max}$ et la capacité d'endurance aérobie [80, 81, 321]. Cependant, contrairement à l'épreuve du 3 000 m plat, le coureur de 3 000 m *steeple* ralentit généralement lors la phase d'impulsion, tant pour le franchissement des barrières que pour la rivière, même chez les steepleurs de niveau élite [322]. Par conséquent, en plus d'exiger un niveau élevé de $\dot{V}O_{2max}$ (environ 95%) lors de l'effort [80], la performance sur 3 000 m *steeple* exige également que les athlètes aient des capacités techniques afin de franchir les barrières et rivières sans occasionner de perte de vitesse.

Les caractéristiques techniques du 3 000 m *steeple*, telles que la hauteur des barrières, soit une différence de 0,152 m entre les hommes et les femmes [77], peuvent entraîner des différences de capacité de performance entre les sexes. En effet, les résultats ont montré que la variable « sexe » améliorait légèrement la précision de $Perf_{Steeple}$. Ce résultat pourrait s'expliquer par une technique de course variant selon le sexe (e.g., vitesse d'approche vers l'obstacle, zone de réception après la barrière, angle des segments du membre inférieur lors de l'appui sur la rivière) pour franchir les barrières et les rivières [82]. Le ratio, hauteur des barrières et taille des athlètes, inférieure chez les femmes pourrait notamment leur permettre de limiter les changements de rythme lors du franchissement (e.g., ratio de $45,9 \pm 1,4$ vs $51,6 \pm 1,9\%$ chez les femmes et les hommes, respectivement, obtenu dans cette présente étude) [323]. Cependant, bien que le rythme de course des femmes puisse être moins affecté par le franchissement des barrières [323], ce dernier peut être davantage affecté par le franchissement de la rivière (e.g., en lien avec une vitesse d'approche plus faible, distance d'envol moins importante, soit une sortie de la fosse d'eau dans un plan incliné lors de la réception) en comparaison à celui des hommes [82]. Ceci peut être expliqué par le fait que les femmes aient besoin d'une foulée relativement plus ample pour sortir de la fosse d'eau en raison d'une longueur de jambes généralement plus courte que celle des hommes. De plus en prenant appui sur une hauteur de barrière plus petite, les femmes ont une hauteur d'envol inévitablement plus basse pour traverser la rivière en comparaison aux hommes.

Bien que la présente étude montre que l'âge contribue de manière significative à la prédiction de la performance sur le 3 000 m *steeple*, l'inclusion de cette variable dans notre équation simple n'apporte qu'un faible gain de précision. Ce résultat peut être expliqué par le fait que la plupart des steepleurs (95% de la population incluse dans l'étude actuelle) avaient entre 20 et 35 ans (Tableau 22) soit une tranche

d'âge assez hétérogène (*i.e.*, large) au regard des travaux réalisés par Haugen et al. [13] ayant démontré un âge de performance à hauteur de 25 et 27 ans chez les hommes et chez les femmes, respectivement dans cette discipline. De plus, il est bien établi que l'âge peut être un facteur limitant de performance en raison des modifications physiologiques en lien avec le vieillissement caractérisant les capacités humaines [95, 127, 157]. Une réduction progressive du $\dot{V}O_{2\max}$ semble être le principal mécanisme associé au déclin de la performance avec le vieillissement chez l'adulte [95].

Enfin, bien que la taille et l'IMC auraient pu être des variables prédictives de la $Perf_{Steeple}$ au regard de l'aspect technique et du coût énergétique lors du franchissement des barrières ou des rivières [81], les résultats de cette présente étude suggèrent le contraire. Ces derniers pourraient possiblement être expliqués par l'intégration de la $Perf_{3000}$ dans l'équation de prédiction de la $Perf_{Steeple}$. En effet, bien que la taille ne soit que légèrement associée à la vitesse dans les épreuves de course à pied, l'IMC semble être un meilleur indicateur de la vitesse sur le 3 000 m piste [15]. Par conséquent, la taille et l'IMC ayant déjà un impact sur la $Perf_{3000}$, ceci pourrait expliquer le fait que ces variables n'aient pas été incluses dans l'équation de prédiction de la $Perf_{Steeple}$ afin d'éviter la redondance dans les données.

Enfin, l'autodéclaration de la taille et de la masse corporelle permettant l'obtention de l'IMC des athlètes peut être une limite de l'étude. En effet, cette étude s'est notamment contentée des déclarations réalisées par les athlètes sur la base de données de la FFA. Toutefois, de manière générale, lorsque les athlètes déclarent leurs données anthropométriques, ces dernières sont reconnues pour être précises [324].

5. Conclusion

Les résultats ont montré que cette équation simple est un outil valable, juste et précis pour prédire la $Perf_{Steeple}$ à partir de variables individuelles (*i.e.*, la $Perf_{3000}$, l'âge et le sexe). Les résultats ne montrent aucune différence significative entre les performances réelles et prédites chez les coureurs de *steeple-chase*. Les prédictions de cette équation simple peuvent par conséquent être utilisées par les athlètes et les entraîneurs. Avoir une idée juste et précise de la capacité de performance d'un athlète lors d'une compétition spécifique comme l'épreuve de *steeple-chase* peut permettre l'optimisation de la performance et faciliter la détection des athlètes au fort potentiel.

CHAPITRE V. PRÉDICTION DES PERFORMANCES SUR MARATHON À L'AIDE DE L'INTELLIGENCE ARTIFICIELLE (IA)

❖ Publication scientifique (cf. Annexe 6)

- **Lerebourg L**, Saboul D, Cléménçon M, Coquart JB. *Prediction of marathon performance using artificial intelligence. Artificial Intelligence in Medicine*. Soumis.

1. Introduction

Le marathon, épreuve de course à pied d'endurance de 42,195 km, a été créé pour les premiers JO modernes à Athènes en 1896. Depuis le premier marathon « urbain » à New York en 1976 [87], le marathon a gagné en popularité en passant du statut d'épreuve olympique à celui d'évènement social mondial [325]. L'enthousiasme pour cet événement n'a cessé de croître [87, 325, 326] et les temps de course se sont régulièrement améliorés [7, 325, 326]. Indépendamment de l'âge, que ce soient des coureurs amateurs ou des athlètes de niveau élite, les objectifs sur cette épreuve diffèrent comme être « *finisher* », parcourir la distance aussi vite que possible, gagner la course et/ou battre des records (e.g., des records personnels, nationaux ou mondiaux). Bien que les performances de course à pied de longue distance, comme le marathon, puissent être influencées par les conditions environnementales, ces dernières sont principalement dépendantes de caractéristiques personnelles (qualités physiques, état psychologique) et de variables d'entraînement (programmation, gestion d'allure, tactique de course) [5, 7, 18, 201]. Au regard de ces paramètres, les athlètes et les entraîneurs doivent avoir conscience de l'importance du développement de programmes d'entraînement appropriés, mettant notamment l'accent sur la gestion des allures de course et sur la mise en place de stratégies adéquates afin d'optimiser les performances [23, 256,

265, 327]. Pour ces raisons, la capacité de prédiction des performances sur marathon peut être intéressante pour la calibration des séances d'entraînement et la définition des limites potentielles de vitesse de l'athlète afin d'obtenir les meilleures performances.

Récemment, les algorithmes d'apprentissage supervisé, et plus particulièrement les algorithmes de *machine learning*, ont suscité un intérêt croissant. Ces méthodologies intelligentes ont donné des résultats prometteurs dans la prédiction de variables continues dans de nombreux domaines tels que la météorologie [328], la santé [286] et le sport [317]. La littérature indique notamment que le sport est l'un des domaines en pleine expansion dans lequel il peut être intéressant d'obtenir une « bonne » précision prédictive [256, 284, 317]. Cependant, bien que des modèles de régression par apprentissage supervisé tels que les ANN [284-286, 317] ou KNN [23] aient été utilisés pour prédire les performances avec une précision remarquable dans certaines activités sportives, la validité et la précision de la prédiction des performances individuelles ou collectives à l'aide de l'IA méritent d'être explorées davantage [288-291].

Bien que des études se soient concentrées sur l'utilisation d'algorithmes d'apprentissage supervisé (e.g., « *bagging* », « *local matrix completion* ») pour prédire les performances sur marathon [23, 131, 327] et les potentiels risques de ralentissement lors de la course [256], aucune étude à notre connaissance n'a utilisé et validé les algorithmes d'apprentissage automatique supervisé ANN et KNN dans cette discipline de course à pied ni même comparé la justesse ou la précision.

Les objectifs de cette étude étaient donc de tester la validité de deux méthodes d'apprentissage automatique supervisé (ANN et KNN) et de comparer la justesse et

la précision des prédictions de performance sur marathon afin de déterminer laquelle était la plus performante.

2. Méthodes

Tous les bilans FFA (hommes et femmes) de 2019 sur 10 km ($n = 217\ 669$) et marathon ($n = 92\ 813$) ont été analysés. En France, le marathon n'étant pas ouvert aux catégories d'athlètes plus jeunes, seuls les athlètes âgés de plus de 21 ans ont été retenus pour les deux courses ($n = 201\ 990$ sur le 10 km et $n = 92\ 813$ sur le marathon). Les athlètes n'ayant pas autodéclaré leur masse corporelle et/ou leur taille, ont été exclus de l'analyse. Ainsi, 7 716 athlètes ayant réalisé une performance sur 10 km et 4 130 sur marathon ont été inclus. Ainsi, 1 728 performances ont été collectées. Cependant, comme l'objectif était de prédire les performances sur marathon à partir des performances sur 10 km, les athlètes ayant couru un marathon avant leur 10 km ont été retirés ($n = 833$). De plus, les athlètes ayant maintenu une vitesse plus élevée lors du marathon que lors de la course de 10 km ont également été éliminés ($n = 11$). Enfin, parmi les 884 athlètes restants, ceux dont la performance sur 10 km était hors barème FFA ont été éliminés (c'est-à-dire une performance > 50 et 60 min, respectivement, pour les hommes et les femmes). L'analyse a donc été réalisée avec un ensemble de données de 820 athlètes. Pour chaque athlète, le sexe (femme ou homme), la date de naissance (pour calculer l'âge), la masse corporelle et la taille (pour calculer l'IMC), et les temps de course (c'est-à-dire les performances sur le 10 km et sur le marathon) ont été enregistrés.

Deux algorithmes de régression par apprentissage automatique supervisé ont été utilisés : ANN et KNN. Les deux algorithmes ont été « implémentés » en langage R. Le logiciel R (version R X64 3.6.1 - R *Development Core Team*, Vienne, Autriche) a été utilisé pour notre analyse, et les bibliothèques R (*library*) suivantes pour les

approches d'apprentissage automatique ont été utilisées : *dplyr* version 0.8.3, *neural net* version 1.44.2 [329, 330].

Toutes les données ont été normalisées pour répondre aux exigences de la fonction de transfert sigmoïde (pour les ANN) et pour supprimer les différences d'échelle entre les variables d'entrée (pour les KNN).

Les données des 820 athlètes ont été séparées de façon aléatoire en une répartition « *train/test* » pour les processus de formation (entraînement) et de test, respectivement. Le rapport 90:10 a été utilisé, signifiant que 90% des données ont été sélectionnées au hasard pour le processus d'entraînement (soit 738 des 820 performances), tandis que 10% des données ont été sélectionnées aléatoirement pour le processus de test (soit 82 performances) [331, 332].

Pour les deux algorithmes d'apprentissage automatique supervisé, les mêmes entrées (*i.e.*, temps de course sur 10 km, IMC, âge et sexe) ont été utilisées pour résoudre le problème de régression linéaire, consistant à estimer la valeur de la même sortie continue (temps de course sur marathon). Il a également été précisé que les mêmes données d'entraînement ($n = 738$) et de test ($n = 82$) ont été utilisées pour les deux algorithmes.

Dans le réseau ANN, un perceptron multicouche a été utilisé avec quatre entrées et une sortie (Figure 43) [290]. Dans ce réseau, les unités de calcul sont disposées en trois couches bien ordonnées. L'information circule en amont des quatre neurones de la couche d'entrée vers les deux neurones de connexion de la couche cachée et, enfin, vers l'unique neurone de la couche de sortie, sans connexion en retrait. La première couche (la couche d'entrée) correspond aux variables indépendantes (*i.e.*, performance sur 10 km, IMC, âge et sexe), tandis que la troisième couche (la couche de sortie) correspond au score de la variable dépendante (*i.e.*, performance sur marathon). La couche intermédiaire, soit la couche cachée, est constituée de toutes les connexions possibles entre les couches d'entrée et de sortie et permet l'impact

combiné d'un ensemble multiple de variables indépendantes sur la couche de sortie. Ce réseau ANN utilise la méthode « Rprop », abréviation de « *Resilient Back Propagation* », une technique de formation sans retour en arrière de l'apprentissage supervisé [329]. Le point d'arrêt de la formation des données (c'est-à-dire le seuil) a été fixé à 0,01.

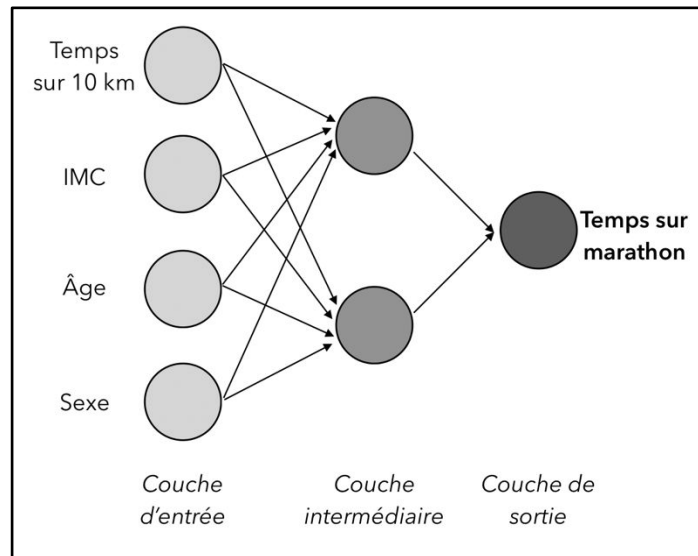


Figure 43. Architecture d'un réseau de neurones

Légende : Indice de masse corporelle (IMC)

Comme pour tester l'avantage de l'ANN, KNN a été appliqué en utilisant les quatre mêmes variables d'entrée ainsi que la même variable de sortie afin de pouvoir les comparer ultérieurement. Dans cette étude, l'algorithme KNN a été testé en sélectionnant les trois plus proches voisins ($k = 3$). En d'autres termes, pour chaque athlète de l'échantillon test, les trois athlètes de l'ensemble des données d'entraînement ayant la plus petite distance euclidienne (somme des différences entre leurs quatre entrées respectives) ont été retenus. Le résultat estimé (*i.e.*, temps sur marathon) a été calculé en utilisant la moyenne des temps de marathon des trois plus proches voisins (*i.e.*, athlètes) pondérés par l'inverse de leurs distances euclidiennes respectives par rapport à l'athlète testé.

Des méthodes statistiques standards ont été utilisées pour calculer les moyennes et les SD.

La distribution normale gaussienne a été vérifiée par le test de Shapiro-Wilk. Un test t de Student sur des échantillons appariés a été utilisé pour les données normalement distribuées afin de comparer les performances réelles et prédites sur marathon pour chaque algorithme d'apprentissage automatique. Lorsque ces données ne passaient pas le test de normalité, un test de Wilcoxon a été utilisé. L'amplitude des différences a été évaluée par l'ES, qui a été interprétée selon l'échelle de Cohen [301].

La relation entre les performances réelles et prédites a été testée à l'aide des corrélations de Pearson ou de Spearman, selon la normalité des données.

Le coefficient de corrélation a été interprété par l'échelle de Munro [309]. Les critères, tels que le coefficient de détermination (r^2) et l'erreur moyenne absolue (*i.e.*, abrégé MAE pour « *Mean Absolute Error* »), basés sur une répartition commune 90:10 des données d'entraînement et de test, respectivement, ont été choisis pour évaluer le succès de l'analyse de régression *via* les deux processus de régression par modélisation. La MAE est la moyenne des erreurs absolues des écarts où des valeurs d'erreur plus faibles signifient généralement que le modèle est plus précis et que les prédictions correspondent étroitement aux valeurs réelles [23, 286] ; r^2 détermine la précision des prédictions et la façon dont le modèle s'adapte aux données [333]. Cela facilite également la comparaison et l'évaluation des résultats [284, 332].

En outre, le biais (c'est-à-dire la différence entre les performances réelles et prédites, permettant d'évaluer la justesse) et les 95% LoA (c'est-à-dire $\pm 1,96$ SD, permettant d'évaluer la précision) ont été calculés selon la méthode de Bland et Altman [311, 312].

Les sorties des performances prédites sur marathon ont été comparées avec les performances réelles sur marathon. Le modèle a été considéré comme valide si la MAE était inférieure à 5% et si les biais et les LoA étaient acceptables. Enfin, les modèles KNN et ANN ont été comparés pour déterminer lequel était le plus juste et précis : le modèle le plus performant a été identifié sur la base de la MAE la plus petite, ainsi que sur la justesse (*i.e.*, biais) et la précision (*i.e.*, LoA) les plus élevées.

Le seuil de significativité statistique a été fixé à $p < 0,05$ et toutes les analyses ont été effectuées avec le logiciel *Statistical Package for the Social Sciences* (version 20.0, Chicago, IL, USA).

3. Résultats

Les moyennes et les SD des performances réelles et prédites sur le marathon sont présentés dans le Tableau 23.

Aucune différence significative n'a été constatée entre les performances réelles et prédites, quel que soit l'algorithme ($p > 0,05$). De plus, l'amplitude du biais des performances prédites était systématiquement insignifiante ($ES \leq -0,091$).

Toutes les performances de course prédites étaient corrélées avec les performances réelles, avec un coefficient de corrélation très élevé ($p < 0,001$, $r \geq 0,918$).

Les MAE étaient, respectivement, de 11 min 16 s (soit 5,6%) et 4 min 48 s (soit 2,4%) pour ANN et KNN.

Les biais \pm 95% LoA sont présentés dans les Figures 44 et 45. Les biais \pm 95% LoA de ANN et KNN étaient, respectivement : 3 min 14 s \pm 28 min 30 s (soit $1,5 \pm 14,1\%$) et -47 s \pm 14 min 21 s (soit $-0,4 \pm 7,2\%$) (Tableau 23 et Figures 44 et 45).

Tableau 23. Résultats des prédictions à partir des algorithmes ANN et KNN

	Performance moyenne (min) ± SD	p	Corrélation		Biais (min) ± 95% LoA	Biais (%) ± 95% LoA	Amplitude du biais	
			r	Interprétation			ES	Interprétation
Performances réelles	199,75 ± 36,01							
Prédictions avec ANN	202,77 ± 30,24	0,063	0,918*	Très élevé	3,023 ± 28,492	1,5 ± 14,1	-0,091	Insignifiante
Prédictions avec KNN	198,96 ± 32,74	0,333	0,982*	Très élevée	-0,79 ± 14,35	-0,4 ± 7,2	-0,024	Insignifiante

Légende : Valeurs moyennes et écart-type (SD) des performances réelles et prédites sur marathon pour chaque algorithme (*i.e.*, réseau de neurones artificiel : ANN, et k-plus proches voisins : KNN) chez 82 athlètes (*i.e.*, 10% des données sélectionnées pour le processus de test), différence entre les performances réelles et prédites (p), amplitude de la différence), coefficient de corrélation de Pearson avec les performances réelles (r), biais et LoA à 95% (min et %) et erreur moyenne absolue (EMA) (%)

Note : Biais = différence entre les performances réelles et prédites ; r = coefficient de corrélation ; 95% LoA = limites d'agrément à 95% ; ES = taille de l'effet. *Corrélation significative à $p < 0,001$

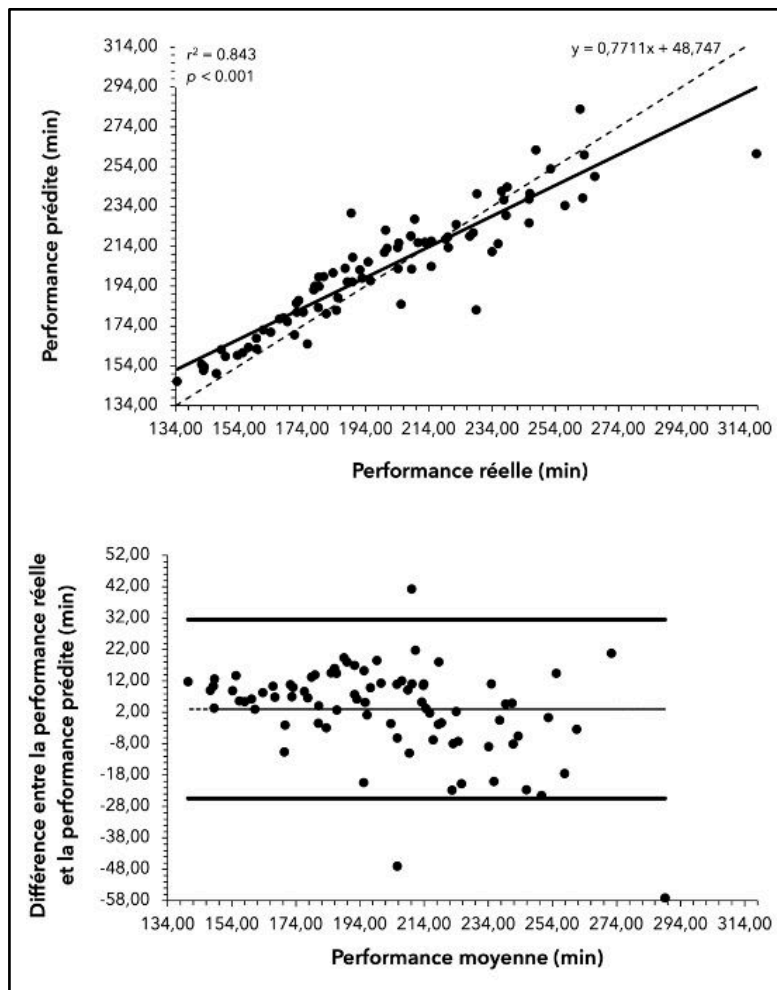


Figure 44. Validité des mesures avec l'algorithme ANN pour prédire la performance

Partie supérieure : relation entre les performances réelles et prédites par l'algorithme ANN sur marathon chez 82 athlètes. La ligne continue représente la régression linéaire. r^2 est le coefficient de détermination.

Partie inférieure : Graphiques de Bland et Altman comparant la performance réelle et la performance prédite sur marathon chez 82 athlètes. La ligne en pointillés représente le biais, les lignes en continu sont les limites d'accord à 95%.

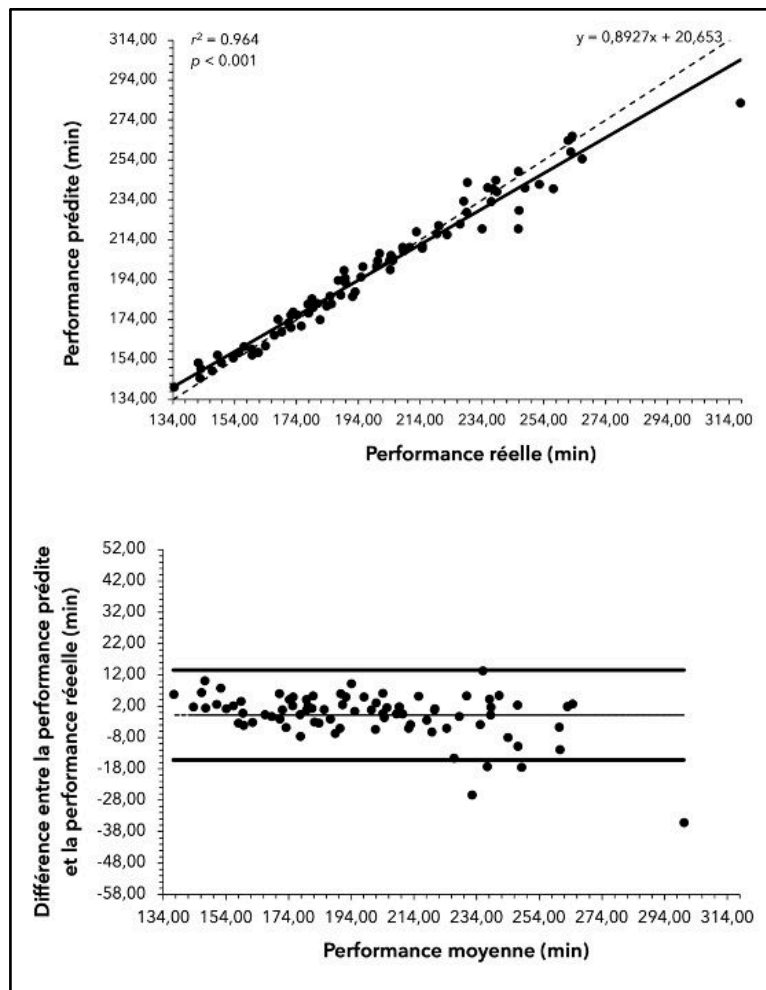


Figure 45. Validité des mesures avec l'algorithme KNN pour prédire la performance

Partie supérieure : relation entre les performances réelles et prédites par l'algorithme KNN sur marathon chez 82 athlètes. La ligne continue représente la régression linéaire. r^2 est le coefficient de détermination.

Partie inférieure : Graphiques de Bland et Altman comparant la performance réelle et la performance prédite sur marathon chez 82 athlètes. La ligne en pointillés représente le biais, les lignes en continu sont les limites d'accord à 95%.

4. Discussion

Les objectifs de cette étude étaient de tester la validité de deux algorithmes d'apprentissage automatique supervisé (ANN et KNN) et de les comparer afin de déterminer lequel était le meilleur pour prédire les performances sur marathon en termes de justesse (à partir de la MAE et du biais) et de précision (à partir de 95% LoA).

L'une des principales conclusions est que les deux algorithmes peuvent effectivement être considérés comme valides, justes et précis pour prédire les performances sur marathon au regard des résultats obtenus. Ces derniers ont confirmé leur validité de prédiction de performances à partir de variables indépendantes (*i.e.*, performances sur 10 km, IMC, âge et sexe) et les deux algorithmes ont affiché une précision de prédiction supérieure à 94%. Ces résultats étaient comparables à ceux d'une autre étude [23] ayant utilisé d'autres méthodes d'IA et ayant montré une précision de 97% chez des coureurs élités.

La MAE était plus faible avec KNN qu'avec ANN, signifiant ainsi que l'algorithme KNN était plus précis (jusqu'à 98% au lieu de 94%) et que les prédictions correspondaient plus étroitement aux performances réelles. Malgré des domaines d'application différents (*i.e.*, algorithme de classification et non de régression), les résultats obtenus étaient en accord avec ceux des travaux de Mustafa et al. [334] et Tamilarasi et Porkodi [335]. Toutefois, nos résultats obtenus étaient en désaccord avec ceux réalisés par Peace et al. [292], Musa et al. [336] et Anyama et al. [337, 338], ayant montré de meilleures prédictions avec ANN que KNN. Quel que soit le domaine, il n'y a pas de réel consensus sur le meilleur modèle d'algorithme en termes de régression ou de classification, et chaque modèle présente des avantages ainsi que des limites [286, 317, 339]. En effet, contrairement aux ANN, les KNN sont plus performants sur de petits échantillons de données et n'ont pas de

risque de tomber dans le surapprentissage. Par conséquent, même si la taille de l'échantillon de données (*i.e.*, 820 athlètes) était relativement importante dans cette étude (dans le champ des sciences du sport, mais faible dans celui de l'IA), le modèle ANN a pu présenter certaines limites (*i.e.*, généralisation ou le risque de surapprentissage des données), ce qui expliquerait la plus grande précision du modèle KNN dans la prédiction des performances sur marathon. Les résultats de la prédiction peuvent être partiellement expliqués non seulement par l'ampleur des données (*e.g.*, la taille de l'ensemble de données) [340], mais aussi par tous les paramètres (*e.g.*, le rapport utilisé pour les ensembles de données d'entraînement et de test, le nombre de couches cachées ou le taux d'apprentissage pour l'entraînement d'un réseau de neurones, le nombre de k dans KNN, le type de distance dans KNN) [292, 293, 317] et la méthode du « *data scientist* » (*e.g.*, la procédure de modélisation définissant les algorithmes) [269].

De plus, les résultats ont révélé que les coefficients de corrélation étaient, respectivement, de 0,918 et 0,982 pour ANN et KNN, et également que dans 95% des résultats, les biais de ANN et KNN étaient de 3 min 14 s \pm 28 min 30 s (*i.e.*, 1,5 \pm 14,1%) et -47 s \pm 14 min 21 s (*i.e.*, -0,4 \pm 7,2%), respectivement. Par conséquent, l'algorithme KNN s'est avéré être plus performant pour prédire la performance sur marathon.

Coquart et al. [265] ont montré que la performance sur marathon pouvait être prédite à partir du nomogramme de Mercier et al. [26] avec des niveaux de justesse et de précision acceptables. En effet, les auteurs ont démontré un biais et une LoA de 95% de -1 min 25 s \pm 27 min 3 s (*i.e.*, -0.7 \pm 13.2%) [265], ce qui est comparable à l'ANN de la présente étude. Cependant, pour un même niveau de justesse et de précision, Coquart et al. [265] ont eu recours à deux performances maximales de longues distances (10 km et 20 km), alors que les prédictions de cette présente étude ont été obtenues à partir d'une seule performance de course à pied (*i.e.*, 10

km). Plus récemment, Vickers et Vertosick [17] ont exploré plusieurs méthodes (*i.e.*, formule de Riegel [113], et deux modèles basés sur une ou deux courses antérieures, respectivement) pour prédire les temps de course sur marathon chez des coureurs amateurs. Entre les temps de marathon prédits et réels, ils ont montré une erreur quadratique moyenne (EQM, mesure caractérisant la précision d'un estimateur) de 6 min 21 s avec la formule de Riegel [113], de 3 min 48 s avec le modèle basé sur une course antérieure, et une EQM de 3 min 28 s avec le modèle basé sur deux courses antérieures. Hormis, l'association de certaines variables (*e.g.*, sexe, âge, IMC) avec le temps de course, le nombre de performances incluses dans un modèle de prédiction de la performance peut améliorer la précision des prédictions. En outre, il est particulièrement intéressant de noter que dans l'étude actuelle, intégrant des facteurs similaires dans les algorithmes (*i.e.*, sexe, âge, IMC), les EQM étaient plus faibles que ceux de l'étude antérieure de Vickers et Vertosick [17] avec une EQM de 2 min 39 s pour ANN et une EQM de 39 s pour KNN, et en ne prenant qu'une seule performance (*i.e.*, temps sur 10 km). Par conséquent, il pourrait être intéressant d'ajouter une autre performance (*e.g.*, semi-marathon) à ANN et KNN pour voir si la justesse et la précision de la prédiction sont significativement meilleures que les méthodes de prédiction ne reposant pas sur l'IA. Pour limiter la fatigue des athlètes, les ANN et surtout les KNN devraient néanmoins être préférés aux autres méthodes de prédiction qui nécessitent deux performances.

Les principales limites potentielles de cette étude étaient la taille de l'échantillon et le nombre de variables incluses. Comme mentionné dans la discussion, KNN a l'avantage de mieux fonctionner sur de petits échantillons de données (pas de risque de tomber dans le surapprentissage) que l'algorithme ANN. En outre, le choix et l'inclusion des variables d'entrée (limitées, avec seulement quatre variables, à savoir la performance sur 10 km, l'IMC, l'âge et le sexe) ont structuré les

algorithmes et influencé les résultats de la prédiction. En d'autres termes, la possibilité d'ajuster l'algorithme pour mieux modéliser le domaine problématique sera toujours envisageable. En effet, bien que la performance en marathon puisse être affectée par une multitude de facteurs puisque la course implique plusieurs éléments de performance, des variables physiologiques (e.g., $\dot{V}O_{2max}$, économie de course et capacité d'endurance aérobie...) [131], psychologiques (e.g., la motivation, l'anxiété et le stress...) et environnementales peuvent également être déterminantes dans la performance en course de fond [5, 7, 201]. L'autre limite principale de cette étude concerne la qualité des données récupérées. En effet, il est probable que les performances (*i.e.*, temps sur 10 km et marathon) aient été influencées par le profil de la course (e.g., profil du parcours : montée, descente), les conditions météorologiques, le terrain d'opposition ou la stratégie de course (e.g., courir pour gagner ou courir pour un temps), mais ces informations n'étaient pas disponibles.

Les applications de l'étude actuelle pourraient être étendues en réalisant des études de validation sur d'autres distances ou profil de course (*i.e.*, course de demi-fond ou de fond sur piste ou sur route). Il serait également intéressant d'utiliser d'autres techniques d'apprentissage automatique pour identifier les facteurs déterminants sur marathon (*i.e.*, utilisation de techniques de classification comme le « *random forest* » ou les « *naive Bayes* ») afin d'aider les athlètes, le staff et les professionnels à concevoir des programmes d'entraînement et détecter de potentiels athlètes aux fortes capacités athlétiques, en plus de prédire les performances.

5. Conclusion

Peu d'études sur la prédiction des performances de course à pied, en particulier sur le marathon, ont utilisé l'IA avec les algorithmes ANN et/ou KNN. Les résultats de la présente étude ont démontré que KNN et ANN étaient capables de prédire les performances des marathoniens avec un niveau très acceptable. Les deux modèles, valides, ont été capables d'atteindre une précision de prédiction supérieure à 94%, bien que KNN semble être supérieur à ANN avec une précision de prédiction des performances sur marathon supérieure à 98%. Ces approches peuvent donc être utilisées pour prédire les performances. Les futures études devraient être orientées vers l'utilisation de techniques d'apprentissage supervisé afin de mieux appréhender la performance. Ces travaux pourraient notamment inclure d'autres paramètres de performance de course à pied ou encore utiliser d'autres techniques d'IA telles que la classification, afin de pouvoir au-delà de la prédiction de performance, détecter les athlètes au fort potentiel.

CHAPITRE VI. PRÉDICTION DU POTENTIEL DES COUREURS DE DEMI-FOND VIA L'ÉVALUATION DU TAUX D'AMÉLIORATION DES PERFORMANCES

- ❖ Publication scientifique (cf. Annexe 7)
 - **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. *Predicting the potential of middle-distance runners by the evaluation of performance improvement rate*. À soumettre.

- ❖ Communication orale sans acte
 - **Lerebourg L**, Guignard B, L'Hermette M, Coquart JB. Prédire le potentiel des athlètes de demi-fond français via l'évaluation du taux d'amélioration des performances. 5^{ème} Doctoriales du CETAPS. 24-25 juin 2021. Mont-Saint-Aignan, France.

1. Introduction

Certains auteurs ont cherché à examiner la progression des performances humaines au fil des décennies, en comparant les sexes [3, 10, 13, 22] et/ou les distances des épreuves de course [3, 13, 22]. Haugen et al. [13] ont par exemple observé une amélioration des performances plus importante chez les femmes que chez les hommes sur 800 m et sur 1 500 m à hauteur de 0,4, \pm 0,2% et à hauteur de 1,2, \pm 0,2% sur des distances plus longues telles que le 5 000 m entre 2002 et 2016. De manière générale, si les résultats montrent une progression plus importante des performances chez les femmes par rapport aux hommes ces dernières années, ils montrent également que les performances en course à pied ont atteint un pic dans de nombreuses disciplines [3, 10]. Cependant, la plupart des études susceptibles de faciliter la compréhension des performances, ainsi que l'évolution de ces

dernières entre les sexes et les épreuves de demi-fond et de fond sur piste, se sont principalement intéressées à l'élite mondiale masculine.

La performance en demi-fond, incluant les distances entre 800 m et 5 000 m, nécessite la contribution des métabolismes aérobie et anaérobie, dont la participation relative évolue en fonction de la distance ou de la durée [14, 63, 66]. En effet, les distances plus longues, comme le 5 000 m, nécessitent une contribution plus importante de la voie de la phosphorylation oxydative alors que les distances plus courtes, comme le 800 m et le 1 500 m, sont caractérisées par une prépondérance plus importante de la voie glycolytique pour maintenir une vitesse accrue [14, 63, 65, 66, 341]. De plus, au-delà des paramètres physiologiques caractérisant les capacités humaines, avec l'entraînement et sur le long terme, il semble possible d'atteindre des pics de performance à des âges plus avancés sur les distances les plus longues (e.g., 5 000 m) que sur les distances les plus courtes (e.g., 800 m et 1 500 m) [13, 93-95, 313]. Par conséquent, être capable d'identifier et de prédire le potentiel futur des jeunes coureurs talentueux dans le but de réaliser des performances en demi-fond et à plus long terme en fond peut être d'un grand intérêt pour les entraîneurs et les athlètes.

Une étude récente de Santos et al. [342] a proposé une nouvelle méthode (basée sur un taux d'amélioration de la performance) pour évaluer si une stratégie d'entraînement induit une progression optimale des performances en course à pied. Par exemple, pour battre le record du monde d'athlétisme du 400 m, des améliorations de 0,97 à 6,83% et de 0,41 à 5,85% par an (calculs basés sur les performances réalisées au moins 5 ans avant le record du monde) ont été indiquées chez les femmes et chez les hommes, respectivement [342]. Cependant, seules les épreuves de *sprint* (i.e., 100 m et 400 m) chez les athlètes élités de niveau mondial ont été étudiées par Santos et al. [342] ; et il ne semble pas y avoir, à ce jour, dans

la littérature de méthode d'évaluation pour les épreuves de demi-fond (e.g., 800 m, 1 500 m, 5 000 m), ce qui fournirait aux entraîneurs un outil alternatif pour le suivi de leurs athlètes.

Dans un premier temps, l'objectif de cette étude était d'analyser le taux d'amélioration de la performance des athlètes composant le Top 10 national français dans les épreuves de demi-fond (e.g., 800 m, 1 500 m et 5 000 m) à travers une analyse de leur meilleure performance personnelle par saison sur plusieurs années. Dans un second temps, cette étude avait pour visée de proposer, en application pratique, un index de performance, afin d'évaluer la progression des athlètes, de prédire le potentiel, et/ou éventuellement de détecter les athlètes au fort potentiel.

2. Méthodes

Les données ont été extraites de la base de données de la FFA. Les caractéristiques et les performances des dix premiers athlètes féminins et masculins figurant dans les dix premières places des bilans officiels « tous temps » (*i.e.*, sur toutes les années depuis l'homologation des records) sur les épreuves de 800 m, 1 500 m et 5 000 m ont été recueillies. Pour chaque athlète, la date de naissance, la nationalité, la meilleure performance de la carrière : MP (*i.e.*, le record personnel, celui qui a permis à l'athlète de figurer dans le Top 10 du classement national de tous les temps), la MP de la saison précédant immédiatement la MP : MP-1, et la MP de chaque saison précédant la MP afin d'obtenir la moyenne des MP des saisons précédentes: MP-MS, ont été enregistrées (Figure 46). Les valeurs aberrantes, identifiées par les boîtes à moustaches et les performances qui n'ont pas été réalisées dans la continuité des saisons précédant la MP ont été exclues. Conformément à l'étude de Santos et al. [342], la différence entre la MP et la MP-1

correspondait au taux d'amélioration de la dernière saison (TA-1S, soit le taux d'amélioration sur une seule saison et celle précédant la MP), tandis que le taux d'amélioration sur plusieurs saisons (TA-MS) représentait la moyenne de l'ensemble des taux d'amélioration précédant la MP, comme le montre la Figure 46.

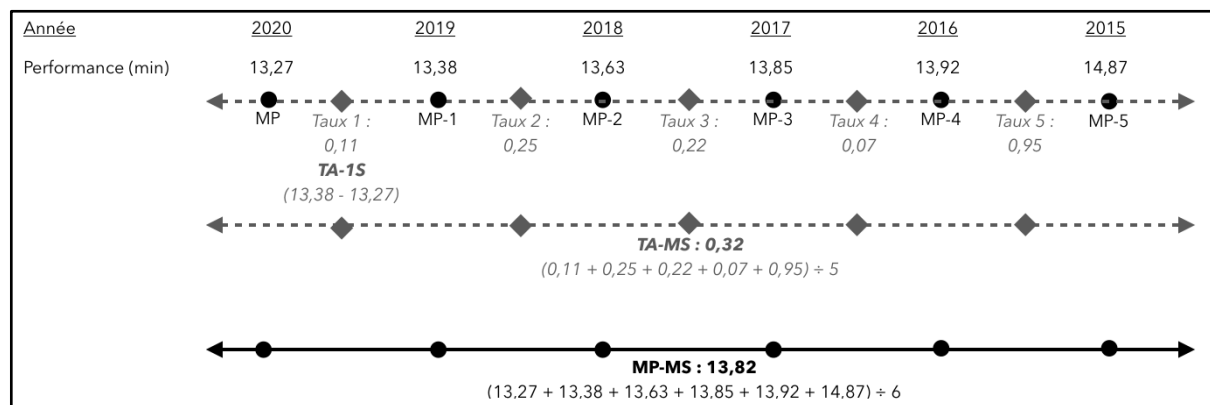


Figure 46. Définitions des termes : MP, MP-1, MP-MS, TA-1S et TA-MS chez un coureur masculin de 5 000 m (sur 6 ans)

Légende : MP : meilleure performance de la carrière (*i.e.*, record personnel) ; MP-1 : MP de la saison précédant immédiatement la MP ; MP-MS : moyenne des MP des saisons précédant la MP ; TA-1S : différence entre la MP et la MP-1 ; TA-MS : moyenne du taux d'amélioration de la saison précédant immédiatement la MP ainsi que des saisons précédentes.

Cette étude a été approuvée par les autorités compétentes (CERSTAPS n°2019-22-02-31).

Les performances de 60 athlètes (*i.e.*, dix meilleurs français, hommes et femmes, sur les trois distances étudiées) ont été prises en considération (Tableau 24). Les plages d'années de MP (en accord avec le Top 10 national de tous les temps) par sexe et par distance s'étendaient de 1986 à 2020, de 1987 à 2010 et de 1994 à 2014 sur 800 m, 1 500 m et 5 000 m respectivement chez les femmes ; et s'étendaient de 1979 à 2020, de 1995 à 2019 et de 1995 à 2019 sur 800 m, 1 500 m et 5 000 m, respectivement, chez les hommes (Tableau 24). Chez les femmes, ces performances sur 800 m, 1 500 m et 5 000 m représentaient respectivement $95 \pm 1\%$, $94 \pm 1\%$ et $93 \pm 1\%$ du record du monde extérieur actuel. Chez les hommes, ces performances

correspondaient, respectivement, $97 \pm 1\%$, $97 \pm 1\%$ et $95 \pm 1\%$ des records du monde en plein air actuels.

Tableau 24. Données extraites des dix meilleurs records nationaux retenus par sexe et distance (Nom, MP, date et âge de la MP)

Distance	Rang	Hommes			Femmes		
		MP (min et s)	Date de la MP	Âge lors de la MP (ans)	MP (min et s)	Date de la MP	Âge lors de la MP (ans)
800 m	1	1 min 42 s	18/07/2014	22	1 min 56 s	09/09/1995	24
	2	1 min 43 s	08/09/2002	24	1 min 58 s	05/06/2016	23
	3	1 min 43 s	12/09/1979	25	1 min 58 s	15/09/2020	24
	4	1 min 43 s	13/07/1987	27	1 min 58 s	09/09/2010	31
	5	1 min 44 s	27/08/2016	25	1 min 58 s	18/07/2008	23
	6	1 min 44 s	02/07/2004	23	1 min 59 s	27/06/1992	25
	7	1 min 44 s	14/08/2020	22	1 min 59 s	15/07/1986	20
	8	1 min 44 s	29/08/2001	21	1 min 59 s	31/07/2004	27
	9	1 min 45 s	16/08/1995	27	1 min 59 s	21/07/2004	31
	10	1 min 45 s	06/08/2004	21	1 min 59 s	19/07/2002	27
1 500 m	1	3 min 28 s	05/09/2003	25	3 min 59 s	16/07/2010	31
	2	3 min 30 s	15/08/2003	25	4 min 00 s	11/07/2006	33
	3	3 min 31 s	19/07/2002	33	4 min 01 s	16/07/2010	30
	4	3 min 32 s	21/07/1999	26	4 min 01 s	09/09/2005	26
	5	3 min 32 s	12/07/1995	29	4 min 02 s	23/08/1996	25
	6	3 min 32 s	19/07/2013	35	4 min 04 s	04/07/2003	29
	7	3 min 32 s	02/07/2004	21	4 min 04 s	18/07/1994	29
	8	3 min 33 s	10/08/1996	25	4 min 05 s	04/08/1999	30
	9	3 min 33 s	24/08/2019	32	4 min 05 s	19/08/1987	21
	10	3 min 33 s	02/07/2013	28	4 min 07 s	06/07/2001	26
5000 m	1	12 min 58 s	28/07/2000	28	14 min 43 s	03/09/2004	30
	2	13 min 02 s	28/08/1998	28	14 min 47 s	01/08/2000	26
	3	13 min 10 s	06/07/2012	32	14 min 58 s	07/09/1999	30
	4	13 min 10 s	25/08/2000	26	15 min 07 s	23/05/2014	24
	5	13 min 11 s	20/06/2006	29	15 min 11 s	14/07/1998	24
	6	13 min 12 s	18/05/2014	36	15 min 12 s	14/07/2006	27
	7	13 min 12 s	07/09/1996	21	15 min 13 s	20/06/1999	34
	8	13 min 13 s	17/06/1995	35	15 min 14 s	06/07/2013	32
	9	13 min 13 s	01/08/1998	35	15 min 16 s	02/08/2014	40
	10	13 min 13 s	05/07/2019	23	15 min 16 s	21/08/1994	32

Légende : MP : meilleure performance

Des méthodes statistiques standards ont été utilisées pour calculer les moyennes et les SD.

La distribution normale gaussienne a été vérifiée par le test de Shapiro-Wilk et l'homogénéité de la variance par le test de Levene. Une ANOVA à deux facteurs pour mesures indépendantes a été utilisée pour les données normalement

distribuées afin de comparer la variable dépendante (*i.e.*, l'âge, le nombre de performances, la MP, la MP-MS, le TA-1S ou le TA-MS) en fonction du sexe (*i.e.*, hommes vs femmes) et de la distance (*i.e.*, 800 m vs 1 500 m vs 5 000 m). Lorsque ces données n'ont pas passé le test de normalité, un test de Kruskal-Wallis a été utilisé. Ensuite, les analyses *post-hoc* ou les tests de Mann-Whitney ont été utilisés pour identifier les différences significatives pour les données suivant une loi normale et les données non-paramétriques, respectivement. Le seuil de significativité a été fixé à $p < 0,05$. Lorsque des analyses *post-hoc* ont été effectuées, une correction de Bonferroni a été appliquée en tenant compte du nombre de comparaisons. L'amplitude des différences a été évaluée par l'ES et interprétée selon l'échelle de Cohen [301] ou selon un coefficient de corrélation de rangs bivariés (r_B) [343] pour les données suivant une loi normale ou non-normale, respectivement. Selon le d de Cohen, l'effet a été considéré comme insignifiant pour $ES < 0,2$; faible pour $0,2 \leq ES < 0,5$; modéré pour $0,5 \leq ES < 0,8$, et important pour $ES \geq 0,8$; et selon le r_B l'effet a été considéré comme insignifiant pour $ES < 0,1$; faible pour $0,1 \leq ES < 0,3$; modéré pour $0,3 \leq ES < 0,5$ ou important $ES \geq 0,5$.

En outre, en accord avec l'étude de Santos et al. [342], à partir des données descriptives du TA-MS (en min), une liste de mille entrées aléatoires a été générée dans chaque groupe (800 m, 1 500 m et 5 000 m ; femmes et hommes) afin de créer une courbe de Gauss et obtenir l'index de performance. Bien que les nombres générés soient aléatoires, la moyenne (*i.e.*, valeur centrale) de chaque courbe et la distance (*i.e.*, valeur du SD) de TA-MS sont réelles, ce qui servira aux fins de ce manuscrit. En particulier, la valeur de SD indiquera si un certain athlète s'améliore plus ou moins que les meilleurs athlètes sur la distance choisie.

Toutes les analyses ont été effectuées avec le logiciel *Jeffrey's Amazing Statistics Program* (JASP Team, version 0.14.1, computer software).

Conception de l'index de performance et procédures de traitement

Conformément à l'étude de Santos et al. [342], les variables TA-1S et TA-MS ont été calculées pour chaque athlète. Ensuite, ces mesures ont été rapportées sous forme de valeur relative (%). Une analyse descriptive (valeurs moyennes, SD et plage de données) de l'ensemble des données a été effectuée (*cf.* résultats, Tableau 25 et Figure 47). L'index permettant l'évaluation du taux d'amélioration de la performance optimale a été généré à partir des données réelles des dix meilleurs athlètes français figurant dans les bilans « tous temps » féminins et masculins sur les épreuves de 800 m, 1 500 m et 5 000 m.

3. Résultats

Les valeurs moyennes et les SD des variables (*i.e.*, l'âge, le nombre de performances, la MP, la MP-MS, le TA-1S et le TA-MS) sont présentés dans le Tableau 25 et la Figure 47.

Tableau 25. Données descriptives des dix meilleures performances françaises de tous les temps en demi-fond (valeur moyenne, SD et étendue des données)

	Hommes		Femmes	
	Moyenne (SD)	Étendue (min-max)	Moyenne (SD)	Étendue (min-max)
Âge (ans)				
800 m	23,7 (2,3)	21 - 27	25,5 (3,5)	20 - 31
1 500 m	27,9 (4,4)	21 - 35	28,0 (3,5)	21 - 33
5 000 m	29,3 (5,2)	21 - 36	29,9 (4,9)	24 - 40
Nombre de Performances enregistrées par athlète (n)				
800 m	5,3 (2,2)	2 - 8	6,2 (3,6) ^c	1 - 12
1 500 m	7,3 (4,1)	1 - 13	4,8 (4,7)	1 - 16
5 000 m	4,1 (3,3)	1 - 11	2,3 (1,6) ^a	1 - 6
MP (min et s)				
800 m	1 min 44 s (1,2 s) ^{*b,c}	1 min 42 s - 1 min 45 s	1 min 58 s (1,2 s) ^{*b,c}	1 min 56 s - 1 min 59 s
1 500 m	3 min 32 s (1,2 s) ^{*a,c}	3 min 28 s - 3 min 33 s	4 min 3 s (2,6 s) ^{*a,c}	3 min 59 s - 4 min 7 s
5 000 m	13 min 9 s (5,4 s) ^{*a,b}	12 min 58 - 13 min 13 s	15 min 6 s (12 s) ^{*a,b}	14 min 43 s - 15 min 16 s
MP-MS (min et s)				
800 m	1 min 47 s (1,8 s) ^{*b,c}	1 min 44 - 1 min 49 s	2 min 3 s (2,4 s) ^{*b,c}	1 min 58 - 2 min 5 s
1 500 m	3 min 39 s (5,4 s) ^{*a,c}	3 min 33 s - 3 min 46 s	4 min 10 s (6,6 s) ^{*a,c}	4 min 1 s - 4 min 21 s
5 000 m	13 min 19 s (12 s) ^{*a,b}	13 min 18 s - 13 min 40 s	15 min 20 s (15 s) ^{*a,b}	14 min 55 s - 15 min 39 s
TA-1S (min et s)				
800 m	1,8 s (1,2 s) ^c	1,2 s - 4,2 s	3 s (2,4 s) ^c	0 s - 7,2 s
1 500 m	2,4 s (1,8 s) ^c	0 s - 4,8 s	5,4 s (3 s) ^c	1,8 s - 12 s
5 000 m	10,8 s (5,4 s) ^{a,b}	4,2 s - 21 s	31,8 s (20,4 s) ^{a,b}	13,8 s - 1 min 6 s
TA-1S (%)				
800 m	2,02 (1,06)	0,95 - 3,88	2,43 (1,80)	0 - 5,88
1 500 m	1,18 (0,71)	0 - 2,35	2,11 (1,37)	0,82 - 4,98
5 000 m	1,36 (0,72)	0,50 - 2,65	3,57 (2,30)	1,58 - 7,47
TA-MS (min et s)				
800 m	1,8 s (0,6 s) ^c	0,6 s - 3 s	1,8 s (1,2 s) ^{b,c}	1,2 s - 4,2 s
1 500 m	3 s (1,2 s) ^c	1,2 s - 6 s	5,4 s (4,2 s) ^{a,c}	1,8 s - 13,2 s
5 000 m	7,8 s (4,2 s) ^{a,b}	2,4 s - 13,8 s	28,2 s (21 s) ^{a,b}	9,6 s - 1 min 6 s
TA-MS (%)				
800 m	1,64 (0,72)	0,57 - 2,72	1,66 (0,78)	0,57 - 3,26
1 500 m	1,41 (0,64)	0,62 - 2,68	2,13 (1,57)	0,81 - 5,27
5 000 m	0,97 (0,49) [*]	0,32 - 1,77	3,10 (2,37) [*]	1,04 - 7,47

Légende : SD : écart-type ; MP : meilleure performance ; MP-MS : moyenne des meilleures performances sur plusieurs saisons ; TA-1S : taux d'amélioration de la dernière saison ; TA-MS : moyenne des taux d'amélioration sur plusieurs saisons ; ^{*}Différences entre les sexes ($p < 0,05$) ; ^aDifférence significative avec le 800 m ; ^bDifférence significative avec le 1 500 m ; ^cDifférence significative avec le 5 000 m ($p < 0,05$).

Prédiction de la performance en course à pied : demi-fond et fond

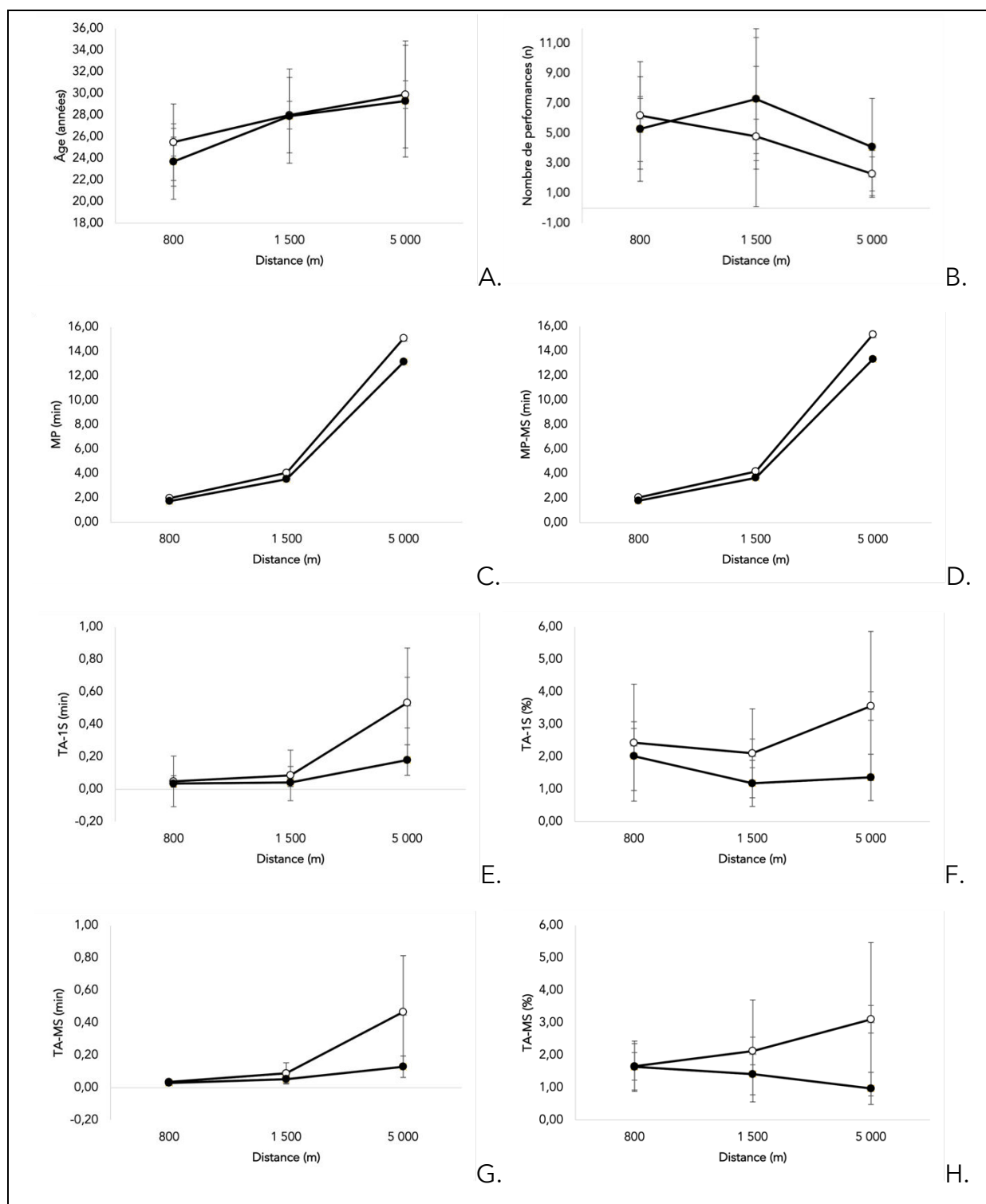


Figure 47. Descriptifs de chaque variable en fonction du sexe et de la distance des dix meilleurs athlètes français de tous les temps en demi-fond (moyenne et SD)

Légende : SD : écart-type. Les points blancs et noirs représentent respectivement les femmes et les hommes de chaque variable : âge (A), nombre de performances (B), MP : meilleure performance (min) (C), MP-MS : moyenne des meilleures performances sur plusieurs saisons (min) (D), TA-1S : taux d'amélioration de la dernière saison (min) (E), TA-1S (%) (F), TA-MS : moyenne des taux d'amélioration sur plusieurs saisons (min) (G) et TA-MS (%) (H).

Âge

Une différence significative a été obtenue entre les distances uniquement (hommes et femmes confondus) ($p = 0,001$). Une analyse *post-hoc* a montré que les coureurs de 800 m étaient plus jeunes que les coureurs de 1500 m et 5000 m respectivement ($p < 0,01$) avec une amplitude de différence importante (ES = 0,967 et ES = 1,221, respectivement). Il n'y avait pas d'effet significatif entre les sexes ($p = 0,433$) et pas d'interaction significative entre le sexe et la distance ($p = 0,796$).

Nombre de performances enregistrées

Il n'y avait pas d'effet significatif pour le sexe ($p = 0,130$) mais il y avait une différence significative avec la distance ($p = 0,017$). L'analyse *post-hoc* a montré que le nombre de performances enregistrées chez les femmes était significativement plus élevé dans l'épreuve du 800 m que dans celle du 5 000 m ($p = 0,012$) avec une amplitude de différence importante ($r_B = 0,670$).

MP

Il y avait des effets significatifs entre les deux sexes ($p = 0,026$) et entre les distances ($p < 0,01$). L'analyse *post-hoc* a montré que la MP était significativement plus faible chez les hommes que chez les femmes, quelle que soit la distance de l'épreuve ($p < 0,001$) avec une amplitude de différence importante ($r_B = 1,000$). Le test *post-hoc* a révélé également que la MP était significativement différente entre chaque distance, pour chaque sexe ($p < 0,01$) avec une amplitude de différence importante ($r_B = 1,000$).

MP-MS

Il y avait des effets significatifs entre les deux sexes ($p = 0,027$) et entre les distances ($p < 0,01$). L'analyse *post-hoc* a montré que la MP-MS était significativement plus faible chez les hommes que chez les femmes, quelle que soit la distance de

l'épreuve ($p < 0,01$), avec une amplitude de différence importante ($r_B = 1,000$) ; et a révélé que la MP-MS était significativement différente entre chaque distance, pour chaque sexe ($p < 0,01$), avec également une importante amplitude de cette différence ($r_B = 1,000$).

TA-1S (min)

Il n'y avait pas d'effet significatif entre les sexes ($p = 0,159$) mais un effet significatif entre les distances ($p < 0,01$). L'analyse *post-hoc* a montré que le TA-1S (min) était significativement plus élevé dans les épreuves de 5 000 m par rapport aux épreuves de 800 m et de 1 500 m, pour chaque sexe ($p < 0,01$) avec une amplitude de différence importante entre les distances ($r_{B800-5000} = 1,000$ et $r_{B1500-5000} = 1,000$ pour les femmes ; $r_{B800-5000} = 0,989$ et $r_{B1500-5000} = 0,944$ pour les hommes).

TA-1S (%)

Il y avait un effet significatif entre les sexes ($p = 0,036$). Une analyse *post-hoc* a révélé que le TA-1S (%) était plus élevé chez les femmes que chez les hommes sur 5 000 m uniquement ($p = 0,021$). Cependant, après une correction de Bonferroni, compte tenu du nombre de comparaisons, ce résultat n'est pas apparu significatif, bien que l'amplitude de la différence sur 5 000 m entre les hommes et les femmes soit importante ($r_B = 0,741$). Il n'y avait pas d'effet significatif entre les distances ($p = 0,361$).

TA-MS (min)

Il n'y avait pas d'effet significatif entre les sexes ($p = 0,297$) mais il y avait un effet significatif entre les distances ($p < 0,01$). L'analyse *post-hoc* a montré que les TA-MS (min) étaient significativement plus élevés sur la course de 5 000 m par rapport aux courses de 800 m et 1 500 m pour chaque sexe ($p < 0,01$) avec une amplitude de différence entre les distances importante ($r_{B800-5000} = 1,000$ et $r_{B1500-5000} = 0,926$ pour

les femmes ; $r_{B800-5000} = 0,944$ et $r_{B1500-5000} = 0,764$ pour les hommes), et le test *post-hoc* a également montré que les TA-MS (min) étaient significativement plus élevés sur les épreuves de 1 500 m par rapport à celles de 800 m chez les femmes ($p = 0,012$) avec une amplitude de différence importante ($r_B = 0,704$).

TA-MS (%)

Il y avait un effet significatif entre les sexes ($p = 0,045$). L'analyse *post-hoc* a montré que les TA-MS (%) étaient significativement plus élevés chez les femmes que chez les hommes sur le 5 000 m ($p = 0,008$) et que l'amplitude de la différence était importante ($r_B = 0,815$). Il n'y avait pas d'effet significatif entre les distances ($p = 0,712$).

Index de performance

L'index de performance permettant l'évaluation du taux d'amélioration de la performance optimale est présenté dans la Figure 48. Les résultats de la présente étude suggèrent que pour qu'une femme atteigne le record national du 800 m, elle doit améliorer sa performance de 1,02 à 3,06 s à l'année. En revanche, un homme doit améliorer sa performance entre 1,02 et 2,58 s. Au 1 500 m, les femmes doivent améliorer leur performance entre 1,32 à 9,36 s et les hommes entre 1,62 et 4,5 s. Enfin sur 5 000 m, les femmes doivent s'améliorer annuellement entre 7,14 et 48,9 s et les hommes entre 3,72 et 11,52 s.

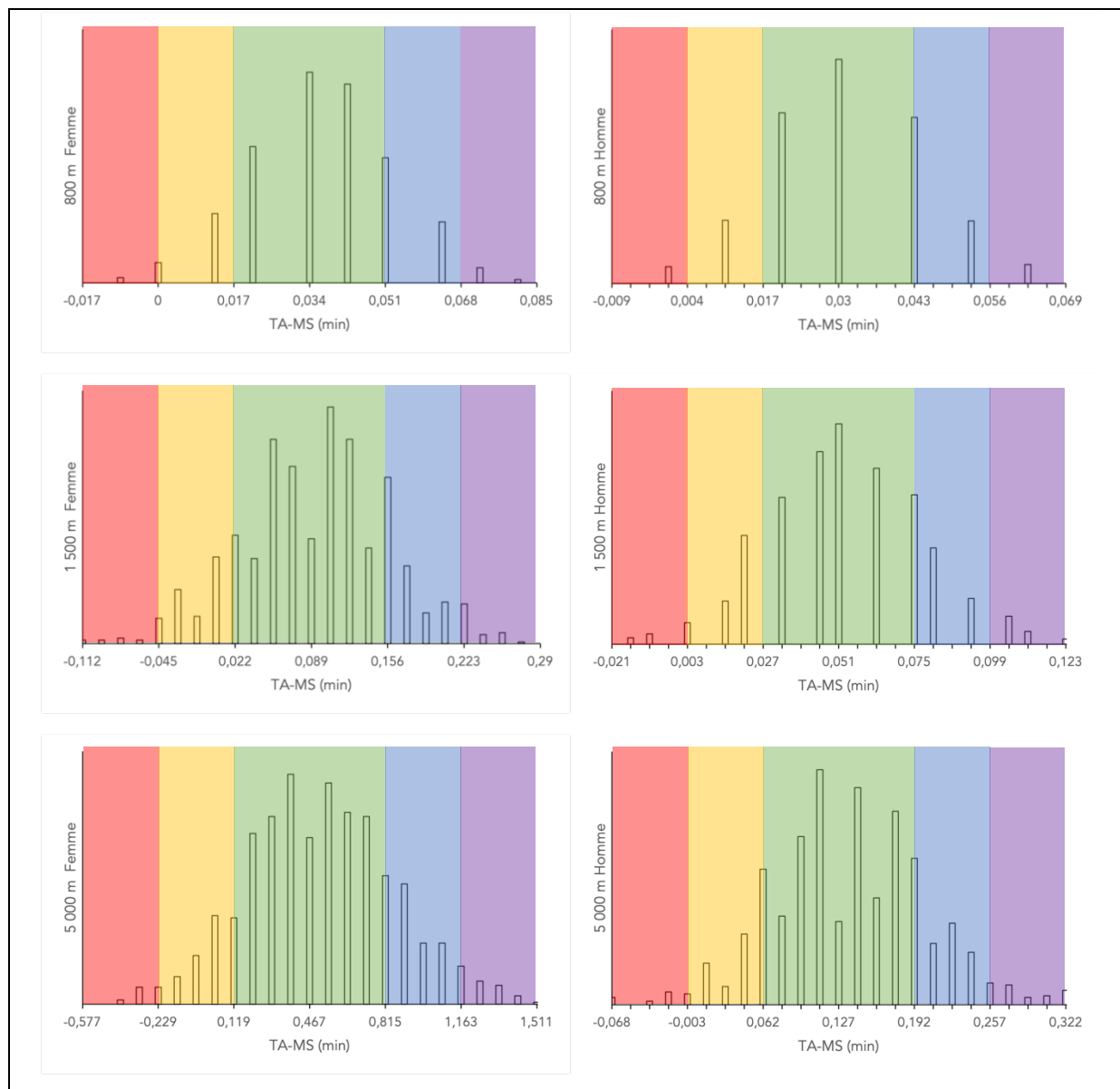


Figure 48. Index du taux d'amélioration pour les courses de 800 m, 1 500 m et 5 000 m pour les femmes et les hommes

Légende : SD : écart-type ; TA-MS : moyenne du taux d'amélioration multi saison (min). Zone verte (± 1 SD) : amélioration dans la moyenne des coureurs du Top 10 ; Zone jaune (- 2 SD) : amélioration inférieure à la moyenne des coureurs du Top 10, un changement de stratégie d'entraînement doit être envisagé ; Zone rouge (- 3 SD) : amélioration très inférieure à la moyenne des coureurs du Top 10, un changement majeur de stratégie d'entraînement doit être envisagé ; Zone bleue (+ 2 SD) : amélioration supérieure à la moyenne des coureurs du Top 10, la stratégie d'entraînement est efficace et les causes possibles doivent être étudiées ; Zone violette (+ 3 DS) : amélioration très supérieure à la moyenne des coureurs du Top 10, la stratégie d'entraînement est très efficace et les causes possibles doivent être étudiées.

4. Discussion

L'objectif principal de cette étude était d'analyser le taux d'amélioration de la performance des athlètes composant le Top 10 national français dans les épreuves de demi-fond (e.g., 800 m, 1 500 m et 5 000 m) à travers une analyse de leur meilleure performance personnelle par saison sur plusieurs années. Ensuite, l'objectif était de proposer un index de performance, permettant d'évaluer ce taux d'amélioration de la performance optimale, qui serait utilisé comme un outil pour les entraîneurs désireux d'évaluer une progression chez des athlètes, de prédire le potentiel et éventuellement de détecter les athlètes à haut potentiel.

Bien que les résultats généraux de cette étude aient montré un taux de progression plus élevé chez les femmes que chez les hommes, la différence entre les sexes était significative pour le TA-MS (%) sur 5 000 m uniquement. En effet, les résultats ont montré que les femmes ont amélioré en moyenne leur performance à hauteur de 3,10% avant d'atteindre le Top 10 national « tous temps », alors que les hommes ont eu une amélioration moyenne de 0,97%.

Naturellement et conformément aux études précédentes [10, 13, 22, 24], les résultats de la présente étude ont montré que pour les mêmes distances de course (i.e., 800 m, 1 500 m et 5 000 m), les femmes ont des performances (i.e., MP et MP-MS) significativement inférieures à celles des hommes, probablement en raison de leur $\dot{V}O_{2max}$, de leur capacité d'endurance aérobie plus faibles et des différences physiologiques existantes (e.g., force musculaire plus faible et masse grasse plus importante) [116, 119]. Cependant, la comparaison du taux d'amélioration entre les hommes et les femmes pour toutes les distances a montré, globalement, un taux d'amélioration plus élevé chez les femmes, notamment sur le 5 000 m. Au-delà des différences physiologiques entre les femmes et les hommes, le contexte historique, les facteurs socio-culturels et politiques, [3, 61], les facteurs psychologiques (e.g.,

motivation, résistance psychologique, gestion de l'anxiété et du stress) [9, 105, 182] voire tactiques (e.g., gestion d'allure, stratégie de course) [12, 205, 344] pourraient jouer un rôle potentiellement important sur les performances. Nous ne pouvons que spéculer sur ces explications possibles des différences de progression entre les sexes, mais il est possible que les changements dans l'acceptation sociétale de la course à pied féminine soient en cause et qu'ils aient entraîné plus d'opportunités d'entraînement et de participation à des compétitions afin de permettre aux femmes d'améliorer davantage leur record [1, 3, 10, 16, 61]. Bien qu'il n'y ait pas de différence significative dans la présente étude, le nombre de performances moyennes enregistrées pour les coureurs de 5 000 m était plus faible chez les femmes que chez les hommes (*i.e.*, 2,3 et 4,1 performances, respectivement) et pouvant ainsi être l'une des explications possibles de cette différence dans le taux d'amélioration des performances entre les sexes [1]. Les femmes peuvent être plus réticentes à courir cette distance que les hommes, mais lorsqu'elles se sont entraînées pour courir cette distance en compétition, elles ont créé un écart plus important avec leur dernière MP. Les résultats actuels relatifs aux différences entre les sexes en matière de taux de progression sont principalement conformes à ceux rapportés par des études précédentes sur différentes distances athlétiques [3, 10, 13, 22, 93, 342]. Par exemple, Haugen et al. [13] ont observé une plus grande amélioration chez les femmes que chez les hommes sur 800 m et 1 500 m à hauteur de $0,4 \pm 0,2\%$ mais aussi sur des distances plus longues telles que le 5 000 m et à hauteur de $1,2 \pm 0,2\%$, entre 2002 et 2016. Toutefois, ces études antérieures [3, 10, 13, 22, 93, 342] ont essentiellement étudié des performances d'athlètes de haut niveau (*i.e.*, de niveau mondial ; e.g., Top 12-16 pour Hollings et al. [93] ou top 10 ou top 11-100 pour Haugen et al. [13]), alors que les présentes analyses ont inclus les meilleurs résultats de la saison du Top 10 français de tous les temps, ce qui pourrait expliquer des différences dans les valeurs de progression des performances. De plus, si les données peuvent être altérées par le nombre plus

faible de performances féminines, notamment sur la distance la plus longue comme le 5 000 m, les différences de taux d'amélioration pourraient également s'expliquer par les années de performance étudiées. Dans cette étude, aucune limite de temps par athlète n'a été fixée (si ce n'est que les performances qui n'ont pas été réalisées dans la continuité des saisons précédant la MP n'ont pas été prises en compte) contrairement aux études de Haugen et al. [13] ayant analysé les meilleures performances individuelles des athlètes classés au niveau mondial de 2002 à 2016 ou de Cheuvront et al. [10] ayant analysé les performances de course des records du monde sur 1500 m et 5000 m des années allant de 1910 à 2004, par exemple. En effet, il semble important de ne pas exclure la professionnalisation, la recherche, mais aussi les avancées technologiques conduisant à une efficacité accrue de certains médicaments androgènes sur la physiologie des athlètes féminines dans le taux d'amélioration des performances. En fait, les améliorations technologiques des dernières décennies (e.g., les chaussures de course telles que les chaussures de *running* ou les pointes d'athlétisme) sont probablement un facteur important à prendre en compte [179, 181]. Cependant, même si au cours de la dernière décennie, certaines études ont montré que les performances des femmes ont tendance à s'améliorer davantage que celles des hommes et à rattraper une partie de l'écart entre les sexes [13], certains travaux antérieurs suggèrent que les différences entre les sexes ont atteint un plateau et que les spéculations selon lesquelles les femmes se rapprochent des temps des hommes ne sont pas soutenues [120, 345]. En effet, il a été suggéré précédemment que les différences de performance en course à pied (i.e., temps de course) entre les sexes diminuaient avec l'augmentation des distances [346, 347]. Bien que, la différence physiologique entre les sexes soit à l'avantage des hommes dans la performance de la course à pied (e.g., $\dot{V}O_{2max}$) [10, 116, 119], certains facteurs physiologiques en faveur des femmes peuvent également entrer en jeu, comme une meilleure tolérance aux conditions de course chaudes et humides (en raison de leur taille plus petite) [348,

349], et une meilleure utilisation des substrats lors d'un exercice d'endurance (consommation plus élevée des lipides et consommation plus faible de glucides pour les athlètes entraînées) [350]. Cependant, des études plus récentes [10, 22] contrairement à ce qui était suggéré, ont montré que l'écart de performance augmente avec la distance plutôt que de diminuer sur des distances supérieures au marathon. Étant donné que les paramètres physiologiques clés [11, 14, 132, 351] sont dépendants du sexe [17, 319] et déterminants sur les courses de demi-fond (telles que le 800 m, 1 500 m, 5 000 m), cette explication semble relativement limitée et plus adaptée aux courses de fond ou d'ultra-fond (supérieures au marathon).

Enfin, les résultats de cette étude ont montré un taux d'amélioration significativement plus élevé (*i.e.*, TA-1S et TA-MS en min) sur la plus longue distance étudiée principalement (*i.e.*, 5 000 m) que sur les plus courtes distances (*i.e.*, 800 m et 1 500 m). Ces résultats concordent avec la première étude réalisée durant ces travaux de recherche (*cf.* deuxième partie, chapitre I) ayant montré une augmentation significative du niveau de performance des femmes sur les plus longues distances sur piste telles que le 5 000 m. L'expertise en course à pied permettrait de progresser sur de plus longues distances, notamment grâce à une meilleure gestion de l'allure de course ou encore au choix d'une stratégie de course optimale, et permettrait ainsi d'expliquer un pic de performance à un âge plus avancé sur les plus longues distances.

La mesure du taux d'amélioration de la performance proposée pour connaître le gain annuel de performance des athlètes figurant dans le Top 10 français pourrait aider les entraîneurs ou les athlètes dans leur pratique afin d'optimiser la progression [342]. En effet, pour les athlètes qui se situent dans une dynamique de progression (Figure 48), il est raisonnable de supposer que la stratégie d'entraînement est efficace [342]. Cependant, les athlètes ne parviennent pas

toujours à améliorer leurs performances chaque année, et pour ceux qui affichent une mauvaise dynamique ou une contre-progression (Figure 48), un changement de stratégie d'entraînement doit être envisagé [342].

La présente étude vise à aider les entraîneurs de demi-fond qui collaborent avec des athlètes masculins et féminins (*i.e.*, 800 m, 1 500 m et 5 000 m). Par conséquent, en accord avec l'étude de Santos et al. [342], des graphiques permettant l'analyse de taux d'amélioration des performances ont été générés, afin de permettre aux entraîneurs voire aux athlètes de vérifier si la méthode d'entraînement mise en place génère un taux d'amélioration de la performance comparable aux athlètes figurant dans le Top 10 français, voire un taux d'amélioration comparable au détenteur du record national sur la distance choisie (Figure 48).

Pour appliquer les résultats de l'index de performance, un athlète doit considérer sa MP de la saison en cours ainsi que chaque MP réalisée lors des saisons précédentes permettant de calculer ainsi les TA-1S et TA-MS. Les taux d'amélioration de la performance peuvent ensuite être comparés aux dix meilleurs coureurs nationaux de tous les temps dans la discipline choisie (Figure 48, Tableau 26). Si l'on prend pour exemple, un coureur de 5 000 m ayant couru cette distance lors de 6 saisons, et ayant réalisé une MP de 13,25 min lors de la saison en cours (*i.e.*, 13 min 15 s) et que l'on considère les résultats affichés dans le Tableau 26 (*i.e.*, TA-MS de 0,32 min, soit 19,2 s) ainsi que l'index présenté en Figure 48, le TA-MS de cet homme se situe dans la zone « violette ». Cela signifie que la stratégie d'entraînement semble être très efficace, et qu'une projection tenant compte de son âge et de l'épreuve de course visée pourrait être planifiée.

Tableau 26. Exemple d'un coureur de 5 000 m homme (sur 6 ans)

Année	Performance (min)	Taux d'amélioration (min)
2015	14,87	-
2016	13,92	0,95
2017	13,85	0,07
2018	13,63	0,22
2019	13,38	0,25
2020 (MP)	13,27	0,11 (TA-1S)
Taux moyen d'amélioration		0,32 (TA-MS)

La principale limite de cette étude était que les performances en course à pied n'ont pas été réalisées sur la même échelle de temps (*i.e.*, sans limite de temps par athlète, les meilleurs coureurs de tous les temps figurant dans le Top 10 français) et n'ont pas été soumises à des conditions technologiques ou environnementales similaires, avec des conditions physiques et psychologiques (*e.g.*, motivation et stress) identiques, et en utilisant le même terrain d'opposition et/ou la même stratégie de course. Une autre limite de l'étude pourrait être le nombre limité de performances enregistrées pour certains athlètes, pour diverses raisons (*e.g.*, blessures, raisons personnelles, changement de distance athlétique, conduite dopante), permettant de calculer les taux de progression (*i.e.*, TA-1S ou TA-MS).

5. Conclusion

Les athlètes féminines figurant dans le Top 10 français de tous les temps des courses de demi-fond ont montré des taux d'amélioration supérieurs à ceux des athlètes masculins avant d'atteindre leur record personnel. Quel que soit le sexe, lorsque les distances étaient comparées, le taux d'amélioration était plus élevé sur les distances les plus longues. En outre, cette étude fournit des indications qui peuvent aider les athlètes et les entraîneurs à optimiser les performances et/ou qui cherchent à atteindre le Top 10 national voire le record de France sur 800 m, 1 500 m et 5 000 m. L'index proposé peut aider à déterminer si certains ajustements sont nécessaires

dans le processus d'entraînement. Les recherches futures pourraient se focaliser sur d'autres distances de course à pied telles que les courses sur route de 10 km, de semi-marathon et de marathon.

DISCUSSION GÉNÉRALE

L'objectif initial de cette thèse CIFRE était d'examiner le possible effet du port d'orthèses plantaires sur la récupération du membre inférieur en période de post-effort chez des sportifs. Après la réalisation d'une première revue de littérature systématique sur ce sujet, qui a été publiée [57] (cf. Annexe 1), et le développement d'une gamme de sandales orthopédiques « sur-mesure » (cf. Annexe 8), nous avons mis au point deux expérimentations visant à tester l'hypothèse que les semelles et sandales de récupération permettraient d'améliorer la récupération et la performance sportive de golfeurs et de coureurs à pied. Ces protocoles expérimentaux ont ensuite obtenu l'accord d'un comité d'éthique (IRB00012476-2020-15-07-62). En outre, dans le cadre des expérimentations, l'obtention d'un « Bonus Qualité Recherche » a permis l'achat d'un pléthysmographe (Pletix, Callegari, Parma, Italia). Cependant, ces expérimentations ont à deux reprises été reportées du fait des nombreux confinements liés à la crise sanitaire due à la pandémie de la COVID-19. En plus de ces reports, les contraintes imposées par la gestion de la crise ne nous ont pas permis de rassembler suffisamment de sujets. De ce fait, ces études sont encore en cours de réalisation. Néanmoins pour pouvoir soutenir cette thèse dans les temps, nous avons recentré notre travail sur la prédiction de la performance en course à pied. L'analyse de la littérature, effectuée dans le cadre de cette thématique de recherche, a notamment mis en avant les multiples paramètres (*i.e.*, physiologiques, biomécaniques, psychologiques, environnementaux et technologiques) à prendre en considération afin d'appréhender et de comprendre la performance en course à pied dans son ensemble [3, 5, 6, 9-19]. Aux prémices de ce travail doctoral a été étudiée l'évolution des performances en courses de demi-fond et fond chez les athlètes féminines françaises. De cette première étude, ont pu être constatées une augmentation du nombre de performances françaises enregistrées chez les femmes ainsi qu'une

amélioration de ces performances sur les plus longues distances disputées officiellement sur piste (5 000 m et 10 000 m). Comme discuté dans cette première étude, l'amélioration des performances pourrait être expliquée par des paramètres physiologiques liés à l'âge caractérisant les capacités humaines à produire une puissance aérobie ou à développer une capacité d'endurance aérobie sur une distance donnée, mais aussi par des paramètres psychologiques où les connaissances et l'expérience de course évoluent avec la maturité de l'athlète [93-95, 127]. Les paramètres environnementaux tels que la mise en place d'un plan de développement territorial par la FFA, la restructuration des clubs d'athlétisme, la communication pour promouvoir l'athlétisme, le développement des compétitions afin d'offrir davantage de densités aux athlètes, ont également été discutés. Ces facteurs pourraient être une explication plausible à l'augmentation et l'amélioration des performances françaises en demi-fond et fond [89, 302, 303]. Il est, en effet, bien établi que les facteurs environnementaux (e.g., mode de vie, météorologie, société, infrastructure, règlement...) peuvent impacter la performance et son évolution en course à pied [3, 83, 74, 6, 9, 105, 206]. Au vu de l'enjeu que représente la performance sportive à l'heure actuelle (e.g., préparation pour les JO de Paris 2024), il nous a semblé intéressant de réaliser des études complémentaires portant sur la prédiction de la performance voire sur la prédiction du potentiel de performance chez des athlètes pratiquant la course à pied. Dans la littérature antérieure, sur la base de comparaisons de différents modèles (e.g., logarithmique, exponentiel, hyperbolique, loi de puissance), procédures (e.g., méthode d'interpolation ou d'extrapolation) et variables intégrées (e.g., distance de course, sexe, niveau de performance, nombre de performances), il a été observé des résultats de prédiction fluctuant selon les modèles en termes de justesse et de précision par rapport aux temps réels de performance à l'avantage généralement des modèles complexes (i.e., multiparamètres) [28, 31, 261]. Toutefois, il a été relevé que l'application relativement difficile des modèles complexes (e.g., modèle

hyperbolique à trois paramètres), en lien par exemple avec le nombre de données (*i.e.*, multiples performances maximales) à intégrer et à ajuster, pouvait constituer vraisemblablement l'une des principales limites d'utilisation en comparaison aux modèles intégrant un plus faible nombre de paramètres (*e.g.*, formule de Riegel, modèle hyperbolique à deux paramètres, modèle de loi de puissance à deux paramètres...) mais souvent plus limités par la justesse et la précision des prédictions [28, 31, 261]. De nombreux modèles ont principalement ciblé un profil d'athlète en particulier (*e.g.*, athlètes de niveau élite) pour modéliser et tenter de prédire la performance en course à pied. De manière générale, les records du monde traduisent l'optimisation des capacités athlétiques de l'espèce humaine. Cependant, au regard des travaux réalisés antérieurement il semblait intéressant de tester et de comparer différents modèles de prédiction de la performance en course à pied chez des athlètes de profil hétérogène [279]. Dans le cadre de ces travaux, la convention de recherche obtenue par mon directeur de thèse avec la FFA a notamment permis l'accès à une grande base de données. Cette convention a donné la possibilité de collecter, d'analyser et d'interpréter un grand nombre de performances françaises féminines et masculines, quels que soient l'âge et le niveau de pratique des athlètes. Les données exploitées durant ce doctorat ont ainsi permis de tester et de comparer des approches ou modèles mathématiques simples (*e.g.*, nomogrammes, modèle linéaire de vitesse critique) mais aussi relativement complexes (*e.g.*, IA) qui ont été peu étudiés, voire non investigués à ce jour dans le domaine de la course à pied sur certaines distances. Les résultats obtenus (*cf.* deuxième partie, chapitres II à V) ont notamment montré que l'ensemble des méthodes de prédiction testées sur les distances allant du demi-fond au fond (*e.g.*, 3 000 m, 5 000 m, 10 000 m, 3 000 m *steep*/*e* et marathon) chez des athlètes de tous niveaux (*i.e.*, du niveau régional à international) et de sexe féminin et/ou masculin étaient valides et qu'elles pouvaient prédire avec précision les performances en course à pied (Tableau 27).

Tableau 27. Récapitulatif des biais et LoA à 95% pour l'ensemble des méthodes testées : nomogrammes (de Mercier et al. [26] et de Vandewalle [27]), MLDT, loi de puissance, équation simple et IA (ANN et KNN)

Distance	Nomogramme			MLDT	Loi de puissance	Équation simple	IA	
	Mercier et al.	1 ^{er} de Vandewalle	2 ^{ème} de Vandewalle				ANN	KNN
3 000 m	-0,0 ± 6,2	0,1 ± 6,0	0,1 ± 6,0	-3,7 ± 10,8	-0,1 ± 6,3	-	-	-
5 000 m	0,0 ± 3,7	-0,1 ± 3,6	0,0 ± 3,7	1,4 ± 4,3	0,1 ± 3,7	-	-	-
10 000 m	-0,1 ± 8,9	0,2 ± 9,3	0,1 ± 9,3	-2,5 ± 7,4	-0,1 ± 8,9	-	-	-
3 000 m steeple	-	-	-	-	-	-0,5 ± 7,1	-	-
Marathon	-	-	-	-	-	-	1,5 ± 14,1	-0,4 ± 7,2

Légende : Biais et LoA exprimés en %. MLDT (modèle linéaire distance-temps), ANN (*artificial neural networks*), KNN (*k-nearest-neighbors*), LoA (limites d'agrément)

Si pour chaque méthode testée, les résultats ont été discutés individuellement dans les chapitres précédents, il semblait intéressant de mettre en avant dans cette discussion générale les principales connaissances développées ainsi que les principaux intérêts de chaque approche étudiée afin de pouvoir guider les athlètes et les entraîneurs vers la prédiction voire l'optimisation de la performance en course à pied.

Conformément aux travaux réalisés antérieurement, les nomogrammes, le concept de vitesse critique et la loi de puissance ont eu pour avantage de permettre la prédiction d'une performance sur une distance donnée à partir de performances sur deux autres distances [24, 25, 31, 32, 270]. En accord également avec les études précédentes ces résultats valides ont soutenu le fait que l'exactitude et la précision des prédictions étaient accrues lorsque ces dernières étaient réalisées par interpolation plutôt que par extrapolation [24, 25, 27, 265, 270]. Bien que les prédictions aient été considérées comme justes et précises dans l'ensemble, les résultats ont indiqué un avantage pratique des nomogrammes (en comparaison au modèle linéaire de vitesse critique et de la loi de puissance), et plus particulièrement en faveur du nomogramme de Mercier et al. [26] lorsqu'il s'agissait de prédire une performance sur une longue distance comme le 10 000 m. Ces résultats pourraient

notamment être expliqués par les différences de conception entre les nomogrammes étudiés (*i.e.*, nomogramme de Mercier et al. [26] développé à partir de connaissances empiriques et nomogrammes de Vandewalle [27] basés sur le modèle logarithmique de Péronnet et Thibault [112]), qui jusqu'à présent n'avaient pas fait l'objet de comparaison. En outre, à la différence de l'étude de Vandewalle [27] consacrée uniquement aux athlètes élités de sexe masculin, il est important d'indiquer que les résultats obtenus (de l'étude comparant les différents nomogrammes) se sont révélés être également valides pour des athlètes féminines indépendamment de leur niveau de course à pied. Enfin, il semble aussi intéressant de souligner le fait que toutes les prédictions (*via* les nomogrammes, le modèle linéaire de vitesse critique et la loi de puissance) ont été réalisées à l'aide de programmes (ou de formules) mathématiques assistés par ordinateur, éliminant ainsi le risque potentiel d'erreur humaine [24], ce qui pourrait par conséquent être l'une des raisons plausibles d'une précision de prédiction légèrement supérieure dans les résultats obtenus.

Bien que certaines études aient validé des méthodes de prédiction de performances sur 3 000 m piste (*e.g.*, le nomogramme, le concept de vitesse critique) [24, 25, 270], il est apparu qu'aucune étude n'ait tenté de prédire la performance sur 3 000 m *steeple*. Les résultats de prédiction obtenus suite au développement d'une équation simple (*i.e.*, $Perf_{Steeple} = -89,740 + 1,153 \times Perf_{3000} + 0,952 \times \text{âge} + 14,810 \times \text{sexe}$) à partir de variables individuelles (*e.g.*, $Perf_{3000}$, sexe, âge), ont suggéré que cette méthode était valide et précise. En effet, malgré la dimension technique liée à la présence de barrières et de rivières différenciant le 3 000 m plat du 3 000 m *steeple*, et bien que la prédiction des performances sur *steeple* ait été réalisée à partir d'une autre méthode que celle utilisée pour la distance homologue sans obstacle, les résultats de cette étude étaient relativement comparables avec ceux obtenus dans des études antérieures portant sur le 3 000 m piste [24, 25, 270]. Comme discuté dans les chapitres

précédents, la justesse et la précision obtenues dans les prédictions pouvaient être liées au choix des variables individuelles composant cette équation comme la $Perf_{3000}$ (principal prédicteur), le sexe et l'âge. En effet, l'intensité et la durée de l'effort des épreuves de 3 000 m et 3 000 m *steeple* étant relativement proches (*i.e.*, prédicteurs physiologiques communs) [80, 81, 321], la variable $Perf_{3000}$ semblait largement justifiable. En outre, la hauteur et la technique de course pour franchir les barrières et les rivières (*e.g.*, vitesse d'approche vers l'obstacle, zone de réception après la barrière, angle des segments du membre inférieur lors de l'appui sur la rivière) [82], ainsi que les modifications physiologiques (en lien avec le vieillissement) caractérisant les capacités humaines [95, 127, 157], étaient autant d'arguments qui pouvaient respectivement justifier la présence des variables prédictives du sexe et de l'âge dans l'équation de prédiction des performances sur 3 000 m *steeple*.

Au même titre que pour l'équation simple développée pour prédire la performance sur 3 000 m *steeple*, le recours à des variables indépendantes a notamment permis d'obtenir des résultats valides, justes et précis pour prédire la performance sur l'épreuve de marathon en utilisant notamment deux techniques d'IA (ANN et KNN). En effet, au-delà de l'IMC qui différait des variables retenues pour l'équation simple, les prédictions *via* l'utilisation d'algorithmes d'apprentissage supervisé étaient également réalisées à partir d'une performance sur une distance donnée (*i.e.*, 10 km), de l'âge et du sexe. Malgré le fait que les algorithmes d'ANN et de KNN n'aient pas fait l'objet d'étude pour prédire la performance sur marathon, les résultats obtenus étaient comparables à ceux d'une autre étude [23] ayant utilisé l'IA chez des coureurs élités sur cette même distance. En outre, si dans le cadre de cette étude les prédictions obtenues sont apparues plus précises avec KNN que ANN, il a toutefois été souligné dans la littérature antérieure (dans des champs d'application différents : algorithmes de classification dans le domaine médical, la météorologie...) qu'il n'y avait pas de réel consensus sur le meilleur modèle

d'algorithmes d'IA, chaque modèle étant étroitement sensible aux données et à la méthode du « *data scientist* » définissant et modélisant, respectivement ces algorithmes [269, 292, 293, 317, 340]. À noter également que si la prédiction de performance sur marathon s'est avérée valide et relativement précise avec d'autres méthodes que l'IA (e.g., nomogramme, formule de Riegel) [17, 265], il semble très intéressant de souligner le fait que dans cette présente étude, malgré la complexité d'application des algorithmes d'IA, les prédictions n'ont été obtenues qu'à partir d'une seule performance de course à pied (*i.e.*, 10 km) et qui plus est à partir une épreuve relativement éloignée en termes de distance et de durée de l'épreuve prédite (*i.e.*, marathon).

Enfin, dans la continuité des travaux menés durant ce doctorat, les objectifs de la dernière étude étaient d'appréhender l'accession à un record personnel *via* une analyse du taux d'amélioration de la performance, basée sur la MP établie par saison sur plusieurs années, des athlètes composant le Top 10 national dans les épreuves de demi-fond ; puis de proposer un index, permettant l'évaluation de ce taux d'amélioration de la performance optimale, comme réalisé par Santos et al. [342] sur des distances de sprint chez des coureurs composant le Top 10 mondial. Si les taux d'amélioration se sont avérés supérieurs chez les athlètes de sexe féminin et conformes à ceux rapportés par des études antérieures sur différentes distances athlétiques [3, 10, 13, 22, 93, 342], il a également été montré qu'indépendamment du sexe, ces taux étaient plus élevés sur les distances les plus longues ce qui rejoint notamment les résultats obtenus lors de la première étude (*i.e.*, augmentation significative du niveau de performance des femmes sur les plus longues distances sur piste ; *cf.* deuxième partie, chapitre I) [313]. Cette dernière étude a notamment fourni des indications, *via* l'identification de plages de taux d'amélioration, que l'on pourrait également qualifier de « couloirs de performances », dans l'objectif d'aider les athlètes et les entraîneurs à optimiser la stratégie d'entraînement en demi-fond. En effet, l'index généré à partir des données réelles des meilleurs athlètes français

pourrait être utilisé comme un outil pour les entraîneurs désireux d'évaluer une progression chez des athlètes, de prédire leur potentiel et éventuellement de détecter les athlètes au fort potentiel (e.g., médaillables en grands championnats). Cependant, si les études menées durant ce doctorat avaient pour objectif d'élargir le champ de recherche en apportant notamment de nouvelles connaissances dans le domaine de la prédiction de la performance en course à pied, certaines limites générales ont pu être soulignées. En effet, pour l'ensemble des travaux réalisés, analysant soit l'amélioration et/ou la prédiction des performances en course à pied, nous ne pouvons négliger le fait que les performances de course réelles (collectées depuis la base de données FFA) n'aient pas été réalisées de manière exhaustive pour la distance donnée (e.g., course tactique) ou n'aient pas été réalisées dans des conditions environnementales similaires (e.g., mêmes conditions climatiques et altitude), dans des conditions physiologiques (e.g., capacité d'endurance aérobie, économie de course, vitesse de course) et psychologiques (e.g., motivation et stress) identiques, et en utilisant le même lieu d'opposition et/ou la même stratégie de course. En outre, il n'était pas impossible que les variables susmentionnées aient pu affecter positivement ou négativement la justesse et la précision des prédictions. De manière plus spécifique en fonction des études menées, d'autres limites ont pu être évoquées telles que le choix de la procédure de modélisation définissant les algorithmes, l'autodéclaration de certaines données (i.e., taille et masse corporelle renseignées directement par les athlètes) sur la base de données FFA, notamment pour les études ayant intégré la variable de l'IMC. En outre, la prise en considération des meilleures performances réalisées sur des périodes de temps espacées (i.e., bilan FFA « tous temps ») laissant parfois envisager des conditions technologiques peu comparables dans le temps (e.g., chaussures de course, mode de récupération...) ou encore le nombre de performances enregistrées relativement limité pour certains athlètes pour diverses raisons possibles (e.g., blessures, raisons

personnelles, changement de distance athlétique, conduite dopante...), sont autant de limites qui peuvent être discutées.

Par conséquent, en identifiant et considérant ces différentes limites, au regard de l'ensemble des résultats obtenus, sur la base de validation et de comparaison de différents modèles (*i.e.*, nomogramme de Mercier et al. [26], nomogrammes de Vandewalle [27], modèle linéaire de vitesse critique, loi de puissance, équation simple, ANN et KNN) ces travaux de thèse ont pu montrer, qu'au-delà de la validité des méthodes testées, que la justesse et la précision des performances prédites étaient étroitement liées aux différents paramètres définissant les modèles (*e.g.*, distance de course, prédiction par interpolation ou extrapolation, nombre de variables individuelles intégrées...), et étaient en l'occurrence conformes aux conclusions générales apportées par des travaux antérieurs dans ce domaine [28, 31, 261]. De ce fait les modèles de prédiction de la performance en course à pied ne doivent pas être considérés comme une vérité absolue et sont assujettis aux différents facteurs (*e.g.*, physiologiques, biomécaniques, psychologiques, environnementaux et technologiques) associés à la performance. Certaines performances prédites pourraient, par exemple, s'avérer être relativement très éloignées de la réalité sans pour autant que le modèle développé fasse l'objet d'une remise en question. De même, dans une optique de détection de potentiel de performance, certains athlètes pourraient être amenés à ne pas correspondre, voire échapper, aux modèles établis sur la base d'autres athlètes aux caractéristiques possiblement éloignées. Tenter de prédire la performance en course à pied n'est pas sans poser de problèmes liés aux intérêts et limites, aux choix d'approche pour l'obtention de résultats valides, ainsi qu'aux connaissances développées. La prédiction de la performance s'établit sur l'identification des « bons » paramètres permettant le développement des meilleures méthodes valides et précises capables de s'appliquer à différents échantillons tels que les coureurs de profils variés (*e.g.*, sexe, âge, niveau d'expertise). En outre, et comme évoqué

antérieurement, si les athlètes de catégorie « *senior* », indépendamment de leur niveau en course à pied, ont principalement été ciblés dans les premiers travaux menés, par exemple, nous pourrions justifier le fait que ces derniers âgés de 23 à 34 ans sont théoriquement plus matures physiologiquement que leurs homologues plus jeunes (*i.e.*, cadets, *juniors* ou espoirs) en demi-fond et de fond [13, 93-95] et qu'ils sont majoritairement impliqués et désireux de se dépasser et/ou performer lorsque ces derniers choisissent de pratiquer en compétition.

Les perspectives de ce travail sont multiples. Compte tenu du fait que la prédiction de performance (ou la prédiction du potentiel de performance) en course à pied peut être utilisée à des fins explicatives (*i.e.*, appréhender les facteurs de performance) et interventionnelles (*i.e.*, détecter de jeunes athlètes au fort potentiel et orienter l'entraînement), il pourrait être intéressant d'exploiter davantage la base de données de la FFA voire celle de *World Athletics* en appliquant les modèles, précédemment testés ou développés, sur de nouveaux échantillons de données (*e.g.*, athlètes en catégorie jeunes, athlètes spécialisés en ultra-fond). Le recours à d'autres modèles existants voire d'autres paramètres de performances ainsi que le développement de nouvelles approches de prédiction dans une optique d'optimisation de la performance en course à pied pourraient également être envisageables. Enfin, si la stagnation chronométrique des performances en course à pied (justifiée principalement par l'atteinte de limites physiologiques humaines) a pu être soulignée et discutée dans différents travaux antérieurs [16, 73, 74, 115, 120, 124], il apparaît que les progressions des performances dans différentes épreuves de course récemment constatées, possiblement liées aux innovations technologiques, puissent offrir également de belles perspectives de travail...

CONCLUSION GÉNÉRALE

Appréhender et comprendre la performance en course à pied afin de tenter de prédire la performance (ou le potentiel), tels étaient les enjeux de ce travail doctoral. La prédiction de la performance peut être utile pour planifier des séances d'entraînement, déterminer des allures de course en compétition et détecter des athlètes disposant d'un fort potentiel. Les études menées, non interventionnelles, ont notamment bénéficié d'une convention avec la FFA permettant l'accès à une large base de données répertoriant l'ensemble des performances athlétiques, dont les disciplines de course à pied. Les épreuves de demi-fond et de fond, allant du 800 m au marathon, ont particulièrement été ciblées. Différents modèles de prédiction ont été analysés afin de tester leur validité, leur justesse et leur précision sur des performances individuelles chez des coureurs hommes et/ou femmes, de niveau départemental à international, pratiquant la course à pied sur piste ou sur route. Les résultats se sont avérés valides et précis pour l'ensemble des approches testées, et les avantages ainsi que les limites de chaque modèle ont été soulignés et discutés. Un index de progression à visée pratique afin d'évaluer l'évolution des performances, d'évaluer la stratégie d'entraînement mise en place et permettre une éventuelle détection et orientation des athlètes au fort potentiel a également été développé. En l'occurrence, les études réalisées avaient pour objectifs de fournir différentes méthodologies de prédiction de performance et d'apporter de nouvelles connaissances aux athlètes et entraîneurs de course à pied.

RÉFÉRENCES

- [1] Charlet S. *L'Athlétisme à travers les siècles*. Saint-Cyr-Sur-Loire: Editions Alan Sutton, 2004.
- [2] FFA. Histoire de l'athlétisme. *Fédération Française d'Athlétisme*, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=418> (accessed 11 July 2021).
- [3] Berthelot G, Sedeaud A, Marck A, et al. Has athletic performance reached its peak? *Sports Med* 2015; 45: 1263-1271.
- [4] Gembris D, Taylor JG, Suter D. Evolution of athletic records: statistical effects versus real improvements. *J Appl Stat* 2007; 34: 529-545.
- [5] El Helou N, Tafflet M, Berthelot G, et al. Impact of environmental parameters on marathon running performance. *PloS One* 2012; 7: e37407.
- [6] Haïda A, Dor F, Guillaume M, et al. Environment and scheduling effects on sprint and middle distance running performances. *PloS One* 2013; 8: e79548.
- [7] Marc A, Sedeaud A, Guillaume M, et al. Marathon progress: demography, morphology and environment. *J Sports Sci* 2014; 32: 524-532.
- [8] Billat V. *Physiologie et méthodologie de l'entraînement*. De Boeck Supérieur, 2017.
- [9] Smith DJ. A framework for understanding the training process leading to elite performance. *Sports Med Auckl NZ* 2003; 33: 1103-1126.
- [10] Cheuvront SN, Carter R, Deruisseau KC, et al. Running performance differences between men and women:an update. *Sports Med Auckl NZ* 2005; 35: 1017-1024.
- [11] Di Prampero PE, Atchou G, Brückner JC, et al. The energetics of endurance running. *Eur J Appl Physiol* 1986; 55: 259-266.
- [12] Filipas L, Nerli Ballati E, Bonato M, et al. Elite male and female 800-m runners' display of different pacing strategies during season-best performances. *Int J Sports*

Physiol Perform 2018; 1-5.

[13] Haugen TA, Solberg PA, Foster C, et al. Peak age and performance progression in world-class track-and-field athletes. *Int J Sports Physiol Perform* 2018; 13: 1122-1129.

[14] Kenney WL, Hodgson JL. Variables predictive of performance in elite middle-distance runners. *Br J Sports Med* 1985; 19: 207.

[15] Sedeaud A, Marc A, Marck A, et al. BMI, a performance parameter for speed improvement. *PloS One* 2014; 9: e90183.

[16] Thibault V, Guillaume M, Berthelot G, et al. Women and men in sport performance: the gender gap has not evolved since 1983. *J Sports Sci Med* 2010; 9: 214-223.

[17] Vickers AJ, Vertosick EA. An empirical study of race times in recreational endurance runners. *BMC Sports Sci Med Rehabil* 2016; 8: 26.

[18] Boullosa D, Esteve-Lanao J, Casado A, et al. Factors affecting training and physical performance in recreational endurance runners. *Sports* 2020; 8: 35.

[19] Salmela J, Durand-Bush N, Durand M. La détection des talents ou le développement de l'expertise en sport. *Enfance* 1994; 47: 233-245.

[20] Casado A, Hanley B, Jiménez-Reyes P, et al. Pacing profiles and tactical behaviors of elite runners. *J Sport Health Sci* 2021; 10: 537-549.

[21] Haïda A. *Blessure, environnement et performance de haut niveau*. Université de Rouen, France, 2014.

[22] Coast JR, Blevins JS, Wilson BA. Do gender differences in running performance disappear with distance? *Can J Appl Physiol Rev Can Physiol Appl* 2004; 29: 139-145.

[23] Blythe DAJ, Király FJ. Prediction and quantification of individual athletic performance of runners. *PloS One* 2016; 11: e0157257.

[24] Coquart JB, Mercier D, Tabben M, et al. Influence of sex and specialty on the prediction of middle-distance running performances using the Mercier et al.'s

nomogram. *J Sports Sci* 2015; 33: 1124-1131.

[25] Tabben M, Bosquet L, Coquart JB. Effect of performance level on the prediction of middle-distance-running performances using a nomogram. *Int J Sports Physiol Perform* 2016; 11: 623-626.

[26] Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le VO₂max et l'endurance relative en course de fond. *Méd Sport* 1984; 58: 181-187.

[27] Vandewalle H. A nomogram of performances in endurance running based on logarithmic model of péronnet-thibault. *Am J Eng Res* 2017; 6: 78-85.

[28] Gamelin FX, Coquart JM, Ferrari N, et al. Prediction of one-hour running performance using constant duration tests. *J Strength Cond Res* 2006; 20: 735-739.

[29] Hughson RL, Orok CJ, Staudt LE. A high velocity treadmill running test to assess endurance running potential. *Int J Sports Med* 1984; 5: 23-25.

[30] Kennelly AE. An approximate law of fatigue in the speeds of racing animals. In: *Proceedings of the American Academy of Arts and Sciences*. JSTOR, 1906, pp. 275-331.

[31] Vandewalle H. Modelling of running performances: comparisons of power-law, hyperbolic, logarithmic, and exponential models in elite endurance runners. *BioMed Res Int* 2018; 2018: 8203062.

[32] Zinoubi B, Vandewalle H, Driss T. Modeling of running performances in humans: comparison of power laws and critical speed. *J Strength Cond Res* 2017; 31: 1859-1867.

[33] Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. *J Physiol* 2008; 586: 35-44.

[34] Speed CA, Ingham SA. Research in high-performance sports medicine: from the bench, to the bedside...to the podium. *Br J Sports Med* 2011; 45: 608-610.

[35] Yarrow K, Brown P, Krakauer JW. Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nat Rev Neurosci* 2009; 10: 585-

596.

[36] MACRI C. Ministère des affaires sociales et de la santé.

[37] Lopez AA, Goldcher A. *Historique de la compréhension de la biomécanique du pied nu*. Elsevier Masson. Paris: Elsevier Masson SAS, Paris), Podologie, 2010.

[38] Viel E. Biomécanique des fonctions majeures du pied humain: amortissement, équilibre, propulsion et pivotement. *Ann Kinésithér* 1985; 12: 35-49.

[39] Skipper Jr JK, Hughes JE. Podiatry: A medical care specialty in quest of full professional status and recognition. *Soc Sci Med* 1983; 17: 1541-1548.

[40] Dananberg HJ. Functional hallux limitus and its relationship to gait efficiency. *J Am Podiatr Med Assoc* 1986; 76: 648-652.

[41] Eng JJ, Pierrynowski MR. The effect of soft foot orthotics on three-dimensional lower-limb kinematics during walking and running. *Phys Ther* 1994; 74: 836-844.

[42] Lysens R, Steverlynck A, van den Auweele Y, et al. The predictability of sports injuries. *Sports Med* 1984; 1: 6-10.

[43] Lavigne A, Noviel D. *Etude clinique du pied et thérapeutique par orthèse*. Masson, 1992.

[44] Lavigne Alain, Noviel Daniel. *Traité de la semelle orthopédique et autres orthèses en podologie : études cliniques des troubles de l'orthostatisme*. Boulogne (B.P. 26, 92101, cedex): SCERDES, 1981.

[45] Zing E. *Examen clinique élémentaire en podologie*. Paris: Elsevier Masson SAS, 2008.

[46] Goldcher A. Indications, principes d'action et de prescription des orthèses plantaires. *Rev Rhum Monogr* 2014; 81: 203-208.

[47] Goldcher A. Traitement médical de l'hallux valgus. *EMC Elsevier Masson SAS Paris Podol.*

[48] Zygas P. Les pathologies de l'arrière et de l'avant-pied. *Rev Med Brux* 2006; 27: 307-14.

- [49] Williams III DS, McClay IS, Hamill J. Arch structure and injury patterns in runners. *Clin Biomech* 2001; 16: 341-347.
- [50] Couzan S, Pouget JF. Le sportif: un insuffisant veineux potentiel. *Cardio Sport* 2006; 8: 7-20.
- [51] Kemoun G. Activité physique en pathologie vasculaire: indications et contre-indications. *EMC-Cardiol-Angéiologie* 2005; 2: 351-357.
- [52] Stick C, Hiedl U, Witzleb E. Venous pressure in the saphenous vein near the ankle during changes in posture and exercise at different ambient temperatures. *Eur J Appl Physiol* 1993; 66: 434-438.
- [53] Barnett A. Using recovery modalities between training sessions in elite athletes: does it help? *Sports Med Auckl NZ* 2006; 36: 781-796.
- [54] Autrusson MC, Perrein D. Dossier du patient en pédicurie-podologie. *EMC-Podol-Kinésithérapie* 2004; 1: 10-13.
- [55] Fernandes R, Armada-da-Silva P, Pool-Goudaazward A, et al. Three dimensional multi-segmental trunk kinematics and kinetics during gait: Test-retest reliability and minimal detectable change. *Gait Posture* 2016; 46: 18-25.
- [56] Hayot C, Sakka S, Lacouture P. Contribution of the six major gait determinants on the vertical center of mass trajectory and the vertical ground reaction force. *Hum Mov Sci* 2013; 32: 279-289.
- [57] Lerebourg L, L'Hermette M, Menez C, et al. The effects of shoe type on lower limb venous status during gait or exercise: A systematic review. *PloS One* 2020; 15: e0239787.
- [58] Viel E. *La marche humaine, la course et le saut: biomécanique, explorations, normes et dysfonctionnements*. Elsevier Masson, 2000.
- [59] Dyer KF, Dwyer T. *Running out of time: an examination of the improvement in running records*. Australia: New South Wales University Press, 1984.
- [60] Aubert F, Choffin T. *Athlétisme. Tome 3. Les courses*. Revue EP.S, 2007.
- [61] Capranica L, Piacentini MF, Halson S, et al. The gender gap in sport

performance: equity influences equality. *Int J Sports Physiol Perform* 2013; 8: 99-103.

[62] FFA. Les disciplines, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=526> (accessed 14 July 2021).

[63] Spencer MR, Gustin PB, Payne W. Energy system contribution during 400 to 1500 metres running. *New Stud Athl* 1996; 11: 59-66.

[64] Hill DW. Energy system contributions in middle-distance running events. *J Sports Sci* 1999; 17: 477-483.

[65] Duffield R, Dawson B, Goodman C. Energy system contribution to 400-metre and 800-metre track running. *J Sports Sci* 2005; 23: 299-307.

[66] Duffield R, Dawson B, Goodman C. Energy system contribution to 1500- and 3000-metre track running. *J Sports Sci* 2005; 23: 993-1002.

[67] Duffield R, Dawson B. Energy system contribution in track running. *New Stud Athl IAAF* 2003; 18: 47-56.

[68] Di Prampero PE, Capelli C, Pagliaro P, et al. Energetics of best performances in middle-distance running. *J Appl Physiol Bethesda Md* 1985 1993; 74: 2318-2324.

[69] FFA. Présentation de la FFA. *Fédération Française d'Athlétisme*, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=4277> (accessed 12 July 2021).

[70] Thompson MA. Physiological and biomechanical mechanisms of distance specific human running performance. *Integr Comp Biol* 2017; 57: 293-300.

[71] Howald H. Effects of sport activities on the metabolism. *Schweiz Med Wochenschr* 1974; 104: 1535-1538.

[72] International Olympic Committee. Olympic charter. In force as From 17 July 2020., http://www.olympic.org/Documents/olympic_charter_en.pdf. (2020, accessed 16 July 2021).

[73] Berthelot G, Len S, Hellard P, et al. Exponential growth combined with

exponential decline explains lifetime performance evolution in individual and human species. *Age Dordr Neth* 2012; 34: 1001-1009.

[74] El Helou N. *Evolution des performances sportives: Apport de l'épidémiologie pour l'analyse des performances sportives et des influences physiologiques, génétiques et environnementales sur les progressions humaines au cours de l'ère Olympique*. PhD thesis, Université Paris 5 Renée Descartes, 2010.

[75] Bryson L. Sport, drugs and the development of modern capitalism. *Sport Tradit* 1990; 6: 135-153.

[76] Norton K, Olds T. Morphological evolution of athletes over the 20th century: causes and consequences. *Sports Med Auckl NZ* 2001; 31: 763-783.

[77] IAAF. *Les règles des compétitions 2018-2019.*, <http://www.iaaf.org> (2017).

[78] Tamini N. *La saga des pédestriens: la grande histoire de la course à pied*. Editions Edior, 1997.

[79] World Athletics. Competition Rules. *World Athletics*, <https://www.worldathletics.org/library#rules> (accessed 18 July 2021).

[80] Slawinski J, Chiron F, Millot B, et al. Effect of a 16-days altitude training camp on 3000-m steeplechase running energetics and biomechanics: A case study. *Front Sports Act Living* 2019; 1: 63.

[81] Earl S, Hunter I, Mack GW, et al. The relationship between steeplechase hurdle economy, mechanics, and performance. *J Sport Health Sci* 2015; 4: 353-356.

[82] Hunter I, Lindsay BK, Andersen KR. Gender differences and biomechanics in the 3000M steeplechase water jump. *J Sports Sci Med* 2008; 7: 218-222.

[83] Billat V. *Course de fond et performance*. Chiron, 1991.

[84] Daniels J. *Daniels' running formula*. Amphora. Editions amphora, 2016.

[85] Conley DL, Krahenbuhl GS. Running economy and distance running performance of highly trained athletes. *Med Sci Sports Exerc* 1980; 12: 357-360.

[86] Billat VL, Demarle A, Slawinski J, et al. Physical and training characteristics of top-class marathon runners. *Med Sci Sports Exerc* 2001; 33: 2089-2097.

- [87] Burfoot A. The history of the marathon. *Sports Med* 2007; 37: 284-287.
- [88] FFA. Rapport d'activités 2012-2016, www.athle.fr/pdf/AG2016/Rapport-activite-2012-2016.pdf (2017).
- [89] FFA. Plan de développement 2013-2017 | Fédération Française d'Athlétisme, http://www.athle.fr/pdf/docffa/Plan_Developpement_2013-2017.pdf (2012).
- [90] FFA. Projet de performance Fédéral 2017-2024.
- [91] FFA. Plan de développement 2018-2024 | Fédération Française d'Athlétisme, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=5763> (2018, accessed 4 December 2019).
- [92] Piasenta J. *L'éducation athlétique*. Amphora. Paris, 1988.
- [93] Hollings SC, Hopkins WG, Hume PA. Age at Peak Performance of Successful Track & Field Athletes. *Int J Sports Sci Coach* 2014; 9: 651-661.
- [94] Schulz R, Curnow C. Peak performance and age among superathletes: track and field, swimming, baseball, tennis, and golf. *J Gerontol* 1988; 43: NaN-NaN.
- [95] Tanaka H, Seals DR. Endurance exercise performance in Masters athletes: age-associated changes and underlying physiological mechanisms. *J Physiol* 2008; 586: 55-63.
- [96] FFA. Les différentes licences. *Fédération Française d'Athlétisme*, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=577> (2021, accessed 20 July 2021).
- [97] FFA. Statistiques. *Fédération Française d'Athlétisme*, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=2195> (2021, accessed 20 July 2021).
- [98] CDH. Records de France- Chronologie des records de France seniors plein air. *Site de la comission de la documnetation et de l'histoire*, <http://cdm.athle.com/asp.net/espaces.html/html.aspx?id=15324> (accessed 23 July 2021).
- [99] FFA. Barèmes et tables de cotations. *Fédération Française d'Athlétisme*,

<https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=62> (2010, accessed 23 July 2021).

[100] IAAF. IAAF scoring tables of athletics.

[101] Hill AV. Muscular movement in man: the factors governing speed and recovery from fatigue. *Muscular Mov Man Factors Gov Speed Recovery Fatigue*; 1927.

[102] King HA, Black DG. Analysis of Olympic and world records in track and field and swimming: past, present, and future. *Med Sport Sci* 1984; 18: 212-230.

[103] Grubb HJ. Models for comparing athletic performances. *J R Stat Soc Ser Stat* 1998; 47: 509-521.

[104] FFA. Tous les records. *Fédération Française d'Athlétisme*, <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=2117> (2021, accessed 25 July 2021).

[105] Anderson O. *Running science*. Human Kinetics, 2013.

[106] Duffield R, Dawson B, Goodman C. Energy system contribution to 100-m and 200-m track running events. *J Sci Med Sport* 2004; 7: 302-313.

[107] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. *Sports Med* 1999; 27: 359-379.

[108] Di Prampero PE. The energy cost of human locomotion on land and in water. *Int J Sports Med* 1986; 7: 55-72.

[109] Ettema JH. Limits of human performance and energy-production. *Int Z Für Angew Physiol Einschließlich Arbeitsphysiologie* 1966; 22: 45-54.

[110] Monod H, Scherrer J. The work capacity of a synergic muscular group. *Ergonomics* 1965; 8: 329-338.

[111] Nevill AM, Whyte G. Are there limits to running world records? *Med Sci Sports Exerc* 2005; 37: 1785-1788.

[112] Péronnet F, Thibault G. Mathematical analysis of running performance and world running records. *J Appl Physiol Bethesda Md* 1985 1989; 67: 453-465.

- [113] Riegel PS. Athletic records and human endurance: A time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. *Am Sci* 1981; 69: 285-290.
- [114] Scherrer J, Monod H. Le travail musculaire local et la fatigue chez l'homme. *Presse Med* 1960; 68: 1717-1717.
- [115] Lippi G, Banfi G, Falavaro EJ, et al. Updates on improvement of human athletic performance: focus on world records in athletics. *Br Med Bull* 2008; 87: 7-15.
- [116] Cheuvront S, Moffatt R, DeRuisseau K. Body Composition and Gender Differences in Performance. 2002, pp. 177-200.
- [117] Péronnet F. Les records du monde de course à pied masculins et féminins: à propos d'un article de la revue 'Nature'. *Rev STAPS* 1993; 32: 47-55.
- [118] Ruby BC, Robergs RA. Gender differences in substrate utilisation during exercise. *Sports Med Auckl NZ* 1994; 17: 393-410.
- [119] Sparling PB, Cureton KJ. Biological determinants of the sex difference in 12-min run performance. *Med Sci Sports Exerc* 1983; 15: 218-223.
- [120] Ostarello J, Suess E. The gender gap : 100m to 42km. In: *ISBS-Conference Proceedings Archive*. 2001.
- [121] Berthelot G, Thibault V, Tafflet M, et al. The citius end: world records progression announces the completion of a brief ultra-physiological quest. *PLoS One* 2008; 3: e1552.
- [122] Berthelot G, Tafflet M, El Helou N, et al. Athlete atypicality on the edge of human achievement: performances stagnate after the last peak, in 1988. *PLoS ONE* 2010; 5: e8800.
- [123] Dilger A, Frick B, Tolsdorf F. Are athletes doped? Some theoretical arguments and empirical evidence. *Contemp Econ Policy* 2007; 25: 604-615.
- [124] Franke WW, Berendonk B. Hormonal doping and androgenization of

athletes: a secret program of the German Democratic Republic government. *Clin Chem* 1997; 43: 1262-1279.

[125] Brandon LJ. Physiological factors associated with middle distance running performance. *Sports Med Auckl NZ* 1995; 19: 268-277.

[126] Houmard JA, Costill DL, Mitchell JB, et al. The role of anaerobic ability in middle distance running performance. *Eur J Appl Physiol* 1991; 62: 40-43.

[127] Nikolaidis PT, Knechtle B. Force-velocity characteristics and maximal anaerobic power in male recreational marathon runners. *Res Sports Med Print* 2019; 1-12.

[128] Costill DL, Fink WJ, Flynn M, et al. Muscle fiber composition and enzyme activities in elite female distance runners. *Int J Sports Med* 1987; 8 Suppl 2: 103-106.

[129] Neptune RR, McGowan CP, Fiandt JM. The influence of muscle physiology and advanced technology on sports performance. *Annu Rev Biomed Eng* 2009; 11: 81-107.

[130] Brewer J. Running Science. In: *Running Science*. University of Chicago Press, 2021.

[131] Alvero-Cruz JR, Carnero EA, García MAG, et al. Predictive performance models in long-distance runners: a narrative review. *Int J Environ Res Public Health*; 17. Epub ahead of print 9 November 2020. DOI: 10.3390/ijerph17218289.

[132] Hill AV, Long CNH, Lupton H. Muscular exercise, lactic acid, and the supply and utilisation of oxygen. *Proc R Soc Lond Ser B Contain Pap Biol Character* 1924; 96: 438-475.

[133] Hawkins MN, Raven PB, Snell PG, et al. Maximal oxygen uptake as a parametric measure of cardiorespiratory capacity. *Med Sci Sports Exerc* 2007; 39: 103-107.

[134] Levine BD. $\dot{V}O_{2\max}$: what do we know, and what do we still need to know? *J Physiol* 2008; 586: 25-34.

[135] Billat V, Berthoin S, Blondel N, et al. La vitesse à $\dot{V}O_{2\max}$, signification et

applications en course à pied. *Staps* 2001; 54: 45-61.

[136] Foster C. V O₂ max and training indices as determinants of competitive running performance. *J Sports Sci* 1983; 1: 13-22.

[137] Foster C, Costill DL, Daniels JT, et al. Skeletal muscle enzyme activity, fiber composition and VO₂ max in relation to distance running performance. *Eur J Appl Physiol* 1978; 39: 73-80.

[138] Noakes TD. Implications of exercise testing for prediction of athletic performance: a contemporary perspective. *Med Sci Sports Exerc* 1988; 20: 319-330.

[139] Costill DL, Thomason H, Roberts E. Fractional utilization of the aerobic capacity during distance running. *Med Sci Sports* 1973; 5: 248-252.

[140] Bosquet L, Léger L, Legros P. Methods to determine aerobic endurance. *Sports Med Auckl NZ* 2002; 32: 675-700.

[141] Daniels J, Daniels N. Running economy of elite male and elite female runners. *Med Sci Sports Exerc* 1992; 24: 483-489.

[142] Foster C, Lucia A. Running economy: the forgotten factor in elite performance. *Sports Med Auckl NZ* 2007; 37: 316-319.

[143] Adelson W, Yaggie JA, Buono MJ. The vertical component of the ground reaction force and running economy. *Clin Kinesiol Online* 2005; 59: 20.

[144] Morgan DW, Martin PE, Krahenbuhl GS. Factors affecting running economy. *Sports Med Auckl NZ* 1989; 7: 310-330.

[145] Saunders PU, Pyne DB, Telford RD, et al. Factors affecting running economy in trained distance runners. *Sports Med Auckl NZ* 2004; 34: 465-485.

[146] Morgan DW, Martin PE, Baldini FD, et al. Effects of a prolonged maximal run on running economy and running mechanics. *Med Sci Sports Exerc* 1990; 22: 834-840.

[147] Svedenhag J, Sjödén B. Maximal and submaximal oxygen uptakes and blood lactate levels in elite male middle- and long-distance runners. *Int J Sports Med* 1984; 5: 255-261.

- [148] Daniels JT. A physiologist's view of running economy. *Med Sci Sports Exerc* 1985; 17: 332-338.
- [149] Pollock ML, Jackson AS, Pate RR. Discriminant analysis of physiological differences between good and elite distance runners. *Res Q Exerc Sport* 1980; 51: 521-532.
- [150] Cavanagh P. Biomechanical studies of elite distance runners: directions for future research. *Future Dir Exerc Sport Sci Res Ed K Skin Corbin P Landers P Martin C Wells Hum Kinet Champaign Ill USA* 1989; 163-179.
- [151] Margaria R, Aghemo P, Rovelli E. Indirect determination of maximal O₂ consumption in man. *J Appl Physiol* 1965; 20: 1070-1073.
- [152] Costill DL, Branam G, Eddy D, et al. Determinants of marathon running success. *Int Z Für Angew Physiol Einschließlich Arbeitsphysiologie* 1971; 29: 249-254.
- [153] Hopkins WG, Edmond IM, Hamilton BH, et al. Relation between power and endurance for treadmill running of short duration. *Ergonomics* 1989; 32: 1565-1571.
- [154] Gollnick PD, Armstrong RB, Saltin B, et al. Effect of training on enzyme activity and fiber composition of human skeletal muscle. *J Appl Physiol* 1973; 34: 107-111.
- [155] Helgerud J, Høydal K, Wang E, et al. Aerobic high-intensity intervals improve VO₂max more than moderate training. *Med Sci Sports Exerc* 2007; 39: 665-671.
- [156] Cureton KJ, Sparling PB, Evans BW, et al. Effect of experimental alterations in excess weight on aerobic capacity and distance running performance. *Med Sci Sports* 1978; 10: 194-199.
- [157] Nikolaidis PT, Rosemann T, Knechtle B. Age-predicted maximal heart rate in recreational marathon runners: a cross-sectional study on Fox's and Tanaka's equations. *Front Physiol* 2018; 9: 226.
- [158] Allen SV, Hopkins WG. Age of peak competitive performance of elite athletes: a systematic review. *Sports Med Auckl NZ* 2015; 45: 1431-1441.

- [159] Longo AF, Siffredi CR, Cardey ML, et al. Age of peak performance in Olympic sports: A comparative research among disciplines. *J Hum Sport Exerc* 2016; 11: 31-41.
- [160] Morgan DW, Bransford DR, Costill DL, et al. Variation in the aerobic demand of running among trained and untrained subjects. *Med Sci Sports Exerc* 1995; 27: 404-409.
- [161] Anderson T. Biomechanics and running economy. *Sports Med Auckl NZ* 1996; 22: 76-89.
- [162] Cunningham R, Hunter I, Seeley MK, et al. Variations in running technique between female sprinters, middle, and distance runners. *Int J Exerc Sci* 2013; 6: 6.
- [163] Elliott BC, Roberts AD. A biomechanical evaluation of the role of fatigue in middle-distance running. *Can J Appl Sport Sci J Can Sci Appl Au Sport* 1980; 5: 203-207.
- [164] Arellano CJ, Kram R. The effects of step width and arm swing on energetic cost and lateral balance during running. *J Biomech* 2011; 44: 1291-1295.
- [165] Arellano CJ, Kram R. The metabolic cost of human running: is swinging the arms worth it? *J Exp Biol* 2014; 217: 2456-2461.
- [166] Moore IS. Is there an economical running technique? A review of modifiable biomechanical factors affecting running economy. *Sports Med* 2016; 46: 793-807.
- [167] Bushnell T, Hunter I. Differences in technique between sprinters and distance runners at equal and maximal speeds. *Sports Biomech* 2007; 6: 261-268.
- [168] Enomoto Y, Kadono H, Suzuki Y, et al. Biomechanical analysis of the medalists in the 10,000 metres at the 2007 World Championships in Athletics. *New Stud Athl* 2008; 3: 61-66.
- [169] Pontzer H, Holloway JH, Holloway JH, et al. Control and function of arm swing in human walking and running. *J Exp Biol* 2009; 212: 523-534.
- [170] Williams KR, Cavanagh PR. Relationship between distance running mechanics, running economy, and performance. *J Appl Physiol* 1987; 63: 1236-

1245.

- [171] Elliot B, Ackland T. Biomechanical effects of fatigue on 10,000 meter running technique. *Res Q Exerc Sport* 1981; 52: 160-166.
- [172] Bishop M, Fiolkowski P, Conrad B, et al. Athletic footwear, leg stiffness, and running kinematics. *J Athl Train* 2006; 41: 387-392.
- [173] Ferris DP, Liang K, Farley CT. Runners adjust leg stiffness for their first step on a new running surface. *J Biomech* 1999; 32: 787-794.
- [174] Hoogkamer W, Kipp S, Frank JH, et al. A comparison of the energetic cost of running in marathon racing shoes. *Sports Med Auckl NZ* 2018; 48: 1009-1019.
- [175] Lussiana T, Hébert-Losier K, Mourot L. Effect of minimal shoes and slope on vertical and leg stiffness during running. *J Sport Health Sci* 2015; 4: 195-202.
- [176] McMahon TA, Greene PR. Fast Running Tracks. *Sci Am* 1978; 239: 148-163.
- [177] Smith G, Watanatada P. Adjustment of vertical displacement and stiffness with changes to running footwear stiffness. *Med Sci Sports Exerc* 2002; 34: S178.
- [178] Stefanyshyn DJ, Nigg BM. Energy and performance aspects in sports surfaces. *Sport Surfaces-Biomechanics Inj Perform Test Install Nigg B Cole G Stefanyshyn Eds* 2003; 31-46.
- [179] Bermon S, Garrandes F, Szabo A, et al. Effect of Advanced Shoe Technology on the Evolution of Road Race Times in Male and Female Elite Runners. *Front Sports Act Living* 2021; 3: 653173.
- [180] Fuller JT, Bellenger CR, Thewlis D, et al. The effect of footwear on running performance and running economy in distance runners. *Sports Med Auckl NZ* 2015; 45: 411-422.
- [181] Hoogkamer W, Kram R, Arellano CJ. How Biomechanical Improvements in Running Economy Could Break the 2-hour Marathon Barrier. *Sports Med Auckl NZ* 2017; 47: 1739-1750.
- [182] Lippi G, Favaloro EJ, Guidi GC. The genetic basis of human athletic performance. Why are psychological components so often overlooked? *J Physiol*

2008; 586: 3017; author reply 3019-3020.

[183] Vealey RS. Mental skills training in sport.

[184] Seiler R, Stock A. Handbuch psychotraining im sport. *Reinbek Rowohlt*.

[185] Birrer D, Morgan G. Psychological skills training as a way to enhance an athlete's performance in high-intensity sports. *Scand J Med Sci Sports* 2010; 20 Suppl 2: 78-87.

[186] Lane AM, Devonport TJ, Friesen AP, et al. How should I regulate my emotions if I want to run faster? *Eur J Sport Sci* 2016; 16: 465-472.

[187] Morgan WP. Psychogenic factors and exercise metabolism: a review. *Med Sci Sports Exerc* 1985; 17: 309-316.

[188] Arent SM, Landers DM. Arousal, anxiety, and performance: a reexamination of the Inverted-U hypothesis. *Res Q Exerc Sport* 2003; 74: 436-444.

[189] Hoos O, Boeselt T, Steiner M, et al. Long-range correlations and complex regulation of pacing in long-distance road racing. *Int J Sports Physiol Perform* 2014; 9: 544-553.

[190] St C, Noakes T. Evidence for complex system integration and dynamic neural regulation of skeletal muscle recruitment during exercise in humans. *Br J Sports Med* 2004; 38: 797-806.

[191] Aftalion A, Bonnans JF. Optimization of running strategies based on anaerobic energy and variations of velocity. *SIAM J Appl Math* 2014; 74: 1615-1636.

[192] Davies CT. Effects of wind assistance and resistance on the forward motion of a runner. *J Appl Physiol* 1980; 48: 702-709.

[193] Jones AM, Whipp BJ. Bioenergetic constraints on tactical decision making in middle distance running. *Br J Sports Med* 2002; 36: 102-104.

[194] Tucker R, Lambert MI, Noakes TD. An analysis of pacing strategies during men's world-record performances in track athletics. *Int J Sports Physiol Perform* 2006; 1: 233-245.

[195] Angus SD. Did recent world record marathon runners employ optimal pacing

strategies? *J Sports Sci* 2014; 32: 31-45.

[196] Renfree A, Casado A. Athletic races represent complex systems, and pacing behavior should be viewed as an emergent phenomenon. *Front Physiol* 2018; 9: 1432.

[197] Renfree A, St Clair Gibson A. Influence of different performance levels on pacing strategy during the Women's World Championship marathon race. *Int J Sports Physiol Perform* 2013; 8: 279-285.

[198] Billat V, Vitiello D, Palacin F, et al. Race Analysis of the World's Best Female and Male Marathon Runners. *Int J Environ Res Public Health* 2020; 17: E1177.

[199] Santos-Lozano A, Collado PS, Foster C, et al. Influence of sex and level on marathon pacing strategy. Insights from the New York City race. *Int J Sports Med* 2014; 35: 933-938.

[200] Ito S. Aerodynamic effects by marathon pacemakers on a main runner. *Trans Jpn Soc Mech Eng Ser B* 2007; 73: 1975-1980.

[201] Nikolaidis PT, Di Gangi S, Chtourou H, et al. The role of environmental conditions on marathon running performance in men competing in Boston marathon from 1897 to 2018. *Int J Environ Res Public Health*; 16. Epub ahead of print 20 2019. DOI: 10.3390/ijerph16040614.

[202] Péronnet F, Thibault G, Cousineau DL. A theoretical analysis of the effect of altitude on running performance. *J Appl Physiol Bethesda Md* 1985 1991; 70: 399-404.

[203] Deaner RO, Carter RE, Joyner MJ, et al. Men are more likely than women to slow in the marathon. *Med Sci Sports Exerc* 2015; 47: 607-616.

[204] Thiel C, Foster C, Banzer W, et al. Pacing in Olympic track races: competitive tactics versus best performance strategy. *J Sports Sci* 2012; 30: 1107-1115.

[205] Hanley B, Williams EL. Successful pacing profiles of olympic men and women 3,000 m steeplechasers. *Front Sports Act Living* 2020; 2: 21.

[206] Petrovic I, Marinković M. Influence of morphological characteristics on

running performance of endurance athletes. *Facta Univ Ser Phys Educ Sport* 2018; 16: 095-106.

[207] Denham J, Marques FZ, O'Brien BJ, et al. Exercise: putting action into our epigenome. *Sports Med Auckl NZ* 2014; 44: 189-209.

[208] Snyder EM. Complexity of genetics in the athlete phenotype: A commentary on Adrenergic- β 2 receptor polymorphism and athletic performance. *J Hum Genet* 2010; 55: 477-478.

[209] Williams AG, Folland JP. Similarity of polygenic profiles limits the potential for elite human physical performance. *J Physiol* 2008; 586: 113-121.

[210] Bouchard C. Genomic predictors of trainability. *Exp Physiol* 2012; 97: 347-352.

[211] Tucker R, Collins M. What makes champions? A review of the relative contribution of genes and training to sporting success. *Br J Sports Med* 2012; 46: 555-561.

[212] Holliss BA, Burden RJ, Jones AM, et al. Eight weeks of intermittent hypoxic training improves submaximal physiological variables in highly trained runners. *J Strength Cond Res* 2014; 28: 2195-2203.

[213] Robertson EY, Saunders PU, Pyne DB, et al. Reproducibility of performance changes to simulated live high/train low altitude. *Med Sci Sports Exerc* 2010; 42: 394-401.

[214] Stray-Gundersen J, Chapman RF, Levine BD. 'Living high-training low' altitude training improves sea level performance in male and female elite runners. *J Appl Physiol Bethesda Md* 1985 2001; 91: 1113-1120.

[215] Ebbeling CB, Clarkson PM. Exercise-induced muscle damage and adaptation. *Sports Med Auckl NZ* 1989; 7: 207-234.

[216] Mujika I, Halson S, Burke LM, et al. An integrated, multifactorial approach to periodization for optimal performance in individual and team sports. *Int J Sports Physiol Perform* 2018; 13: 538-561.

- [217] Semenza GL. Regulation of oxygen homeostasis by hypoxia-inducible factor 1. *Physiology* 2009; 24: 97-106.
- [218] Theofilidis G, Bogdanis GC, Koutedakis Y, et al. Monitoring exercise-induced muscle fatigue and adaptations: making sense of popular or emerging indices and biomarkers. *Sports*; 6. Epub ahead of print 26 November 2018. DOI: 10.3390/sports6040153.
- [219] Turner G. *Hypoxic Exposure to Optimise Altitude Training Adaptations in Elite Endurance Athletes*. University of Brighton, 2016.
- [220] Kellmann M. Preventing overtraining in athletes in high-intensity sports and stress/recovery monitoring. *Scand J Med Sci Sports* 2010; 20 Suppl 2: 95-102.
- [221] Hollings SC, Hopkins WG, Hume PA. Environmental and venue-related factors affecting the performance of elite male track athletes. *Eur J Sport Sci* 2012; 12: 201-206.
- [222] Thomas CE, Chambers TP, Main LC, et al. Factors influencing the early development of world-class caribbean track and field athletes: a qualitative investigation. *J Sports Sci Med* 2019; 18: 758-771.
- [223] Aughey RJ. Applications of GPS technologies to field sports. *Int J Sports Physiol Perform* 2011; 6: 295-310.
- [224] Baker LB, Jeukendrup AE. Optimal composition of fluid-replacement beverages. *Compr Physiol* 2014; 4: 575-620.
- [225] Duffield R, Cannon J, King M. The effects of compression garments on recovery of muscle performance following high-intensity sprint and plyometric exercise. *J Sci Med Sport* 2010; 13: 136-140.
- [226] Muniz-Pardos B, Sutehall S, Angeloudis K, et al. Recent improvements in marathon run times are likely technological, not physiological. *Sports Med Auckl NZ* 2021; 51: 371-378.
- [227] Emig T, Peltonen J. Human running performance from real-world big data. *Nat Commun* 2020; 11: 4936.

- [228] Frederick EC. Physiological and ergonomics factors in running shoe design. *Appl Ergon* 1984; 15: 281-287.
- [229] Senefeld JW, Haischer MH, Jones AM, et al. Technological advances in elite marathon performance. *J Appl Physiol Bethesda Md* 1985 2021; 130: 2002-2008.
- [230] Jakeman JR, Byrne C, Eston RG. Lower limb compression garment improves recovery from exercise-induced muscle damage in young, active females. *Eur J Appl Physiol* 2010; 109: 1137-1144.
- [231] de Glanville KM, Hamlin MJ. Positive effect of lower body compression garments on subsequent 40-kM cycling time trial performance. *J Strength Cond Res* 2012; 26: 480-486.
- [232] Berry MJ, McMurray RG. Effects of graduated compression stockings on blood lactate following an exhaustive bout of exercise. *Am J Phys Med* 1987; 66: 121-132.
- [233] Wang N, Zha A, Wang J. Study on the wicking property of polyester filament yarns. *Fibers Polym* 2008; 9: 97-100.
- [234] Bishop PA, Jones E, Woods AK. Recovery from training: a brief review: brief review. *J Strength Cond Res* 2008; 22: 1015-1024.
- [235] Hausswirth C, Mujika I. *Recovery for performance in sport*. Human Kinetics, 2013.
- [236] Tenforde AS, Sayres LC, McCurdy ML, et al. Overuse injuries in high school runners: lifetime prevalence and prevention strategies. *PM R* 2011; 3: 125-131; quiz 131.
- [237] Filipp F. Is science killing sport? Gene therapy and its possible abuse in doping. *EMBO Rep* 2007; 8: 433-435.
- [238] Baron DA, Martin DM, Magd SA. Doping in sports and its spread to at-risk populations: an international review. *World Psychiatry* 2007; 6: 118.
- [239] Birzniece V. Doping in sport: effects, harm and misconceptions. *Intern Med J* 2015; 45: 239-248.

- [240] Durussel J, Daskalaki E, Anderson M, et al. Haemoglobin mass and running time trial performance after recombinant human erythropoietin administration in trained men. *PloS One* 2013; 8: e56151.
- [241] Haile DW, Durussel J, Mekonen W, et al. Effects of EPO on Blood Parameters and Running Performance in Kenyan Athletes. *Med Sci Sports Exerc* 2019; 51: 299-307.
- [242] Cazorla G, Leger L. Utilisation des tests de terrain dans l'orientation du contenu d'entraînement dans les sports individuels et collectifs.
- [243] Margaria R, Aghemo P, Piñera Limas F. A simple relation between performance in running and maximal aerobic power. *J Appl Physiol* 1975; 38: 351-352.
- [244] Cazorla G. Exploration du métabolisme aérobie sur le terrain. In: *Colloque médico-technique de la Fédération Française d'Athlétisme*. 1985, pp. 8-13.
- [245] Pugh LG. The influence of wind resistance in running and walking and the mechanical efficiency of work against horizontal or vertical forces. *J Physiol* 1971; 213: 255-276.
- [246] Thibaut G, Mercier D. La planification de l'entraînement du coureur de fond. *Rev Qué.*
- [247] Léger L, Mercier D, Gauvin L. The relationship between % VO₂max and running performance time. *Sport Elite Perform Champaign IL Hum Kinet* 1986; 113-20.
- [248] Sordello J. *La Bible du running: Le guide scientifique et pratique pour tous*. Amphora, 2019.
- [249] Schmid W, Knechtle B, Knechtle P, et al. Predictor variables for marathon race time in recreational female runners. *Asian J Sports Med* 2012; 3: 90-98.
- [250] Tanda G. Prediction of marathon performance time on the basis of training indices. *J Hum Sport Exerc* 2011; Volume 6: pag. 521-520.
- [251] Péronnet F, Thibault G. [Physiological analysis of running performance:

revision of the hyperbolic model]. *J Physiol (Paris)* 1987; 82: 52-60.

[252] Esteve-Lanao J, Del Rosso S, Larumbe-Zabala E, et al. Predicting recreational runners' marathon performance time during their training preparation. *J Strength Cond Res*. Epub ahead of print 1 July 2019. DOI: 10.1519/JSC.0000000000003199.

[253] Arcelli E, Massini F. *La mia maratona. Tutti i consigli per correre i 42,195 km per la prima volta. E per la seconda, per la terza...* Correre. 2011.

[254] Keogh A, Smyth B, Caulfield B, et al. Prediction equations for marathon performance: A systematic review. *Int J Sports Physiol Perform* 2019; 14: 1159-1169.

[255] García-Manso JM, Martín-González JM, Vaamonde D, et al. The limitations of scaling laws in the prediction of performance in endurance events. *J Theor Biol* 2012; 300: 324-329.

[256] Berndsen J, Smyth B, Lawlor A. Pace my race: recommendations for marathon running. In: *Proceedings of the 13th ACM Conference on Recommender Systems*. 2019, pp. 246-250.

[257] Ruiz-Mayo D, Pulido E, Martinez-Munoz G. Marathon performance prediction of amateur runners based on training session data. *Mach Learn Data Min Sports Anal* 2016; 8.

[258] Coe PN, Martin DE. *Training distance runners*. Human Kinetics Publishers Inc. Champaign, IL, USA, 1991.

[259] Foster C, Daniels J. *Running by the numbers*. Run world 10: 14-17, 1975.

[260] Slovic P. Empirical study of training and performance in the marathon. *Res Q* 1977; 48: 769-777.

[261] de Smet d'Olbecke D. *Hybrid models to predict recreational runners' performance*. UCL-Université Catholique de Louvain, 2019.

[262] Billat V. *Révolution marathon*. De Boeck Supérieur, 2018.

[263] Run SMART Project. *Jack Daniels' VDOT Running Calculator* |, <https://runsmartproject.com/calculator/> (accessed 13 September 2021).

[264] Scudamore EM, Barry VW, Coons JM. An evaluation of time-trial-based

- predictions of $\dot{V}O_2\text{max}$ and recommended training paces for collegiate and recreational runners. *J Strength Cond Res* 2018; 32: 1137-1143.
- [265] Coquart J, Alberty M, Bosquet L. Validity of a nomogram to predict long distance running performance. *J Strength Cond Res* 2009; 23: 2119-2123.
- [266] Mulligan M, Adam G, Emig T. A minimal power model for human running performance. *PloS One* 2018; 13: e0206645.
- [267] Sylvan Katz J, Katz L. Power laws and athletic performance. *J Sports Sci* 1999; 17: 467-476.
- [268] Hugh Morton R. A 3-parameter critical power model. *Ergonomics* 1996; 39: 611-619.
- [269] Yang L, Shami A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* 2020; 415: 295-316.
- [270] Coquart J, Bosquet L. Precision in the prediction of middle distance-running performances using either a nomogram or the modeling of the distance-time relationship. *J Strength Cond Res* 2010; 24: 2920-2926.
- [271] Saltin B. Oxygen transport by the circulatory system during exercise in man. *Limiting Factors Phys Perform* 1973; 235-252.
- [272] Davies CTM, Thompson MW. Aerobic performance of female marathon and male ultramarathon athletes. *Eur J Appl Physiol* 1979; 41: 233-245.
- [273] Morton RH, Hodgson DJ. The relationship between power output and endurance: a brief review. *Eur J Appl Physiol* 1996; 73: 491-502.
- [274] Bosquet L, Duchene A, Lecot F, et al. $\dot{V}O_2\text{max}$ estimate from three-parameter critical velocity models: validity and impact on 800 m running performance prediction. *Eur J Appl Physiol* 2006; 97: 34-42.
- [275] Vandewalle H, Vautier JF, Kachouri M, et al. Work-exhaustion time relationships and the critical power concept. *Crit Rev J Sports Med Phys Fit* 1997; 37: 89-102.
- [276] Bishop D, Jenkins DG, Howard A. The critical power function is dependent on

the duration of the predictive exercise tests chosen. *Int J Sports Med* 1998; 19: 125-129.

[277] Vautier JF, Vandewalle H, Arabi H, et al. Critical power as an endurance index. *Appl Ergon* 1995; 26: 117-121.

[278] Vandewalle H, Zinoubi B, Driss T. Modelling of running performances: comparison of power laws and Endurance Index. In: *Communication to the 14th Annual Conference of the Society of Chinese Scholars on Exercise Physiology and Fitness (SCSEPF) Macau*. 2015, pp. 22-23.

[279] Zinoubi B, Vandewalle H, Zbidi S, et al. Estimation of running endurance by means of empirical models: A preliminary study. *Sci Sports* 2019; 34: 24-29.

[280] Poole DC, Burnley M, Vanhatalo A, et al. Critical power: an important fatigue threshold in exercise physiology. *Med Sci Sports Exerc* 2016; 48: 2320-2334.

[281] Hill DW, Vingren JL, Nakamura FY, et al. Relationship between speed and time in running. *Int J Sports Med* 2011; 32: 519-522.

[282] Bunker RP, Thabtah F. A machine learning framework for sport result prediction. *Appl Comput Inform* 2019; 15: 27-33.

[283] Cust EE, Sweeting AJ, Ball K, et al. Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance. *J Sports Sci* 2019; 37: 568-600.

[284] Wiseman O. *Using machine learning to predict the winning score of professional golf events on the PGA Tour*. National College of Ireland, 2016.

[285] McCabe A, Trevathan J. Artificial intelligence in sports prediction. In: *Fifth International Conference on Information Technology: New Generations (itng 2008)*. IEEE, 2008, pp. 1194-1197.

[286] Uddin S, Khan A, Hossain ME, et al. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak* 2019; 19: 281.

[287] Bunker R, Susnjak T. The application of machine learning techniques for

predicting results in team sport: a review. *Prepr Submitt ArXiv* 2019; 1-48.

[288] Claudino JG, de Oliveira Capanema D, de Souza TV, et al. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. *Sports Med-Open* 2019; 5: 28.

[289] Dhanabal S, Chandramathi S. A review of various k-nearest neighbor query processing techniques. *Int J Comput Appl* 2011; 31: 14-22.

[290] Maszczyk A, Gołaś A, Pietraszewski P, et al. Application of neural and regression models in sports results prediction. *Procedia - Soc Behav Sci* 2014; 117: 482-487.

[291] Velichkov B, Koychev I, Boytcheva S. Deep learning contextual models for prediction of sport event outcome from sportsman's interviews. In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*. 2019, pp. 1240-1246.

[292] Peace IC, Uzoma AO, Ita SA, et al. A comparative analysis of K-NN and ANN techniques in machine learning. *Int J Eng Res Technol* 2015; 4: 420-425.

[293] Ruiz JR, Ramirez-Lechuga J, Ortega FB, et al. Artificial neural network-based equation for estimating VO₂max from the 20 m shuttle run test in adolescents. *Artif Intell Med* 2008; 44: 233-245.

[294] Op De Beéck T, Meert W, Schütte K, et al. Fatigue prediction in outdoor runners via machine learning and sensor fusion. In: *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*. 2018, pp. 606-615.

[295] Cuk I, Nikolaidis PT, Knechtle B. Sex differences in pacing during half-marathon and marathon race. *Res Sports Med Print* 2019; 1-10.

[296] Cushman DM, Markert M, Rho M. Performance trends in large 10-km road running races in the United States. *J Strength Cond Res* 2014; 28: 892-901.

[297] Nikolaidis PT, Knechtle B. Pacing in age group marathoners in the 'New York

City Marathon'. *Res Sports Med Print* 2018; 26: 86-99.

[298] Roeh A, Bunse T, Lembeck M, et al. Running effects on cognition and plasticity (ReCaP): study protocol of a longitudinal examination of multimodal adaptations of marathon running. *Res Sports Med Print* 2019; 1-15.

[299] Vieluf S, Scheer V, Hasija T, et al. Multimodal approach towards understanding the changes in the autonomic nervous system induced by an ultramarathon. *Res Sports Med Print* 2019; 1-10.

[300] Marc A, Sedeaud A, Schipman J, et al. Geographic enrolment of the top 100 in athletics running events from 1996 to 2012. *J Sports Med Phys Fitness* 2017; 57: 418-425.

[301] Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1: 98-101.

[302] FFA. Rapport d'activités 2012-2016., www.athle.fr/pdf/AG2016/Rapport-activite-2012-2016.pdf (2017).

[303] FFA. *Règlement des compétitions nationales 2016*.

[304] FFA. Tous les résultats 2019., <http://www.athle.fr/asp.net/main.html/html.aspx?htmlid=5257>.

[305] Noble TJ, Chapman RF. Marathon specialization in elites: a head start for africans. *Int J Sports Physiol Perform* 2018; 13: 102-106.

[306] Onywera VO. East African runners: their genetics, lifestyle and athletic prowess. *Med Sport Sci* 2009; 54: 102-109.

[307] Smyth B, Muniz-Pumares D. Calculation of critical speed from raw training data in recreational marathon runners. *Med Sci Sports Exerc* 2020; 52: 2637-2645.

[308] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155-163.

[309] Munro BH. Correlation. *Stat Methods Health Care Res* 1997; 224-245.

[310] Freund Y. Concordance entre deux méthodes de mesure d'une même variable: diagramme de Bland et Altman. *Ann Fr Médecine Urgence* 2016; 6: 143-146.

- [311] Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet* 1986; 327: 307-310.
- [312] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8: 135-160.
- [313] Lerebourg L, Coquart JB. Changes in performances/characteristics of French female runners over the last 12 years. *Res Sports Med* 2020; 1-11.
- [314] Tokmakidis SP, Léger L, Mercier D, et al. New approaches to predict VO₂max and endurance from running performances. *J Sports Med Phys Fitness* 1987; 27: 401-409.
- [315] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. From empirical models to physiological models. *Sports Med Auckl NZ* 1999; 27: 359-379.
- [316] Giavarina D. Understanding bland altman analysis. *Biochem Medica* 2015; 25: 141-151.
- [317] Bunker R. The application of machine learning techniques for predicting results in team sport: a review. *Prepr Submitt ArXiv* 2019; 1-48.
- [318] Di Prampero PE. The concept of critical velocity: a brief analysis. *Eur J Appl Physiol* 1999; 80: 162-164.
- [319] Trappe SW, Costill DL, Vukovich MD, et al. Aging among elite distance runners: a 22-yr longitudinal study. *J Appl Physiol Bethesda Md* 1985 1996; 80: 285-290.
- [320] Sánchez Muñoz C, Muros JJ, López Belmonte Ó, et al. Anthropometric characteristics, body composition and somatotype of elite male young runners. *Int J Environ Res Public Health*; 17. Epub ahead of print 20 January 2020. DOI: 10.3390/ijerph17020674.
- [321] Popov T. Hurdling in the steeplechase. *Mod Athlete Coach* 1983; 21: 17-18.
- [322] Kipp S, Taboga P, Kram R. Ground reaction forces during steeplechase hurdling and waterjumps. *Sports Biomech* 2017; 16: 152-165.
- [323] Hunter I, Bushnell TD. Steeplechase barriers affect women less than men. *J*

Sports Sci Med 2006; 5: 318.

[324] Nikolaidis PT, Knechtle B. Validity of recreational marathon runners' self-reported anthropometric data. *Percept Mot Skills* 2020; 127: 1068-1078.

[325] Vitti A, Nikolaidis PT, Villiger E, et al. The "New York City Marathon": participation and performance trends of 1.2M runners during half-century. *Res Sports Med* 2019; 0: 1-17.

[326] Knechtle B, Di Gangi S, Rüst CA, et al. Men's participation and performance in the Boston marathon from 1897 to 2017. *Int J Sports Med* 2018; 39: 1018-1027.

[327] Ruiz-Mayo D, Pulido E, Martínez G. Marathon performance prediction of amateur runners based on training session data. *Proc Mach Learn Data Min Sports Anal* 2016; 8.

[328] Huang Z-Q, Chen Y-C, Wen C-Y. Real-time weather monitoring and prediction using city buses and machine learning. *Sensors*; 20. Epub ahead of print 10 September 2020. DOI: 10.3390/s20185173.

[329] Fritsch S, Guenther F, Suling M, et al. *Package 'neuralnet': training of neural networks*. 2016.

[330] Renganathan V. Overview of artificial neural network models in the biomedical domain. *Bratisl Lek Listy* 2019; 120: 536-540.

[331] Malone C, Fennell L, Folliard T, et al. Using a neural network to predict deviations in mean heart dose during the treatment of left-sided deep inspiration breath hold patients. *Phys Medica PM Int J Devoted Appl Phys Med Biol Off J Ital Assoc Biomed Phys AIFB* 2019; 65: 137-142.

[332] Smith PF, Ganesh S, Liu P. A comparison of random forest regression and multiple linear regression for prediction in neuroscience. *J Neurosci Methods* 2013; 220: 85-91.

[333] Lin L, Torbeck LD. Coefficient of accuracy and concordance correlation coefficient: new statistics for methods comparison. *PDA J Pharm Sci Technol* 1998; 52: 55-59.

- [334] Mustafa M, Taib MN, Murat Z, et al. Comparison between KNN and ANN classification in brain balancing application via spectrogram image. *J Comput Sci Comput Math* 2012; 2: 17-22.
- [335] Tamilarasi R, Porkodi DR. A study and analysis of disease prediction techniques in data mining for healthcare. *Int J Emerg Res Manag Technol* ISSN 2015; 1: 2278-9359.
- [336] Musa RM, Majeed AA, Taha Z, et al. The application of Artificial Neural Network and k-Nearest Neighbour classification models in the scouting of high-performance archers from a selected fitness and motor skill performance parameters. *Sci Sports* 2019; 34: e241-e249.
- [337] Anyama OU, Igiri CP. An application of linear regression & artificial neural network model in the NFL result prediction. *Int J Eng Res Technol* 2015; 4: 457-461.
- [338] Anyama OU, Nwachukwu EO. A hybrid prediction system for american NFL results. *Int J Comput Appl Technol Res* 2015; 4: 42-47.
- [339] Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod* 2019; 20: 41.
- [340] Ajiboye AR, Abdullah-Arshah R, Hongwu Q. Evaluating the effect of dataset size on predictive model using supervised learning technique. *Int J Softw Eng Comput Sci IJSECS* 2015; 1: 75-84.
- [341] Chamari K, Padulo J. 'Aerobic' and 'Anaerobic' terms used in exercise physiology: a critical terminology reflection. *Sports Med - Open* 2015; 1: 9.
- [342] Santos PA, Sousa CV, Barbosa LP, et al. Breaking the athletics world record in the 100 and 400 meters: an alternative method for assessment. *J Sports Med Phys Fitness* 2020; 60: 1317-1321.
- [343] Goss-Sampson M. *Statistical analysis in JASP: A guide for students*. JASP, 2019.
- [344] Emerson J, Hill B. Gender differences in competition: running performance in

1,500 meter tournaments. *East Econ J* 2014; 40: 499-517.

[345] Sparling PB, O'Donnell EM, Snow TK. The gender difference in distance running performance has plateaued: an analysis of world rankings from 1980 to 1996. *Med Sci Sports Exerc* 1998; 30: 1725-1729.

[346] Bam J, Noakes TD, Juritz J, et al. Could women outrun men in ultramarathon races? *Med Sci Sports Exerc* 1997; 29: 244-247.

[347] Whipp BJ, Ward SA. Will women soon outrun men? *Nature* 1992; 355: 25.

[348] Dennis SC, Noakes TD. Advantages of a smaller bodymass in humans when distance-running in warm, humid conditions. *Eur J Appl Physiol* 1999; 79: 280-284.

[349] Marino FE, Mbambo Z, Kortekaas E, et al. Advantages of smaller body mass during distance running in warm, humid environments. *Pflugers Arch* 2000; 441: 359-367.

[350] Tate CA, Holtz RW. Gender and fat metabolism during exercise: a review. *Can J Appl Physiol Rev Can Physiol Appl* 1998; 23: 570-582.

[351] Billat LV, Koralsztein JP. Significance of the velocity at VO₂max and time to exhaustion at this velocity. *Sports Med Auckl NZ* 1996; 22: 90-108.

ANNEXES

Annexe 1. Article 1 : *The effects of shoe type on lower limb venous status during gait or exercise: A systematic review*

PLOS ONE

RESEARCH ARTICLE

The effects of shoe type on lower limb venous status during gait or exercise: A systematic review

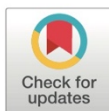
Lucie Lerebourg^{1,2*}, Maxime L'Hermette¹, Charlotte Menez^{1,2}, Jeremy Coquart¹

1 Normandie Univ, UNIROUEN, CETAPS, Rouen, France, **2** Orthodynamica Center, Mathilde Hospital 2, Rouen, France

* lucie.lerebourg@etu-univ-rouen.fr

Abstract

This systematic review evaluated the literature pertaining to the effect of shoes on lower limb venous status in asymptomatic populations during gait or exercise. The review was conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The PubMed-NCBI, EBSCO Host, Cochrane Library and Science Direct databases were searched (March 2019) for words around two concepts: shoes and venous parameters. The inclusion criteria were as follows: (1) the manuscript had to be published in an English-language peer-reviewed journal and the study had to be observational or experimental and (2) the study had to suggest the analysis of many types of shoes or orthotics on venous parameters before, during and/or after exercise. Out of 366 articles, 60 duplications were identified, 306 articles were analyzed, and 13 articles met the eligibility criteria after screening and were included. This review including approximately 211 participants. The methodological rigor of these studies was evaluated with the modified Downs and Black quality index. Nine studies investigated the effect of shoes on blood flow parameters, two on venous pressure and two on lower limb circumferences with exercise. Evidence was found that unstable shoes or shoes with similar technology, sandals, athletic or soft shoes, and customized foot orthotics elicited more improvement in venous variables than high-heeled shoes, firm shoes, ankle joint immobilization and barefoot condition. These venous changes are probably related to the efficiency of muscle pumps in the lower limbs, which in turn seem to be dependent on shoe features associated with changes in the kinetics, kinematics and muscle activity variables in lower limbs during gait and exercise.



OPEN ACCESS

Citation: Lerebourg L, L'Hermette M, Menez C, Coquart J (2020) The effects of shoe type on lower limb venous status during gait or exercise: A systematic review. *PLoS ONE* 15(11): e0239787. <https://doi.org/10.1371/journal.pone.0239787>

Editor: Yih-Kuen Jan, University of Illinois at Urbana-Champaign, UNITED STATES

Received: June 19, 2020

Accepted: September 11, 2020

Published: November 25, 2020

Copyright: © 2020 Lerebourg et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Although the hemodynamics of the arterial system, which is dominated by the pumping heart, is relatively simple, the hemodynamics of the venous system in the lower limbs, which is dominated by contracting calf muscles, is far more complicated [1–4]. Indeed, many conditions can modify the hemodynamics of the venous system, including foot dysmorphism (*e.g.*, flatfoot or *pes cavus*), ambient temperature, and even posture (*e.g.*, standing or seated conditions) [5–12].

In standing and seated positions, gravity increases the hydrostatic pressure [2], causing an accumulation of blood in the lower limbs [4,5]. Several mechanisms return the venous blood accumulated in the lower limbs to the heart, including the activation of the diaphragm pump, the plantar venous pump, and the muscle pumps of the calf and thigh [2,13]. Venous return in the foot is ensured by the plantar venous pump, which propels a quantity of venous blood at each step (in push phase) to both the deep and superficial venous system [14–18]. Therefore, the foot pump should be considered as a real impulse-aspiration system, providing the first push of venous blood in the lower limbs during walking [19,20]. One of the strategies to improve venous function might thus be to act directly on the foot with the use of a device (e.g., micro-mobile foot compression or a pneumatic compression device) or by wearing certain types of shoe (e.g., foot orthotics, sandals, heelless shoes. . .).

For example, Dohm et al. [21] showed that a micro-mobile foot compression device resulted in venous return flow velocity in the posterior tibial and popliteal veins that was greater than the resting velocity. Saggini et al. [9] showed an improvement in the venous emptying capacity of the foot-calf veno-muscular system (in *cavus* foot) with shoes equipped with a specific type of foot orthotic (i.e., visco-elastic insoles). More recently, Lopéz-Lopéz et al. [22] demonstrated an increase in venous return in the feet with shoes equipped with orthotics (i.e., customized with ethylene-vinyl acetate material). However, in addition to these technologies (i.e., foot compression devices, insoles and orthotics) to improve venous function, there are now shoe products (e.g., sandals, heelless shoes) that can improve venous flow [23–25] and might optimize recovery [26]. Indeed, athletes are now wearing "recovery" shoes (e.g., flip-flops, sandals) between series of exercise and after training session or competition.

Currently, shoes come in a wide range of types (e.g., open-toed sandals, flat shoes, heeled shoes, safety shoes, athletic shoes. . .) and each type uses a different technology (e.g., rocker soles, unstable footwear, foot-bed, balance pods and multi-density soles) to reduce discomfort or pain [27,28] or improve recovery [26]. Moreover, several reviews have indicated that shoes particularly affect gait kinematics and kinetics both acutely and chronically [29–34]. For example, some unstable shoes have been reported to improve upright walking posture, reduce lower limb joint motion and loading, and/or even modify muscle activation patterns [24,33,35,36]. Furthermore, some authors have observed or suggested that unstable shoes have an effect on blood flow or the venous system [24,25,37,38]. However, the quantitative evidence on this last observation is limited and the positive effects of shoes on improving venous return remain unknown.

To date, no study has systematically reviewed the effects of a range of shoe types (including foot orthotics, insoles, cast shoes. . .) on lower limb venous status before, during or after exercise. Therefore, the aim of this systematic review was to identify, appraise and summarize the available evidence on the effects of many types of shoes on lower limb venous status in asymptomatic populations.

Materials and methods

This systematic review was executed in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (S1 Checklist and S1 Fig) [39]. Moreover, the review protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO registration number: CRD42019121613).

Identification

A systematic search of the research literature was conducted to look for studies that assessed the effects of all types of shoes on lower limb venous status during exercise. No restriction in

terms of period have been retained. Therefore, this search was undertaken in the PubMed-NCBI (1957–2019), EBSCO Host (1981–2019), Cochrane Library (1978–2019) and ScienceDirect (1973–2019) databases. The search was conducted independently by three of the authors (LL, CM, JC) in March 2019 from keywords identified by all co-authors.

These keywords were based on two concepts: shoes and venous parameters (Table 1). All MeSH (Medical Subject Headings) terms associated with each concept (*i.e.*, synonyms) in the National Library of Medicine's (NLM) controlled vocabulary thesaurus were collected. Moreover, the authors completed this list of terms/synonyms with keywords that are frequently used in the literature to represent both concepts. Then, all identified terms/synonyms in MeSH and the keywords were combined with Boolean operators ("AND" and "OR"), and the search was conducted in the PubMed-NCBI, EBSCO Host and Cochrane Library databases (Table 1). However, for the ScienceDirect database, only the MeSH heading (not the terms/synonyms) and keywords were used because this database is unable to support more than eight Boolean operators per field (Table 1). The search string used the Boolean operator "AND" to combine the two concepts and "OR" to provide a comprehensive set of terms for each concept (Table 1). The search was conducted for titles, abstracts and keywords of citations in each database. This revealed 366 manuscripts. Once duplicate citations were removed ($n = 60$), 306 articles were analyzed (Fig 1).

Screening

Following the initial selection of studies, two experts in the field (LL and CM) assessed the eligibility of each manuscript independently in a blinded standardized manner by screening the titles and abstracts. Disagreements were discussed between three authors (CM, LL and JC) and resolved by consensus. To be considered eligible, the manuscript had to be published in a peer-reviewed journal and the study had to be observational or experimental (*e.g.*, observational, randomized controlled trials . . .). Letters to the editor, short communications, meeting abstracts, editorials, theses, case reports, reviews and meta-analyses were thus excluded, as were animal studies and studies not published in English. The study also had to suggest the analysis of the effects of several types of shoes or orthotics on venous parameters before, during and/or after physical exercise (*i.e.*, inclusion criteria). All outcomes relating to shoe fit or

Table 1. Search strategy.

Databases	Concept 1	Concept 2
PubMed-NCBI, EBSCO Host and Cochrane Library	<i>(footwear OR shoe OR orthoses, foot OR foot orthosis OR orthosis, foot OR foot orthotic devices OR device, foot orthotic OR devices, foot orthotic OR foot orthotic device OR orthotic device, foot OR orthotic devices, foot OR foot arch supports OR arch support, foot OR arch supports, foot OR foot arch support OR support, foot arch OR supports, foot arch OR orthotic shoe inserts OR insert, orthotic shoe OR inserts, orthotic shoe OR orthotic shoe insert OR shoe insert, orthotic OR shoe inserts, orthotic OR orthotic insoles OR insole, orthotic OR insoles, orthotic OR orthotic insole OR orthopedic shoe OR therapeutic shoe)</i>	<i>(venous OR blood flow OR vein)</i>
Science Direct	<i>(footwear OR shoe OR foot orthoses OR orthopedic shoe OR therapeutic shoe)</i>	<i>(venous OR blood flow OR vein)</i>

Legend. **Highlighted words:** MeSH (Medical Subject Headings) terms; **words in bold:** terms/synonymous associated to MeSH terms for "shoe" and "foot orthoses"; **words in italics:** keywords identified by authors and frequently used in the literature.

<https://doi.org/10.1371/journal.pone.0239787.t001>

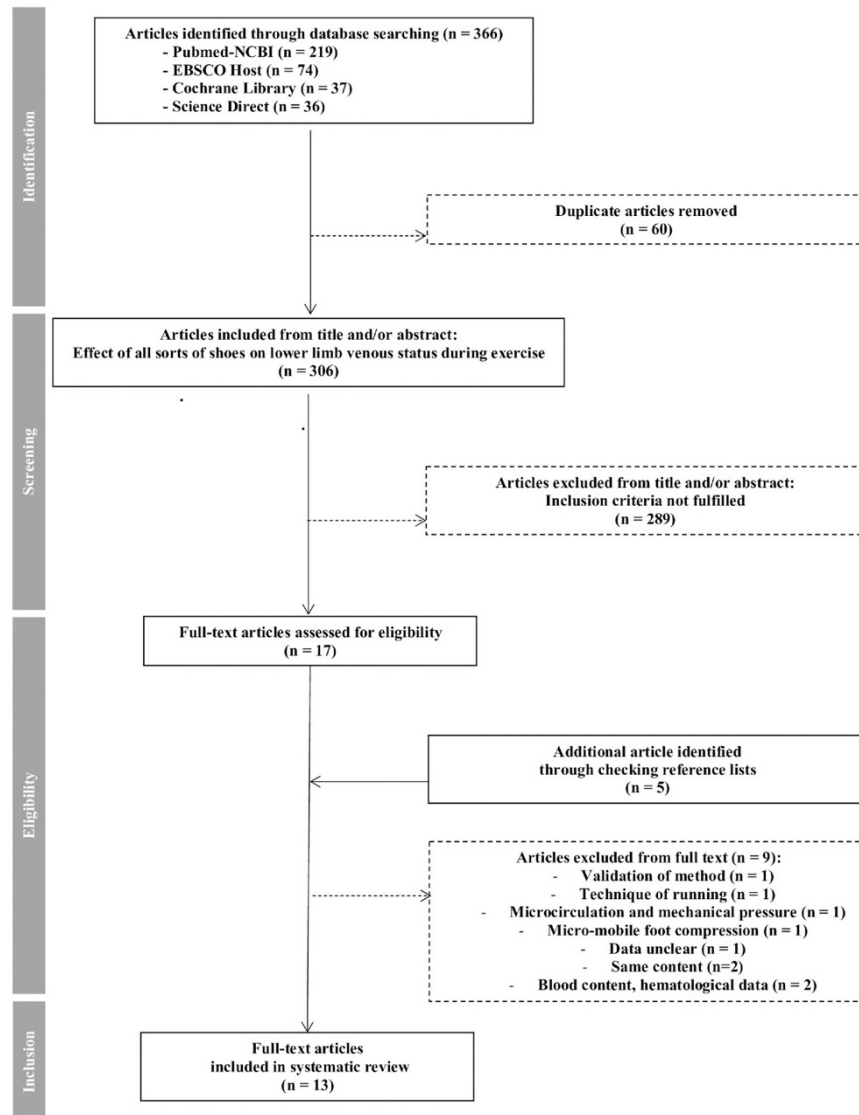


Fig 1. Flow diagram of study selection. Figure presents the flow of information through the different phases (identification, screening, eligibility and inclusion) of a systematic review.

<https://doi.org/10.1371/journal.pone.0239787.g001>

characteristics (e.g., footwear, foot orthotics, insoles, high-heeled shoes . . .) and hemodynamic venous parameters (e.g., blood or venous flow, venous pressure, venous reflux, blood rush . . .) were authorized. Studies were, however, excluded if there were no clear performance measures or if the participants were undergoing medical treatment (excluding orthopedic treatments) or had rheumatic diseases, lacked partial independence in daily activities, or had severe osteoarticular or neurological disease that would compromise posture and walking (*i.e.*, exclusion criteria).

Following the review of the titles and abstracts, 289 articles were further removed from the analysis (Fig 1).

Eligibility

Seventeen full-text articles were initially eligible. The reference lists from these manuscripts were searched manually to identify other possibly eligible articles (CM and LL), and five additional manuscripts were identified. These manuscripts were then screened, according to the procedure described above. None of these new articles was excluded on the basis of the titles and/or abstracts. Therefore, 22 full-text articles were evaluated. Based on the information in these articles, two authors (LL and CM) used a standardized form to select the manuscripts eligible for inclusion in the systematic review. This form was used to determine whether the information provided in the title and abstract matched the text content and met the inclusion criteria (*i.e.*, shoes or insoles, analysis of venous hemodynamic parameters before, during and/or after exercise . . .). Ultimately, the review included 13 studies with approximately 211 participants.

Inclusion

From the 13 included articles, the authors (LL and JC) extracted the following information: study identity (*i.e.*, first author and year of publication), participants (e.g., number, status, sex, age), study design (e.g., randomized, non-randomized), type of shoe or insole intervention (e.g., protective, athletic, high-heeled, custom orthotics . . .), type of exercise or protocol (e.g., walking, prolonged standing position, treadmill running . . .), duration of adaptation (e.g., several minutes, days . . .), venous hemodynamic variables analyzed (e.g., blood flow, venous pressure, circumferences), methods or materials (e.g., plethysmography, Gulick tape measure, Doppler ultrasound), and data to assess venous activity.

Risk of bias and quality appraisal

The methodological rigor of the studies was evaluated with the modified Downs and Black quality index (QI) [40]. The QI is a 27-point checklist [40–42] to assess the methodological quality of both randomized controlled trials and non-randomized studies [42]. This tool has the following four subscales: reporting (items 1–10), external validity (items 11–13), internal validity (items 14–26), and power (item 27). The wording of item 13 was modified from “were the staff, places, and facilities where the patients were treated, representative of the treatment of majority of patients receive?” to “was the intervention representative of those that are used in clinical practice/commercially available?” [33]. The maximum score was 28 because item 5 had a possible score of 2 if scored “yes”, 1 if scored “partially” and 0 if scored “no” or “unable to determine.” For all other items, “yes” was scored 1 and “no” and “unable to determine” were scored 0. Based on the risk of bias assessment scores, studies were classified as very high quality/very low risk of bias (26–28), good quality/low risk of bias (20–25), fair quality/fair risk of bias (15–19) or poor quality/high risk of bias (≤ 14) [43]. Higher scores indicated less risk of bias than lower scores. The QI has been shown to have high internal consistency (Kuder-

Richardson Formula 20 = 0.89), test-retest reliability ($r = 0.88$), and inter-rater reliability ($r = 0.75$) [40].

Two authors (LL and CM) independently appraised all studies in each of the domains of this tool. Disagreements during the quality assessments were discussed until consensus was reached. If consensus was not reached, a third reviewer (JC) provided the final judgment.

Results

The search strategy revealed that 13 articles met the inclusion criteria (Fig 1).

Two studies compared unstable shoes with regular shoes (*i.e.*, conventional, flat-bottomed shoes) or barefoot condition [24,37]. Four studies compared athletic shoes (*e.g.*, soft shoes) with firm shoes (*i.e.*, clogs), heelless shoes or barefoot condition [25,44–46]. Two studies compared high-heeled shoes (*i.e.*, stilettos) with different types of high heel (*i.e.*, medium heels, platform high heels) or barefoot [47,48]. Two studies analyzed the influence of protective shoes (*e.g.*, all-rubber firefighting shoes) during exercise [49,50]. One study compared sandals with conventional shoes and barefoot [23]. Another study analyzed the effect of ankle joint immobilization with different shoes (*i.e.*, cast, pneumatic walking boot) or positions of immobilization (*i.e.*, neutral, equinus) [51]. Only one study collected data on insoles (*i.e.*, customized foot orthotics) vs without insoles [22]. Five studies used consistent treadmill exercise (walking or running) [23,25,47,49,50], one used running exercise [46], five used an upright standing exercise (*e.g.*, simulated work) [22,24,37,44,45], two used tip-toe movements [47,48] and one asked participants to simulate gait [51].

The participant characteristics, study design, interventions (shoe conditions), gait type (type of exercise), duration of adaptation and venous variables analyzed in each study are presented in Table 2. The main statistically significant changes from wearing shoes or insoles in the studies are also presented in Table 2. The intervention duration, whether with shoes or insoles, varied between several minutes to 8 weeks. The participants were between 20 to 50 years old, and 54.3% were female ($n = 119$) and 45.7% were male ($n = 100$). All 13 studies measured hemodynamic venous data and at least one of the venous variables: blood flow, venous pressure or circumference (Table 2). As Table shows, the results of eight studies indicated that intervention with certain shoes resulted in positive blood flow effects (*i.e.*, increase in venous flow or decrease in edema formation) [22,24,25,44,46], positive pressure effects [23,47], or positive circumference effects (*i.e.*, decrease in circumferences) [37]. One study emphasized the harmful effects (increase in residual volume in legs) of some shoes (high-heeled shoes) on venous flow [48]. Three other studies reported that the blood flow changes [49–51] appeared to be dependent on both the shoe type and the exercise conditions (*e.g.*, humidity, temperature, immobilization, weight-bearing). Finally, the studies of Karimi et al. [37] and Lin et al. [45] reported that some shoe types did not have significant clinical effects on venous data, such as circumferences (Table 2).

A summary of the QI scores [40] for each paper is presented in Table 3. The total score of the QI ranged from 16/28 to 23/28, with a mean score of 19.8/28 and a median score of 20/28 for the 13 articles. Five of the studies were of fair quality [23,48–51] and eight were of good quality [22,24,25,37,44–47]. In five studies, the order of the interventions was randomized [23–25,37,45].

Discussion

The aim of this systematic review was to investigate the effects of various shoes and insoles on lower limb venous status, including blood flow, venous pressure and circumferences, in asymptomatic populations before, during or after exercise. The findings suggest that certain

Table 2. Characteristics and primary outcomes of studies.

Author (year)	Participant characteristics	Study design	Intervention/Shoe conditions	Type of Exercise/Gait type	Duration of adaptation	Venous variables analyzed/ Outcome measure(s)	Significant results/Main findings
Knight and Lewis (1977) [23]	9W and 3M Age (range): 29 (20–38) years Individual sample data: • Those who had never worn sandals • Those who habitually wore sandals	Randomized sequences	Barefoot/ Normal shoes/ Sandals	Treadmill walk	14 days	Venous pressures (mmHg) by pressure transducers	Initial: No significant difference in the exercise venous pressure in non-wearers and a significantly lower venous pressure in accustomed wearers when walking with sandals (pressure similar to that of barefoot walking). After 14 days: A significantly lower venous pressure while walking in sandals compared to shoes in both groups. No significant difference in the exercise pressure when walking barefoot compared to that in sandals.
Hansen et al. (1998) [44]	8 W Age (range): 24 (21–29) years Whole sample data	A balanced block design	Clogs (without heel support as the hard shoe)/Sport shoes (soft shoes)	Simulated work (standing and standing/ walking)	2 hours/work	Blood flow by hydro-plethysmography (FV and VV)	Standing work: Significant reduction in edema formation in soft shoes condition (from 3.2% to 2.8% on average). Standing/walking work: Significant decrease in the edema formation (from 2.2 to 1.2%) for soft shoes condition.
Yamamoto et al. (2000) [25]	6 M Age (SD): 19.5 (\pm 4.0) years Whole sample data	Randomized	Regular shoes (regular athletic shoes)/ Heelless shoes	Treadmill walk (at a speed of 60, 80, 100 and 129 m.min ⁻¹)	10 min/test	Blood flow (CBF) by venous occlusion plethysmography	Blood flow: A significantly higher CBF at the speed of 80 m.min ⁻¹ for heelless shoes than regular shoes. No significant difference in CBF in regular shoes or heelless shoes at walking speeds of 60, 100 and 120 m.min ⁻¹ .
Potério-Filho et al. (2006) [47]	10 W Age: 20–50 years Individual sample data: • Genu Valgum and flat feet (n = 5) • No orthopedic problems (n = 5)	N/A	Barefoot/ High-heeled shoes (stiletto)	Tip-toe movements Treadmill walk	1 or 10 extension/ flexions 1 min/ treadmill walk test	Venous pressures (mmHg) by APG (variations of the pressure, hydrostatic leg vein pressures)	No statistical difference between the variation in the cuff pressure values when persons performed 1 or 10 tip-toe foot movements. Statistical differences for the variation of the cuff pressure during static movements (41.05 mmHg), walking barefoot (26.65 mmHg) and walking with high-heeled shoes (52.2 mmHg). Significantly higher variations in the cuff pressure readings were obtained at the leg when compared with walking barefooted (during walking with high-heeled shoes). Significant difference for the final hydrostatic pressure at the end of walking between the 2 groups. Significant difference in the final hydrostatic pressure, for all individuals, when walking with high-heeled shoes (53.5 mmHg) and barefoot (61.5 mmHg).
Irzmańska et al. (2011) [50]	10 M (firemen) Age: 20–45 years Whole sample data	N/A	Protective footwear	Treadmill walk	Several minutes	Blood flow by APG (IR, CT)	Statistically significant changes for the APG parameters and temperature-humidity. Correlation between the level of humidity and temperature and blood flow parameters in the legs (IR, CT) during exercise.

(Continued)

Table 2. (Continued)

Author (year)	Participant characteristics	Study design	Intervention/Shoe conditions	Type of Exercise/Gait type	Duration of adaptation	Venous variables analyzed/ Outcome measure(s)	Significant results/Main findings
Tedeschi Filho et al. (2012) [48]	30 W Age (range): 26,4 (20–35) years Whole sample data	N/A	Barefoot/ Medium heels (3.5cm)/ Stiletto high heels (7cm)/Platform high heels (7cm)	Tip-toe movements	10 orthostatic flexion and extension/test	Blood flow by APG (EF, RVF, VFI)	EF: Higher values for the barefoot group compared with the other 3 groups. A significant decrease in both of the 7-cm high-heeled groups compared with the barefoot. RVF: Significant increase in the groups wearing high heels (stiletto and platform) and medium heels (3.5 cm) compared with barefoot group. Despite the lack of statistical significance, the medium-heel group showed lower values compared with the 2 high-heeled groups. No difference between the 2 7.0-cm high-heeled groups for this parameter. VFI: Similar values in the 3 situations evaluated.
Sousa et al. (2012) [24]	30 W Age: 20–50 years. Individual sample data: • Experimental group with unstable shoe (n = 14), 34.6 (± 7.7) years • Control group with conventional shoe, (n = 16), 34.9 (± 8.0) years	Randomized trials	Barefoot/USW/CSW	Upright standing	8 weeks (Experimental group only: wear USW 8h/day for 8weeks) 30s/ data acquisition of trial	Blood flow by duplex ultrasonography (CFV, PV)	No significant difference in the mean flow rate at CVF and PV between the experimental and control groups for both the first and second evaluations. A comparison between measurements taken with and without the unstable shoe shows a higher level of venous return with the unstable shoe in both groups during the first and second evaluation in PV and CFV. Venous returns increased in subjects wearing USW before and after 8 weeks. No differences between the first and second evaluation, both with and without the unstable shoe, as for the influence of 8 weeks of unstable shoe wearing. The same was verified in subjects who wore conventional footwear.
Lin et al. (2012) [45]	10 M Age (SD): 30.2 (± 2.9) years Whole sample data	Randomized	Barefoot/ Sport shoes	Upright standing	4 hours (1 hour/ experimental session)	Circumferences (thigh and shank) by Gulick Tape	No significant difference in thigh and shank circumferences between the 2 shoe conditions, but the difference in shank circumference was highest when a subject stood for a prolonged period on the hard floor while wearing sport shoes.
Irzmańska et al. (2014) [49]	30 M (firefighters), Age (range): 30.7 (25–35) years Work experience (5–10 years) Whole sample data	N/A	All-rubber firefighter footwear (with protective elements)	Treadmill walk	1 hour 45 minutes	Blood flow by APG (IR, SR, CW, CT, ABF)	Statistically significant parameters of blood flow in the lower limbs were the IR, SR, CW and ABF during the 3 phases (warm-up, exercise and rest) with a rise in values of the IR, SR, CW and ABF during exercise as compared to the initial rest phase and a decrease in the relaxation phase. No significant difference for the CT, remained at similar levels.
Craik et al. (2015) [51]	6 M and 4 W Age (range): 38 (27–61) years Whole sample data	N/A	Without ankle joint immobilization/ ankle immobilized in a neutral cast/ ankle immobilized in a pneumatic walking boot (in both neutral and equinus immobilization)	Simulation of the gait cycle	30 repetitions/ minute	Blood flow (VBF velocity) by duplex ultrasonography (Time averaged peak velocity, time averaged mean velocity, total volume flow)	No significant reduction in VBF measurements between full weight-bearing without ankle joint immobilization and full weight-bearing in a neutral cast or neutral pneumatic walking boot. Reduction of VBF when partial weight-bearing (50%) and when full weight-bearing in a pneumatic walking boot in equinus.

(Continued)

Table 2. (Continued)

Author (year)	Participant characteristics	Study design	Intervention/Shoe conditions	Type of Exercise/Gait type	Duration of adaptation	Venous variables analyzed/ Outcome measure(s)	Significant results/Main findings
Wezenbeek et al. (2017) [46]	25 (experienced runners), 15 M and 10 W, Age (SD): 34.5 (\pm 10.2) years Whole sample data	Non-randomized order (running activity)	Barefoot/ Shod (Pearl Izumi shoes, offset of 7 mm)	Run	30 min (10 min/ phase: rest, barefoot run, shod run)	Blood flow by oxygen-to-see device	A significant increase in blood flow after barefoot and shod running compared to initial tendon blood flow (42.6% after barefoot running and 61.7% after shod running).
Karimi et al. (2017) [37]	10 M Age (SD): 25.3 (\pm 1.49) years Whole sample data	Randomized	Barefoot/Flat-bottomed shoe (stable)/USW	Upright standing	2 hours/shoe condition (session). One-week interval between each test session	Circumferences (lower legs) by Gulick tape	An increase in the mean scores of lower leg volume, from pretest to posttest following 2 hr of standing for both legs, under all 3 footwear conditions. Results revealed that footwear conditions influenced this value for right and left legs after 2 hr of standing. Results showed that percentage of volume change was reduced by USW in relation to barefoot condition for right and left legs. No significant difference for percentage of volume change between USW and flat-bottomed shoes, or between barefoot condition and flat-bottomed shoes in both legs.
López-López et al. (2018) [22]	10M and 10 W Age (SD): 20 (\pm 1.62) years Whole sample data	Non-randomized	CFO/ Without CFO	Standing in a relaxed posture	20 days 20 min /session	Blood flow by APG (VFT and EF).	A significant difference between VFT and the EF with CFO utilization versus without use of CFO.

Subjects: number of subjects per sex (W, women; M, men), characteristics; mean age.

Shoe characteristics: CFO, customized foot orthoses; CSW, conventional shoe wearing; USW, unstable shoe wearing.

APG, Plethysmographic.

Venous parameters (ABF, alternative blood flow; CBF, calf blood flow; CFV, venous velocity of the common femoral; CT, crest time; CW, crest width; EF, ejection fraction; Hct, hematocrit; IR, impedance ratio; MCH, mean cellular hemoglobin; MCV, mean cellular volume; FV, foot volume; PV, popliteal veins; RBC, red blood cell; RVE, residual volume fraction; SR, slope ratio; TIBC, total iron-binding capacity; VBF, venous blood flow; VF, venous flow; VFT, venous filling time; VV, vascular volume; WBC, white blood cell).

<https://doi.org/10.1371/journal.pone.0239787.t002>

types of shoes can influence venous parameters. However, the studies were quite heterogeneous concerning the shoe types, the duration or type of intervention activity, and the venous parameters studied.

Effects of shoes on blood flow parameters

Five of the studies focused on unstable shoes [24], heelless shoes [25], athletic shoes [44,46] or insoles [22] and reported an increase in blood flow and venous return [22,24,25,46] or a decrease in edema formation [44] in the lower limbs. It can be assumed that the respective features (*i.e.*, instability, rocker bottom, soft, customized orthotics) modify the motion or muscle activity in the lower limbs and thus change features of the venous parameters.

Some of the evidence showed that unstable shoes [24] like heelless shoes [25] were associated with kinetic and kinematic changes in lower limb muscle function, as opposed to conventional shoes [30,33]. Indeed, the concept underlying the function of unstable shoes is that instability induces, for example, an increase in the ankle joint's range of motion during walking [28,52] and an increase in the activity of the *gastrocnemius* [33,36] and *tibialis anterior* [28,36]. Athletic shoes (*i.e.*, soft shoes) also modify movement patterns, like step frequency (*e.g.*, increased step/stride length and decreased cadence vs barefoot), joint movements (*e.g.*,

Table 3. Methodological quality of the included studies assessed by the quality index [40].

Author (year)	Reporting score (score/11)	External validity (score/3)	Bias (score/7)	Confounding (score/6)	Power (score/1)	Global (score/28)
Knight and Lewis (1977) [23]	8	2	3	3	0	16
Hansen et al. (1998) [44]	11	2	5	2	0	20
Yamamoto et al. (2000) [25]	10	2	5	4	0	21
Potério-Filho et al. (2006) [47]	10	3	4	3	0	20
Irzmańska et al. (2011) [50]	8	2	4	3	0	17
Tedeschi Filho et al. (2012) [48]	10	2	4	3	0	19
Sousa et al. (2012) [24]	11	3	5	3	0	22
Lin et al. (2012) [45]	11	2	4	3	0	20
Irzmańska et al. (2014) [49]	10	2	4	3	0	19
Craik et al. (2015) [51]	10	3	4	2	0	19
Wezenbeek et al. (2017) [46]	11	2	5	2	0	20
Karimi et al. (2017) [37]	11	2	5	4	0	22
López-López et al. (2018) [22]	10	3	5	4	1	23

Legend for the global score: excellent (26–28); good (20–25); fair (15–19); and poor (< 14).

<https://doi.org/10.1371/journal.pone.0239787.t003>

decrease in plantar-flexion vs barefoot), and muscle activity in the lower limbs [31,32]. These results suggest that blood flow variations are more associated with muscle pump activation. Moreover, it has been demonstrated that athletic shoes vs barefoot [46] can also substantially affect subtalar joint movement. Extreme shoe modifications may affect eversion [53] (*i.e.*, the association of three movements including abduction, pronation and dorsal flexion) [54], and thus the pronation (an aspect of this foot eversion) [55]. Clement et al. [56] first suggested that foot pronation causes partial constriction of the Achilles tendon's vascular network (whipping of torsional action), and Karzis et al. [57] later showed an increase in vascular resistance and reduced Achilles tendon blood flow with foot overpronation. However, it seems that pronation is mainly due to the foot movement itself [57]. Indeed, the study of Stacoff et al. [53] reported no significant difference in eversion between wearing shoes and going barefoot. This result highlights the observation that shoes do not reflect the movement of underlying bone but only of the shoe and/or skin movement. Thus, it is not the shoe that increases the vascular resistance but rather the pronation movement itself. Eversion may therefore be influenced by the shoe characteristics and/or the exercise conditions and thereby exert an influence on the venous flow [46]. Furthermore, foot orthotics are an effective intervention to reduce rear-foot eversion [58] or the pronated foot-type [59]. They also affect kinetic (*e.g.*, reduction in the peak and mean ankle eversion moments) and kinematic (*e.g.*, reduction in peak rearfoot eversion or tibial internal rotation) variables and muscle activity (*e.g.*, increase in *tibialis anterior* and *peroneus longus* electromyographical amplitudes: EMG) [32,59,60]. There was also evidence reported of the relationship of foot and ankle movements to venous return in the lower limb, where combined active movement produced the highest flow velocities [15,61,62]. This would explain the influence of orthotics use on the increased venous return or the decreased edema formation.

Regional blood distribution is dependent on the muscle fiber type [63–66]. The kinetics of venous return are dependent on contraction frequency, with variations in relaxation periods between the muscle contractions [16,64,67,68]. Walkeling et al. [69] suggested that muscle fiber-type recruitment patterns can also be altered by the choice of shoes and the shoe's mid-sole material, and the study of Rao et al. [70] showed lower amplitudes of the *lateral*

gastrocnemius and *medial gastrocnemius* fiber lengths in barefoot conditions compared to shod conditions. Thus, the rhythmic contractions of shank or foot muscles [16,64] coupled possibly with an increased range of joint motion [52] while walking with particular shoes like unstable shoes [24,25], athletic shoes [44,46], and orthotics [22] could enhance venous emptying of the foot compared to conventional shoes (*i.e.*, without this particular feature) or barefoot condition. It is important to note that depending on the type, intensity, duration and order of the exercise protocol, venous return might be affected [64]. For example, in the experiment of Yamamoto et al. [25], calf blood flow was reduced at higher walking speed (120 m.min⁻¹). In the study of Wezenbeek et al. [46], the influence of running on blood flow was higher after shod running (running shoes) than after barefoot running, but the running was performed in non-randomized order.

Two studies that focused on high-heeled shoes [48] and ankle joint immobilization [51] reported increased residual volume and a reduced ejection fraction [48] or decreased blood flow in the lower limbs [51]. It seems that such negative results for venous parameters also depend on motion, gait or muscle activity in the lower limbs caused by the features of the shoes (*e.g.*, high-heel height, as for platform or stiletto shoes, ankle immobilization, cast).

High-heeled shoes were associated with several changes in gait with, for example, a decrease in stride length [71–74], an increase in plantar flexion [72,73,75] or an increase in antagonist muscle activation (*e.g.*, *medial gastrocnemius*, *rectus femoris*) [34,75]. Yet it has been suggested that this increase in muscle activity [34,75] is not systematically linked to better venous return, because muscle activity is only related to systole and not diastole during calf muscle pumping with high heels. Indeed, the studies of Gardner and Fox [61] and Corley et al. [15] showed an increase in the efficiency of venous drainage in the lower limbs when the ankle kinematics under calf muscle activity were acting in unison with the foot pump. However, it is important to note that other characteristics of high-heeled shoes (*e.g.*, platform or stiletto) are important, as Tedeschi Filho et al. [48] showed a higher residual volume for platform shoes compared to stilettos despite the same heel height (7 cm). Thus, the variation in venous return/hemodynamic changes could be linked to kinematic differences [72] like restricted plantar and ankle joint movements with platforms compared to stilettos.

Regarding ankle joint immobilization, Craik et al. [51] showed a decrease in blood flow in the lower limbs under partial weight-bearing conditions (50%) and full weight-bearing in a pneumatic walking boot in equinus. These results can be explained, as previously, by a decrease in the efficiency of muscle pumps (calf and foot). Ankle kinematics work in synergy with the foot pump and each foot compression with weight-bearing permits an effective propulsion of blood [15]. It appears that ankle joint immobilization and a pneumatic walking boot in equinus are not as effective as the natural foot pump in producing venous flow, possibly due to their inability to stress the foot vein plexuses [76], and thus neither provides an active muscle pump [77]. Moreover, when the ankle is immobilized, the movements are clearly limited and flow velocities may thus be decreased. Indeed, Sochart and Hardinge [62] showed that combined active movements produced the highest flow velocities compared to passive movements and are thus required to promote significant increases in venous blood flow.

Last, studies of protective footwear (firefighter footwear with a wool liner) were included [49,50]. In both, the footwear was a model designed for firemen and used in firefighting and emergency actions. The results of these studies showed a correlation between the humidity and temperature and the blood flow parameters in the legs during exercise [49,50]. One of the studies reported an increase in the blood flow in the lower limb vessels due to temperature and humidity inside the footwear and the physical exertion [49]. Indeed, the human foot sweats in response to significant exercise and thermal stimulation [78]. It is well known that the typical vascular reactions in the human body depend on body temperature [64,79], and in response to

the thermal effect of protective footwear during prolonged exercise, the body reacts by the dilation of skin vessels and increased skin blood flows. These protective shoes nevertheless can generate excessive heat, which may result in an excessive load on the vascular system and suggests the deterioration of blood flow in the lower limbs.

Effects of shoes on venous pressure

Two studies that focused on sandals [23] and high-heeled shoes [47] reported a decrease in venous pressure [23,47] and an increase in the variation of venous pressure [47]. The conclusions drawn regarding this variable were similar to those for blood flow. Based on the features (*e.g.*, soft or hard sole, midsole, density, rubber sole, rocker . . .) and the shoe type, there is some evidence that the range of motion, joint movement, and the muscle or muscle pump activity are affected [30–32,34,80–84] by sandals [55,83,85,86] vs conventional shoes [86–88] and by high-heeled shoes [73,89] vs barefoot [34,72,75,90]. Several studies have shown an increase in the activity of *soleus*, *peroneus longus* [84] or *tibialis anterior* during exercise with sandals and flip-flops [83,84,91]. Furthermore, Carl and Barrett [92] demonstrated higher peak plantar pressures with flip-flops than athletic shoes but lower pressures than barefoot. These modifications in muscle activity or plantar pressure could explain the increase in muscle pump activity [16] and consequently the significantly greater fall in venous pressure (*i.e.*, significant difference) with sandals exercise, approximating barefoot walking, than with shoes. The lack of difference (*i.e.*, no significant difference) in venous pressure between sandals and barefoot conditions (*i.e.*, pressure similar) may be explained by broadly similar patterns of foot motion during walking with no significant changes in kinetic or kinematic variables [55,87,88].

We note that one of the major factors that influence muscle activity is walking speed [93]. Indeed, beyond a certain walking speed (*i.e.*, 100m.min⁻¹), Kayaga et al. [94] reported the decrease in calf blood flow when walking. Moreover, gait velocity increases antagonist activation [95]. However, as noted by Blackburn et al. [84], an increase in muscle activity and most likely the level of agonist-antagonist co-contraction may not be positive outcomes. For example, an increase in ankle invertor strength without a concomitant increase in evertor strength could predispose the ankle joint to inversion after invertor-evertor coactivation. This might be the origin of the lower efficiency of muscle venous pumps in the lower limbs.

The results of Potério-Filho et al. [47] (*i.e.*, improved function of the leg muscle pump, which results in lower leg venous pressure during walking) also seem contrary to the general assumption that walking with high-heeled shoes is detrimental to the efficiency of the leg muscle pump. Tedeschi Filho et al. [48] indicated that high heels have a deleterious action on venous return and thus tend to cause venous hypertension in the lower limbs, whereas Potério-Filho et al. [47] reported lower leg venous pressure and greater variation in pressure during walking in high heels compared to barefoot. It has been established that walking in high-heeled shoes is a common cause of venous complaints such as hypertension, pain, fatigue, or the feeling of heavy legs and may be a causal factor of venous disease symptoms. The proposition of Potério-Filho et al. [47] that an increase in muscle contractions or greater expulsion of blood that generates higher pressure in the muscle compartments during walking in high heels might be questioned. Indeed, although several studies have shown an increase in muscle activity with high heels [34,75,96,97] or joint restriction [29,98,99], this does not automatically indicate better efficiency of the venous pump (linked with systole and diastole during calf muscle pumping) during walking and in principle this does not tend toward a better venous return [3,16].

It is important to recall here that a difference in the protocols for high heels (*i.e.*, tip-toe movements vs walking) in the studies of Potério-Filho et al. [47] and Tedeschi Filho et al. [48]

might have affected plantar movements and muscle activity differently and thus affected blood flow parameters or muscle pump efficiency differently [16,95]. It also seems that changes in the muscle pump, as in the blood flow, can modify venous pressure.

Effects of shoes on circumferences

Two studies evaluated circumferences during upright standing exercise. Karimi et al. [37] tested various footwear conditions (*i.e.*, barefoot, stable shoe, unstable shoe) and found lower reduction in lower leg circumferences with unstable shoes vs barefoot only, and Lin et al. [45] showed no significant difference in leg circumferences between barefoot and athletic shoes. We therefore note that in addition to similar exercise conditions (*i.e.*, upright standing), these two studies showed similar results, especially between stable shoes or athletic shoes vs barefoot condition, with no significant difference [37,45].

Thus, the findings of the least amount of swelling (*i.e.*, with unstable shoes vs barefoot only) in leg circumferences [37] could be explained by the features of the shoes, especially unstable shoes, which eliminate stasis and promote venous return. As noted, the rocker sole design, which decreases the degree of the stability surface more than other shoe conditions, induces modifications in the kinetics, kinematics, and muscle activity of the lower limbs [28,30,33,36,52]. Several studies have shown an increase in the activity of *gastrocnemius* [33,36] and *tibialis anterior* [28,36] and the range of motion [28,52] with unstable shoes. Moreover, the findings of Karimi et al. [37], in agreement with those of Sousa et al. [100], reveal that standing with unstable shoes also leads to significantly lower muscle co-contraction in the lower leg (*e.g.*, *tibialis anterior* and *gastrocnemius*) compared to barefoot only, while other studies have shown an increase in muscle co-contraction during walking at self-selected speed with unstable shoes compared to regular shoes [101,102]. These kinetic, kinematic and muscle activity changes might activate the muscle pump mechanism and thus increase venous return and venous emptying and decrease the swelling in the lower legs [103]. Indeed, combined active movements produce higher blood flow velocities and would therefore be the most effective in eliminating venous stasis and promoting venous return [62]. The kinetics of venous return are dependent on contraction frequency, with variations in the relaxation periods between muscle contractions [16,64,67,68] according to the type (*e.g.*, walking, standing), intensity (*e.g.*, speed), and duration of the exercise [64]. Although the level of agonist-antagonist muscle co-contraction is dependent on gait velocity [95], it appears that the lower muscle co-contraction in the lower leg with unstable shoes compared to barefoot during standing may be useful for the venous return mechanism and consequently would decrease lower leg circumferences (*i.e.*, venous stasis).

It thus seems that the change in lower leg circumferences depends on the features of the shoe (barefoot, stable, unstable, athletic shoes), which induce modifications in the muscle activity in the form of dynamic contractions, and the characteristics of the exercise (*e.g.*, floor interface, type, duration . . .). Indeed, Lin et al. [45] showed that floor type and standing duration also significantly affect lower leg (*e.g.*, shank) circumferences.

Quality

The heterogeneous data and good ($n = 8$) or fair ($n = 5$) methodological quality of the included studies nevertheless precluded definitive conclusions being drawn. Importantly, this systematic review highlights that further studies are needed to better identify and understand the effects of a range of shoe types on venous status. Future experiments in this area will need to be randomized and controlled and should feature the current shoe models or new shoe models

that have not yet been studied. The focus should be on venous return for athletes and patients with chronic venous insufficiency or at risk of venous thrombosis, for example.

Conclusion

Lower limb venous status is linked with variations in blood flow, venous pressures and lower limb circumferences, which can be modified by the shoe type. The features of unstable shoes (with rockers) or shoes with similar technology (heelless), athletic or soft shoes (air-cushion, air-sole), and customized foot orthotics can improve these venous variables. The features of shoes like high heels (stiletto, platform) and firm shoes and ankle joint immobilization can degrade these venous variables. The variations in venous status can be explained as follows. Shoe types are associated with several changes in the kinetic and kinematic variables and the muscle activity in the lower limbs; the efficiency of the lower limb skeletal muscle pump and foot pump are likely linked to gait kinetics and kinematics; the kinetics of venous return and the regional blood flow distribution are dependent on muscle contraction frequency and muscle fiber types, respectively; and exercise induces several venous changes that are dependent on the type, intensity and duration of the exercise, as well as the physical and physiological conditions of the subjects. It thus seems that it is important to consider the combination of shoe features, subject characteristics and the exercise conditions to interpret the findings about variations in venous variables. However, the data of these 13 studies showed great heterogeneity concerning the type of shoe, the duration or type of exercise and especially the venous parameters, which probably affected our ability to pool data and draw definitive conclusions from the studies within each category of venous variables. Consequently, this systematic review defined the types of shoes or the shoe features that might be better or less effective for venous return, but it was difficult to rank the shoes. Thus, given the heterogeneity in the data, the limited number of studies on each type of shoe, and the fair quality of some of the studies, further research is needed to reach greater consensus on how shoe types affect venous return in the lower limbs. Further studies could evaluate, for example, the effects of new types of shoes on venous return, like the so-called “recovery” shoes that are now used by many athletes after physical activity.

Supporting information

S1 Checklist. PRISMA checklist.
(DOC)

S1 Fig. Flow diagram of study selection. Figure presents the flow of information through the different phases (identification, screening, eligibility and inclusion) of a systematic review.
(DOC)

Author Contributions

Conceptualization: Lucie Lerebourg, Jeremy Coquart.

Data curation: Lucie Lerebourg, Charlotte Menez, Jeremy Coquart.

Formal analysis: Lucie Lerebourg, Jeremy Coquart.

Investigation: Lucie Lerebourg, Jeremy Coquart.

Methodology: Lucie Lerebourg, Charlotte Menez, Jeremy Coquart.

Project administration: Lucie Lerebourg, Jeremy Coquart.

Resources: Lucie Lerebourg.

Software: Lucie Lerebourg, Jeremy Coquart.

Supervision: Lucie Lerebourg, Jeremy Coquart.

Validation: Lucie Lerebourg, Jeremy Coquart.

Visualization: Lucie Lerebourg, Jeremy Coquart.

Writing – original draft: Lucie Lerebourg.

Writing – review & editing: Lucie Lerebourg, Maxime L’Hermette, Charlotte Menez, Jeremy Coquart.

References

1. Meissner MH. Lower Extremity Venous Anatomy. *Semin Intervent Radiol* 2005; 22: 147–156. <https://doi.org/10.1055/s-2005-921948> PMID: 21326687
2. Meissner MH, Moneta G, Burnand K, et al. The hemodynamics and diagnosis of venous disease. *Journal of Vascular Surgery* 2007; 46: S4–S24. <https://doi.org/10.1016/j.jvs.2007.09.043> PMID: 18068561
3. Recek C. Conception of the venous hemodynamics in the lower extremity. *Angiology* 2006; 57: 556–563. <https://doi.org/10.1177/0003319706293117> PMID: 17067977
4. Secomb TW. Hemodynamics. *Compr Physiol* 2016; 6: 975–1003. <https://doi.org/10.1002/cphy.c150038> PMID: 27065172
5. Antle DM, Cormier L, Findlay M, et al. Lower limb blood flow and mean arterial pressure during standing and seated work: Implications for workplace posture recommendations. *Preventive Medicine Reports* 2018; 10: 117–122. <https://doi.org/10.1016/j.pmedr.2018.02.016> PMID: 29850397
6. Bauchot G. Retentissements hémodynamiques des déséquilibres plantaires. *Act Méd Int—Angiologie* 2010; (16) n°5: 2.
7. Healy B, Levin E, Perrin K, et al. Prolonged work- and computer-related seated immobility and risk of venous thromboembolism. *J R Soc Med* 2010; 103: 447–454. <https://doi.org/10.1258/jrsm.2010.100155> PMID: 21037335
8. Lastowiecka-Moras E. How posture influences venous blood flow in the lower limbs: results of a study using photoplethysmography. *Int J Occup Saf Ergon* 2017; 23: 147–151. <https://doi.org/10.1080/10803548.2016.1256938> PMID: 27813456
9. Saggini R, Bellomo RG, Iodice P, et al. Venous insufficiency and foot dysmorphism: effectiveness of visco-elastic rehabilitation systems on veno-muscle system of the foot and of the calf. *Int J Immunopathol Pharmacol* 2009; 22: 1–8.
10. Siegal D, Lim W. Chapter 142—Venous Thromboembolism. In: Hoffman R, Benz EJ, Silberstein LE, et al. (eds) *Hematology (Seventh Edition)*. Elsevier, pp. 2102–2112.
11. Stick C, Hiedl U, Witzleb E. Venous pressure in the saphenous vein near the ankle during changes in posture and exercise at different ambient temperatures. *Eur J Appl Physiol Occup Physiol* 1993; 66: 434–438. <https://doi.org/10.1007/BF00599617> PMID: 8330612
12. Uhl J-F, Gillot C. Anatomy of the foot venous pump: physiology and influence on chronic venous disease. *Phlebology* 2012; 27: 219–230. <https://doi.org/10.1258/phleb.2012.012b01> PMID: 22847928
13. Dagrada A. La pompe musculaire du mollet: «cœur» du système veineux-Rôle du cycle musculaire systole-diastole. *Ed phlébologiques françaises* 2009; 62: 53–57.
14. Black CM. Anatomy and Physiology of the Lower-Extremity Deep and Superficial Veins. *Techniques in Vascular and Interventional Radiology* 2014; 17: 68–73. <https://doi.org/10.1053/j.tvir.2014.02.002> PMID: 24840960
15. Corley GJ, Broderick BJ, Nestor SM, et al. The anatomy and physiology of the venous foot pump. *Anat Rec (Hoboken)* 2010; 293: 370–378. <https://doi.org/10.1002/ar.21085> PMID: 19957343
16. Horwood A. The biomechanical function of the foot pump in venous return from the lower extremity during the human gait cycle: An expansion of the gait model of the foot pump. *Medical Hypotheses* 2019; 129: 109220. <https://doi.org/10.1016/j.mehy.2019.05.006> PMID: 31371072
17. Ricci S, Moro L, Antonelli Incalzi R. The foot venous system: anatomy, physiology and relevance to clinical practice. *Dermatol Surg* 2014; 40: 225–233. <https://doi.org/10.1111/dsu.12381> PMID: 24372905
18. Uhl J-F, Gillot C. Anatomy of the veno-muscular pumps of the lower limb. *Phlebology* 2015; 30: 180–193. <https://doi.org/10.1177/0268355513517686> PMID: 24415543

19. Uhl JF, Bertier C P C, Gillot C. La pompe veineuse plantaire: anatomie et hypothèses physiologiques. *Ed phlébologiques françaises* 2009; 62: 9–18.
20. White JV, Katz ML, Cisek P, et al. Venous outflow of the leg: anatomy and physiologic mechanism of the plantar venous plexus. *J Vasc Surg* 1996; 24: 819–824. [https://doi.org/10.1016/s0741-5214\(96\)70018-6](https://doi.org/10.1016/s0741-5214(96)70018-6) PMID: 8918329
21. Dohm M, Williams KM, Novotny T. Micro-mobile Foot Compression Device Compared with Pneumatic Compression Device. *Clin Orthop Relat Res* 2011; 469: 1692–1700. <https://doi.org/10.1007/s11999-011-1847-5> PMID: 21403988
22. López-López D, Araújo R, Losa-Iglesias ME, et al. Influence of custom foot orthoses on venous status: A quasi-experimental study. *J Mech Behav Biomed Mater* 2018; 79: 235–238. <https://doi.org/10.1016/j.jmbbm.2017.12.035> PMID: 29331936
23. Knight SE, Lewis JD. The effect of footwear on venous pressure at the ankle. *J Cardiovasc Surg (Torino)* 1977; 18: 117–121. PMID: 853053
24. Sousa A, Tavares JMRS, Macedo R, et al. Influence of wearing an unstable shoe on thigh and leg muscle activity and venous response in upright standing. *Appl Ergon* 2012; 43: 933–939. <https://doi.org/10.1016/j.apergo.2012.01.001> PMID: 22277098
25. Yamamoto T, Ohkuwa T, Itoh H, et al. Walking at moderate speed with heel-less shoes increases calf blood flow. *Arch Physiol Biochem* 2000; 108: 398–404. <https://doi.org/10.1076/apab.108.5.398.4296> PMID: 11262597
26. Nakagawa K, Inami T, Yonezu T, et al. Unstable rocker shoes promote recovery from marathon-induced muscle damage in novice runners. *Scand J Med Sci Sports* 2018; 28: 621–629. <https://doi.org/10.1111/sms.12911> PMID: 28513880
27. Chuter VH, Searle A, Spink MJ. Flip-flop footwear with a moulded foot-bed for the treatment of foot pain: a randomised controlled trial. *BMC Musculoskelet Disord*; 17. Epub ahead of print 11 November 2016. <https://doi.org/10.1186/s12891-016-1327-x> PMID: 27835963
28. Nigg B, Federolf PA, von Tscharnar V, et al. Unstable shoes: functional concepts and scientific evidence. *Footwear Science* 2012; 4: 73–82.
29. Cronin NJ. The effects of high heeled shoes on female gait: a review. *J Electromyogr Kinesiol* 2014; 24: 258–263. <https://doi.org/10.1016/j.jelekin.2014.01.004> PMID: 24508305
30. Farzadi M, Nemati Z, Jalali M, et al. Effects of unstable footwear on gait characteristic: A systematic review. *Foot (Edinb)* 2017; 31: 72–76. <https://doi.org/10.1016/j.foot.2017.04.005> PMID: 28577440
31. Franklin S, Grey MJ, Heneghan N, et al. Barefoot vs common footwear: A systematic review of the kinematic, kinetic and muscle activity differences during walking. *Gait Posture* 2015; 42: 230–239. <https://doi.org/10.1016/j.gaitpost.2015.05.019> PMID: 26220400
32. Murley GS, Landorf KB, Menz HB, et al. Effect of foot posture, foot orthoses and footwear on lower limb muscle activity during walking and running: a systematic review. *Gait Posture* 2009; 29: 172–187. <https://doi.org/10.1016/j.gaitpost.2008.08.015> PMID: 18922696
33. Tan JM, Auhl M, Menz HB, et al. The effect of Masai Barefoot Technology (MBT) footwear on lower limb biomechanics: A systematic review. *Gait Posture* 2016; 43: 76–86. <https://doi.org/10.1016/j.gaitpost.2015.10.017> PMID: 26669956
34. Wiedemeijer MM, Otten E. Effects of high heeled shoes on gait. A review. *Gait Posture* 2018; 61: 423–430. <https://doi.org/10.1016/j.gaitpost.2018.01.036> PMID: 29475153
35. Price C, Smith L, Graham-Smith P, et al. The effect of unstable sandals on instability in gait in healthy female subjects. *Gait Posture* 2013; 38: 410–415. <https://doi.org/10.1016/j.gaitpost.2013.01.003> PMID: 23357759
36. Romkes J, Rudmann C, Brunner R. Changes in gait and EMG when walking with the Masai Barefoot Technique. *Clin Biomech (Bristol, Avon)* 2006; 21: 75–81.
37. Karimi Z, Azghani MR, Allahyari T. Lower leg swelling and muscle co-contraction during prolonged standing: an unstable footwear evaluation. *Footwear Science* 2017; 9: 103–110.
38. Maffioletti Nicola A. Increased lower limb muscle activity induced by wearing MBT shoes: Physiological benefits and potential concerns. *Footwear Science* 2012; 4: 123–129.
39. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery* 2010; 8: 336–341. <https://doi.org/10.1016/j.ijsu.2010.02.007> PMID: 20171303
40. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998; 52: 377–384. <https://doi.org/10.1136/jech.52.6.377> PMID: 9764259

41. Trac MH, McArthur E, Jandoc R, et al. Macrolide antibiotics and the risk of ventricular arrhythmia in older adults. *CMAJ* 2016; 188: E120–E129. <https://doi.org/10.1503/cmaj.150901> PMID: 26903359
42. Wright AA, Hegedus EJ, Tarara DT, et al. Exercise prescription for overhead athletes with shoulder pathology: a systematic review with best evidence synthesis. *Br J Sports Med* 2018; 52: 231–237. <https://doi.org/10.1136/bjsports-2016-096915> PMID: 28404557
43. Hooper P, Jutai JW, Strong G, et al. Age-related macular degeneration and low-vision rehabilitation: a systematic review. *Can J Ophthalmol* 2008; 43: 180–187. <https://doi.org/10.3129/i08-001> PMID: 18347620
44. Hansen L, Winkel J, Jørgensen K. Significance of mat and shoe softness during prolonged work in upright position: based on measurements of low back muscle EMG, foot volume changes, discomfort and ground force reactions. *Appl Ergon* 1998; 29: 217–224. [https://doi.org/10.1016/s0003-6870\(97\)00062-8](https://doi.org/10.1016/s0003-6870(97)00062-8) PMID: 9676339
45. Lin Y-H, Chen C-Y, Cho M-H. Influence of shoe/floor conditions on lower leg circumference and subjective discomfort during prolonged standing. *Appl Ergon* 2012; 43: 965–970. <https://doi.org/10.1016/j.apergo.2012.01.006> PMID: 22342130
46. Wezenbeek E, Willems TM, Mahieu N, et al. Is Achilles tendon blood flow related to foot pronation? *Scand J Med Sci Sports* 2017; 27: 1970–1977. <https://doi.org/10.1111/sms.12834> PMID: 28233345
47. Potério-Filho J, Silveira SAF, Potério GMB, et al. The effect of walking with high-heeled shoes on the leg venous pressure. *Angiology* 2006; 57: 424–430. <https://doi.org/10.1177/0003319706290744> PMID: 17022377
48. Tedeschi Filho W, Dezzotti NRA, Joviliano EE, et al. Influence of high-heeled shoes on venous function in young women. *J Vasc Surg* 2012; 56: 1039–1044. <https://doi.org/10.1016/j.jvs.2012.01.039> PMID: 22483354
49. Irzmańska E, Padula G, Irzmański R. Impedance plethysmography as a tool for assessing exertion-related blood flow changes in the lower limbs in healthy subjects. *Measurement* 2014; 47: 110–115.
50. Irzmańska E, Chartusz M, Irzmański R. The use of impedance plethysmography to evaluate the impact of increasing physical activity on blood flow in the lower extremities involving footwear comfort—a preliminary report. *Med Sci Tech* 2011; 52: RA45–RA51.
51. Craik JD, Clark A, Hendry J, et al. The effect of ankle joint immobilization on lower limb venous flow. *Foot Ankle Int* 2015; 36: 18–23. <https://doi.org/10.1177/1071100714552823> PMID: 25249319
52. Roberts S, Birch I, Otter S. Comparison of ankle and subtalar joint complex range of motion during barefoot walking and walking in Masai Barefoot Technology sandals. *J Foot Ankle Res* 2011; 4: 1. <https://doi.org/10.1186/1757-1146-4-1> PMID: 21194494
53. Stacoff A, Nigg BM, Reinschmidt C, et al. Tibiocalcaneal kinematics of barefoot versus shod running. *Journal of Biomechanics* 2000; 33: 1387–1395. [https://doi.org/10.1016/s0021-9290\(00\)00116-0](https://doi.org/10.1016/s0021-9290(00)00116-0) PMID: 10940397
54. Calais-Germain B. Anatomie pour le mouvement: Introduction à l'analyse des technique corporel. Edition Désiris 2013; 5: 260–261.
55. Morio C, Lake MJ, Gueguen N, et al. The influence of footwear on foot motion during walking and running. *J Biomech* 2009; 42: 2081–2088. <https://doi.org/10.1016/j.jbiomech.2009.06.015> PMID: 19643421
56. Clement DB, Taunton JE, Smart GW. Achilles tendinitis and peritendinitis: etiology and treatment. *Am J Sports Med* 1984; 12: 179–184. <https://doi.org/10.1177/036354658401200301> PMID: 6742297
57. Karzis K, Kalogeris M, Mandalidis D, et al. The effect of foot overpronation on Achilles tendon blood supply in healthy male subjects. *Scand J Med Sci Sports* 2017; 27: 1114–1121. <https://doi.org/10.1111/sms.12722> PMID: 27671520
58. Cheung RTH, Chung RCK, Ng GYF. Efficacies of different external controls for excessive foot pronation: a meta-analysis. *British Journal of Sports Medicine* 2011; 45: 743–751. <https://doi.org/10.1136/bjism.2010.079780> PMID: 21504966
59. Telfer S, Abbott M, Steultjens MPM, et al. Dose-response effects of customised foot orthoses on lower limb kinematics and kinetics in pronated foot type. *J Biomech* 2013; 46: 1489–1495. <https://doi.org/10.1016/j.jbiomech.2013.03.036> PMID: 23631857
60. Mills K, Blanch P, Chapman AR, et al. Foot orthoses and gait: a systematic review and meta-analysis of literature pertaining to potential mechanisms. *British Journal of Sports Medicine* 2010; 44: 1035–1046. <https://doi.org/10.1136/bjism.2009.066977> PMID: 19996330
61. Gardner AM, Fox RH. The venous pump of the human foot—preliminary report. *Bristol Med Chir J* 1983; 98: 109–112. PMID: 6616290
62. Sochart DH, Hardinge K. The relationship of foot and ankle movements to venous return in the lower limb. *J Bone Joint Surg Br* 1999; 81: 700–704. <https://doi.org/10.1302/0301-629x.81b4.8909> PMID: 10463749

63. Armstrong RB, Laughlin MH. Exercise blood flow patterns within and among rat muscles after training. *Am J Physiol* 1984; 246: H59–68. <https://doi.org/10.1152/ajpheart.1984.246.1.H59> PMID: 6696089
64. Korthuis RJ. Exercise Hyperemia and Regulation of Tissue Oxygenation During Muscular Activity. Morgan & Claypool Life Sciences, <https://www.ncbi.nlm.nih.gov/books/NBK57139/> (2011, accessed 16 September 2019).
65. Laughlin MH, Armstrong RB. Rat muscle blood flows as a function of time during prolonged slow treadmill exercise. *Am J Physiol* 1983; 244: H814–824. <https://doi.org/10.1152/ajpheart.1983.244.6.H814> PMID: 6859284
66. Laughlin MH, Korthuis RJ, Sexton WL, et al. Regional muscle blood flow capacity and exercise hyperemia in high-intensity trained rats. *J Appl Physiol* 1988; 64: 2420–2427. <https://doi.org/10.1152/jappl.1988.64.6.2420> PMID: 3403424
67. Hoelting BD, Scheuermann BW, Barstow TJ. Effect of contraction frequency on leg blood flow during knee extension exercise in humans. *J Appl Physiol* 2001; 91: 671–679. <https://doi.org/10.1152/jappl.2001.91.2.671> PMID: 11457780
68. Kagaya A. Reduced exercise hyperaemia in calf muscles working at high contraction frequencies. *Eur J Appl Physiol Occup Physiol* 1992; 64: 298–303. <https://doi.org/10.1007/BF00636215> PMID: 1592053
69. Wakeling JM, Pascual SA, Nigg BM. Altering muscle activity in the lower extremities by running with different shoes. 2002.
70. Rao G, Chambon N, Guéguen N, et al. Does wearing shoes affect your biomechanical efficiency? *Journal of Biomechanics* 2015; 48: 413–417. <https://doi.org/10.1016/j.jbiomech.2014.12.038> PMID: 25575884
71. Chien H-L, Lu T-W, Liu M-W. Control of the motion of the body's center of mass in relation to the center of pressure during high-heeled gait. *Gait & Posture* 2013; 38: 391–396.
72. Di Sipio E, Piccinini G, Pecchioli C, et al. Walking variations in healthy women wearing high-heeled shoes: Shoe size and heel height effects. *Gait Posture* 2018; 63: 195–201. <https://doi.org/10.1016/j.gaitpost.2018.04.048> PMID: 29772495
73. Esenyel M, Walsh K, Walden JG, et al. Kinetics of high-heeled gait. *J Am Podiatr Med Assoc* 2003; 93: 27–32. <https://doi.org/10.7547/87507315-93-1-27> PMID: 12533553
74. Rao S, Ripa R, Lightbourne K. Predictors of walking speed and stride length in high- and low-heeled footwear. *Footwear Science* 2013; 5: 179–184.
75. Mika A, Oleksy L, Mika P, et al. The influence of heel height on lower extremity kinematics and leg muscle activity during gait in young and middle-aged women. *Gait Posture* 2012; 35: 677–680. <https://doi.org/10.1016/j.gaitpost.2011.12.001> PMID: 22300729
76. Broderick BJ, Corley GJ, Quondamatteo F, et al. Venous emptying from the foot: influences of weight bearing, toe curls, electrical stimulation, passive compression, and posture. *J Appl Physiol* 2010; 109: 1045–1052. <https://doi.org/10.1152/jappphysiol.00231.2010> PMID: 20705946
77. Styf J. The venous pump of the human foot. *Clinical Physiology* 1990; 10: 77–84. <https://doi.org/10.1111/j.1475-097x.1990.tb00085.x> PMID: 2302938
78. Taylor NAS, Caldwell JN, Mekjavic IB. The sweating foot: local differences in sweat secretion during exercise-induced hyperthermia. *Aviat Space Environ Med* 2006; 77: 1020–1027. PMID: 17042246
79. Osilla EV, Sharma S. Physiology, Temperature Regulation. In: StatPearls. Treasure Island (FL): StatPearls Publishing, <http://www.ncbi.nlm.nih.gov/books/NBK507838/> (2019, accessed 18 September 2019).
80. Branthwaite H, Chockalingam N, Greenhalgh A, et al. The impact of different footwear characteristics, of a ballet flat pump, on centre of pressure progression and perceived comfort. *Foot (Edinb)* 2014; 24: 116–122. <https://doi.org/10.1016/j.foot.2014.05.004> PMID: 24939663
81. Kim J, Lee J. The relationship of creatine kinase variability with body composition and muscle damage markers following eccentric muscle contractions. *J Exerc Nutrition Biochem* 2015; 19: 123–129. <https://doi.org/10.5717/jenb.2015.15061910> PMID: 26244131
82. Ochsmann E, Noll U, Ellegast R, et al. Influence of different safety shoes on gait and plantar pressure: a standardized examination of workers in the automotive industry. *J Occup Health* 2016; 58: 404–412. <https://doi.org/10.1539/joh.15-0193-OA> PMID: 27488038
83. Price C, Andrejevas V, Findlow AH, et al. Does flip-flop style footwear modify ankle biomechanics and foot loading patterns? *J Foot Ankle Res*; 7. Epub ahead of print 26 September 2014. <https://doi.org/10.1186/s13047-014-0040-y> PMID: 25279003
84. Troy Blackburn J, Hirth CJ, Guskiewicz KM. Exercise Sandals Increase Lower Extremity Electromyographic Activity During Functional Activities. *J Athl Train* 2003; 38: 198–203. PMID: 14608427

85. Granacher U, Roth R, Muehlbauer T, et al. Effects of a new unstable sandal construction on measures of postural control and muscle activity in women. *Swiss Med Wkly* 2011; 141: w13182. <https://doi.org/10.4414/smw.2011.13182> PMID: 21491214
86. James DC, Farmer LJ, Sayers JB, et al. The biomechanical characteristics of wearing FitFlop™ sandals highlight significant alterations in gait pattern: a comparative study. *Clin Biomech (Bristol, Avon)* 2015; 30: 347–354. <https://doi.org/10.1016/j.clinbiomech.2015.02.016> PMID: 25823902
87. Chen TL-W, Wong DW-C, Xu Z, et al. Lower limb muscle co-contraction and joint loading of flip-flops walking in male wearers. *PLoS ONE* 2018; 13: e0193653. <https://doi.org/10.1371/journal.pone.0193653> PMID: 29561862
88. Zhang X, Paquette MR, Zhang S. A comparison of gait biomechanics of flip-flops, sandals, barefoot and shoes. *J Foot Ankle Res* 2013; 6: 45. <https://doi.org/10.1186/1757-1146-6-45> PMID: 24196492
89. Weon JH, Cha HG. The influence of high heeled shoes on balance ability and walking in healthy women. *J Phys Ther Sci* 2018; 30: 910–912. <https://doi.org/10.1589/jpts.30> PMID: 30034095
90. Loredó JB, Aguirre MJB, Ruano EJJ. Influence of high heels on walking motion: gait analysis. *ITSEM* 2015; 6.
91. Shroyer JF. Influence of Various Thong Style Flip-flops on Gait Kinematics and Lower Leg Electromyography. 2009.
92. Carl TJ, Barrett SL. Computerized analysis of plantar pressure variation in flip-flops, athletic shoes, and bare feet. *J Am Podiatr Med Assoc* 2008; 98: 374–378. <https://doi.org/10.7547/0980374> PMID: 18820040
93. Chiu M-C, Wang M-J. The effect of gait speed and gender on perceived exertion, muscle activity, joint motion of lower extremity, ground reaction force and heart rate during normal walking. *Gait & Posture* 2007; 25: 385–392. <https://doi.org/10.1016/j.gaitpost.2006.05.008> PMID: 16814548
94. Kagaya A. Levelling-off of calf blood flow during walking and running, and its relationship to anaerobic threshold. *Ann Physiol Anthropol* 1990; 9: 219–224. <https://doi.org/10.2114/ahs1983.9.219> PMID: 2400463
95. Hortobágyi T, Solnik S, Gruber A, et al. Interaction between age and gait velocity in the amplitude and timing of antagonist muscle coactivation. *Gait & Posture* 2009; 29: 558–564. <https://doi.org/10.1016/j.gaitpost.2008.12.007> PMID: 19147360
96. Barton CJ, Coyle JA, Tinley P. The effect of heel lifts on trunk muscle activation during gait: a study of young healthy females. *J Electromyogr Kinesiol* 2009; 19: 598–606. <https://doi.org/10.1016/j.jelekin.2008.03.001> PMID: 18472278
97. Caria PHF. Relation between Wearing High-Heeled Shoes and Gastrocnemius and Erector Spine Muscle Action and Lumbar Lordosis. *Medical Science and Technology*; 55: 71–76.
98. Back TL, Padberg FT, Araki CT, et al. Limited range of motion is a significant factor in venous ulceration. *J Vasc Surg* 1995; 22: 519–523. [https://doi.org/10.1016/s0741-5214\(95\)70030-7](https://doi.org/10.1016/s0741-5214(95)70030-7) PMID: 7494349
99. Lee C. The Effects of Lower Extremity Angle According to Heel-height Changes in Young Ladies in Their 20s during Gait. *J Phys Ther Sci* 2014; 26: 1055–1058. <https://doi.org/10.1589/jpts.26.1055> PMID: 25140096
100. Sousa ASP, Macedo R, Santos R, et al. Influence of wearing an unstable shoe construction on compensatory control of posture. *Human Movement Science* 2013; 32: 1353–1364. <https://doi.org/10.1016/j.humov.2013.07.004> PMID: 24060226
101. Buchecker M, Wagner H, Pfusterschmied J, et al. Lower extremity joint loading during level walking with Masai barefoot technology shoes in overweight males. *Scand J Med Sci Sports* 2012; 22: 372–380. <https://doi.org/10.1111/j.1600-0838.2010.01179.x> PMID: 20807383
102. Horsak B, Heller M, Baca A. Muscle co-contraction around the knee when walking with unstable shoes. *J Electromyogr Kinesiol* 2015; 25: 175–181. <https://doi.org/10.1016/j.jelekin.2014.07.015> PMID: 25156445
103. Lin Y-H, Chen C-Y, Cho M-H. Effectiveness of leg movement in reducing leg swelling and discomfort in lower extremities. *Appl Ergon* 2012; 43: 1033–1037. <https://doi.org/10.1016/j.apergo.2012.03.002> PMID: 22472344

Annexe 2. Article 2 : *Changes in performances/characteristics of French female runners over the last 12 years*



Research in Sports Medicine
An International Journal



ISSN: 1543-8627 (Print) 1543-8635 (Online) Journal homepage: <https://www.tandfonline.com/loi/gspm20>

Changes in performances/characteristics of French female runners over the last 12 years

Lucie Lerebourg & Jeremy B. Coquart

To cite this article: Lucie Lerebourg & Jeremy B. Coquart (2020): Changes in performances/characteristics of French female runners over the last 12 years, *Research in Sports Medicine*, DOI: [10.1080/15438627.2020.1725009](https://doi.org/10.1080/15438627.2020.1725009)

To link to this article: <https://doi.org/10.1080/15438627.2020.1725009>



Published online: 10 Feb 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Full Terms & Conditions of access and use can be found at
<https://www.tandfonline.com/action/journalInformation?journalCode=gspm20>

Changes in performances/characteristics of French female runners over the last 12 years

Lucie Lerebourg and Jeremy B. Coquart

UFR STAPS, CETAPS, Université de Rouen-Normandie, Mont Saint Aignan, France

ABSTRACT

This study investigated the changes in performance of five official track running distances (800m, 1500m, 3000m, 5000m and 10,000m) and the characteristics of French female runners between 2005 and 2016. 22,839 official track performances were scrutinized. For each performance, the race time, the indication of personal record, age, and performance level were recorded. The main results showed that: (1) the number of French track performances has increased in recent years (except for the 10,000m) and the level of these performances has improved significantly only for the longer distances (5000 and 10,000m), (2) the longest distances were generally performed by older athletes, and (3) the highest percentage of international athletes was for the 10,000m event. These findings may be explained by: (1) the development plan put into place by the Fédération France d'Athlétisme, (2) age-dependent physiological and psychological factors and environmental conditions, and (3) the minima to access the international level.

ARTICLE HISTORY

Received 10 September 2019
Accepted 29 January 2020

KEYWORDS

Run; race; competition;
athlete; olympic cycles

Introduction

Understanding the circumstances in which sports performances are executed may be of great interest for athletes, coaches, sports events organizers and sports federations. This information may be helpful for selecting a speciality, building competitive strategies, choosing the day for a competitive event, or even identifying priorities for developing a sport. However, despite the numerous researches on long-distance road races such as the half-marathon or marathon (Cuk, Nikolaidis, & Knechtle, 2019; Cushman, Markert, & Rho, 2014; Nikolaidis & Knechtle, 2018; Nikolaidis, Rosemann, & Knechtle, 2018; Roeh et al., 2019; Vickers & Vertosick, 2016; Vieluf, Scheer, Hasija, Schreier, & Reinsberger, 2019) the studies for shorter running distances on track (e.g., 10,000m) are scarce, and more especially for athletes of modest level.

Several authors have studied the progression in performances to identify the trainable variables (that is, those variables related to physiology, psychology, tactics, pacing strategy, etc that can be improved through training) and other factors beyond the athlete's control (e.g., demographics, climate conditions, seasonal characteristics ...) that might influence the long distance performances (Cuk et al., 2019; Cushman et al., 2014, 2019;

CONTACT Lucie Lerebourg  lulu.lerebourg@gmail.com  UFR STAPS, CETAPS, Université de Rouen-Normandie, Boulevard Siegfried, Mont Saint Aignan 76 821, France

© 2020 Informa UK Limited, trading as Taylor & Francis Group

Nikolaidis & Knechtle, 2018; Nikolaidis et al., 2018), especially world's best marathon performances (El Helou et al., 2012; Marc et al., 2014; Nikolaidis et al., 2019). For example, peak performance is reached at about 25–27 years old for world-class female athletes (Haugen et al., 2018), probably because of age-related physiological limits. Moreover, a body mass index of $18.2 \text{ kg}\cdot\text{m}^{-2}$ seems optimal for Top100 female marathoners, and these women generally compete in April (27.6%) or October (21.1%) (Marc et al., 2014). These different studies analysing the performances and/or characteristics of international female runners could help to better understand the high-level performance and changes of performances, but they concern only a small number of European athletes, who are under-represented in the World Top100. Indeed, Marc et al. (2014) indicated that 52% of the 100 best female marathoners in 2011 were African, whereas only 24% were European (with no French women among the Top10 Europeans). More recently, Marc et al. (2017) confirmed this superiority of African women athletes on marathon, but equally on the track. The authors further pointed out that 41% of the women who participated in a 3000m event in 2012 and were also in the world's Top100 from East Africa (Marc et al., 2017). However, the majority of athletes have not an international level, and the most of studies did not analyse the performance in athletes of modest level (*i.e.*, national, regional, under-regional ...) (Allen & Hopkins, 2015; Berthelot et al., 2012; Haugen et al., 2018; Hollings, Hopkins, & Hume, 2014; Marc et al., 2017). Consequently, the information included in these articles may not be representative of most French athletes, and an investigation more narrowly focused on the performances of European athletes, such as French female runners, is therefore needed to provide new data for these athletes, their coaches and the federations preparing for future international competitions, like the Olympic Games or the European and/or World championships. To date, no study has examined the progress in French track performances, especially those of women.

The purpose of this study was therefore to examine the progress in performances for five official track distances and the characteristics of French female runners over the last three Olympic cycles.

Methods

The official rankings of the *Fédération Française d'Athlétisme* (FFA) for women between 23 and 39 years (*i.e.*, range of age corresponding to the "senior/adult" category in France) competing in the 800m, 1500m, 3000m, 5000m and 10,000m events from 1 January 2005 to 31 December 2016 were examined from the FFA's official reference table. On-road performances like the 10km were excluded, and only track performances were kept. We limited the study to track events in order to control for any effects of the race profile on performance. For each distance and year, only the best performance of an athlete was retained. A total of 22,839 performances were collected. For each performance, the race time (performance), the indication of personal record, and the athlete's birth date were recorded.

Moreover, from their race times, the FFA official ranking in 2010 (current ranking) was used to class the women into 5 groups of level (*i.e.*, under-regional, regional, inter-regional, national and international; Table 1).

This study was approved by the National Ethics Committee for Research in Sports Sciences (CERSTAPS 2019-22-02-31). Moreover, the protocol for this study was legally declared, in accordance with the European General Data Protection Regulations.

Table 1. Official level of running performances on track in French women.

Distance	International	National	Inter-regional	Regional	Under-Regional
800m	2'01"84	2'17"00	2'28"00	2'46"00	3'08"00
1500m	4'11"00	4'41"00	5'05"00	5'40"00	6'50"00
3000m	9'05"00	10'05"00	11'00"00	12'15"00	14'40"00
5000m	15'30"00	17'40"00	19'00"00	21'00"00	26'00"00
10000m	32'40	37'00	41'00	47'00	1h00'00

Legend: Level according to the FFA official ranking in 2010.

Statistical analysis

Data are reported as number of cases and percentages for categorical data, or as mean \pm standard deviation (SD) for numerical data.

For each year and each distance, the number of athletes, each athlete's best annual performance, the race times, the number of personal records, the athletes' ages, and the number of athletes in each performance level were analysed. The standards for classifying athletes in each performance level are presented in [Table 1](#).

This descriptive statistical analysis was then repeated for only the French Top10 runners to obtain information about the best female runners.

Furthermore, these descriptive analyses were complemented by inferential statistical analyses. To do so, the performances were regrouped into 4-year periods that corresponded to Olympic Game cycles. Thus, 2005–2008 corresponded to Athens-Beijing ($P_{2005-2008}$), 2009–2012 to Beijing-London ($P_{2009-2012}$), and 2013–2016 to London-Rio ($P_{2013-2016}$).

The numerical data were compared between the Olympic cycles with 3-way analysis of variance (groups: global population vs Top10 \times periods: $P_{2005-2008}$ vs $P_{2009-2012}$ vs $P_{2013-2016}$ \times distances: 800m vs 1500m vs 3000m vs 5000m vs 10,000m). When significant differences were obtained, a Tukey post-hoc test was conducted. The magnitude of the difference was assessed by the effect size (ES). The scale proposed by Cohen (1988) was used for interpretation. The magnitude of the difference was considered to be trivial ($ES < 0.2$), small ($0.2 \leq ES < 0.5$), moderate ($0.5 \leq ES < 0.8$), and large ($ES \geq 0.79$).

For the categorical data, the frequencies were compared using descriptive statistical analysis.

Statistical significance was set at $p < 0.05$ and all analyses were performed with the Statistical Package for the Social Sciences (release 20.0, Chicago, IL, USA).

Results

The number of performances was 7400 for $P_{2005-2008}$, 7293 for $P_{2009-2012}$, and 8146 for $P_{2013-2016}$. The highest number of performances occurred in $P_{2013-2016}$ for all distances, except for the 10,000m (90 performances vs 113 for $P_{2005-2008}$ and 102 for $P_{2009-2012}$).

The mean performances for both groups (Top10 and all runners) in the 5000m were significantly improved in $P_{2013-2016}$ compared with $P_{2005-2008}$ ($p = 0.010$, with $ES = 0.20$ for Top10 and $ES = 0.15$ for all runners; [Table 2](#)). Furthermore, for the 10,000m, the mean performances (for both groups) were significantly lower in $P_{2005-2008}$ compared with other periods ($p < 0.001$). Although a trivial effect was observed between $P_{2005-2008}$ and

Table 2. Evolution of performances in French female performers between 2005 and 2016.

Distance	Group	Period		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800m	Best performance	1'59"	1'59"	1'58"
	Top10 performances	2'03"±2"	2'04"±2"	2'04"±2"
	Average performances	2'36"±17"	2'36"±16"	2'35"±16"
1500m	Best performance	4'00"	4'00"	4'03"
	Top10 performances	4'15"±7"	4'15"±7"	4'16"±4"
	Average performances	5'22"±30"	5'22"±30"	5'20"±30"
3000m	Best performance	8'35"	8'56"	8'52"
	Top10 performances	9'17"±12"	9'30"±13"	9'25"±10"
	Average performances	11'52"±1'15"	11'52"±1'13"	11'52"±1'16"
5000m	Best performance	14'59"	15'21"	15'08"
	Top10 performances	16'22"±28"	16'26"±28"	16'17"±26" ^a
	Average performances	19'47"±2'04"	19'38"±2'02"	19'28"±2'00" ^a
10000m	Best performance	31'57"	31'36"	31'53"
	Top10 performances	38'03"±3'19"	37'47"±3'22" ^{a,c}	36'02"±2'03" ^{a,b}
	Average performances	44'07"±5'47"	42'40"±5'05" ^{a,c}	42'04"±6'39" ^{a,b}

Legend: ^aSignificant difference from P₂₀₀₅₋₂₀₀₈ ($p < 0.01$), ^bSignificant difference from P₂₀₀₉₋₂₀₁₂ ($p < 0.001$), ^cSignificant difference from P₂₀₁₃₋₂₀₁₆ ($p < 0.001$).

P₂₀₀₉₋₂₀₁₂ for Top10 (ES = 0.08), ES was higher for all runners (ES = 0.27, *i.e.*, small effect). Moreover, a similar ES was noted for all runners (ES = 0.33) and moderate effect for Top10 (ES = 0.73) between P₂₀₀₅₋₂₀₀₈ and P₂₀₁₃₋₂₀₁₆. Last, the performances for the 10,000m were also significantly higher (for both groups) in P₂₀₁₃₋₂₀₁₆ compared to P₂₀₀₉₋₂₀₁₂ ($p < 0.001$; ES = 0.60 for Top10 and ES = 0.10 for all runners). No significant difference was noted for the 800m, 1500m and 3000m whatever the group and across time periods (Table 2).

For the whole population, the number of personal records for the 800m was higher in P₂₀₀₅₋₂₀₀₈ and for the 1500m the number was slightly lower in P₂₀₀₉₋₂₀₁₂ (Table 3).

In the Top10, the number of personal records for the 800m and 1500m was higher in P₂₀₁₃₋₂₀₁₆ (Table 3).

For the 3000m, 5000m or 10,000m, whatever the group, the number of personal records was close.

Table 3. Evolution of percentages of personal records in French female performers between 2005 and 2016.

Distance	Group	Period		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800m	Best performance	1 (100.0)	1 (100.0)	1 (100.0)
	Top10 performances	9 (22.5)	8 (20.0)	20 (50.0)
	Average performances	936 (41.6)	770 (34.9)	908 (36.7)
1500m	Best performance	0 (0.0)	1 (100.0)	1 (100.0)
	Top10 performances	9 (22.5)	7 (17.5)	21 (52.5)
	Average performances	981 (42.1)	822 (35.6)	1022 (40.7)
3000m	Best performance	1 (100.0)	1 (100.0)	1 (100.0)
	Top10 performances	18 (45.0)	15 (37.5)	16 (40.0)
	Average performances	1035 (47.5)	991 (45.5)	1178 (47.7)
5000m	Best performance	0 (0.0)	0 (0.0)	1 (100.0)
	Top10 performances	10 (25.0)	14 (35.0)	19 (47.5)
	Average performances	273 (52.0)	255 (51.6)	319 (53.0)
10000m	Best performance	0 (0.0)	1 (100.0)	1 (100.0)
	Top10 performances	26 (65.0)	25 (62.5)	25 (62.5)
	Average performances	79 (69.9)	71 (69.6)	65 (72.2)

Legend: The numbers represent the number of athletes (%) who realized a personal record.

Table 4. Evolution of ages in French female performers between 2005 and 2016.

Distance	Group	Period		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800m	Best performance	23	31	26
	Top10 performances	28.7 ± 3.7	28.0 ± 3.6	26.7 ± 2.7
	Average performances	28.7 ± 4.7	28.7 ± 4.8	28.6 ± 4.8 ^a
1500m	Best performance	26	31	23
	Top10 performances	29.6 ± 3.9	29.6 ± 3.7	27.7 ± 3.6
	Average performances	30.1 ± 5.0	29.9 ± 5.0	29.9 ± 5.0 ^a
3000m	Best performance	26	27	29
	Top10 performances	30.0 ± 3.6	30.0 ± 4.3	29.3 ± 4.5
	Average performances	31.4 ± 4.9	31.5 ± 5.1	31.2 ± 5.0 ^a
5000m	Best performance	31	38	24
	Top10 performances	31.4 ± 2.7	30.1 ± 5.2	29.7 ± 4.7
	Average performances	31.4 ± 4.8	30.8 ± 5.1	30.9 ± 4.8 ^a
10000m	Best performance	34	38	24
	Top10 performances	32.7 ± 3.6	33.4 ± 4.7	31.4 ± 5.4
	Average performances	32.8 ± 4.4	33.4 ± 5.1	32.5 ± 5.2 ^a

Legend: ^a Significant difference from Top10 ($p < 0.05$). Note: Significant differences between each distance, except between 3000 and 5000m ($p < 0.05$).

The longer the distances were, the older the athletes were (whatever the group; $p < 0.001$), except between the 3000m and the 5000m ($p = 0.103$, Table 4).

The athletes in the Top10 were significantly younger than the group of all runners in P₂₀₁₃₋₂₀₁₆ ($p < 0.05$; ES was 0.40 for 800m, 0.44 for 1500m, 0.38 for 3000m, 0.25 for 5000m and 0.21 for 10,000m; Table 4). Significant differences between each distance were noted, except between 3000m, 5000m and no significant difference was noted for each distance whatever the period ($p > 0.05$).

For the 10,000m, the percentage of international French female performers was higher (3.9% to 6.2%) whatever the period than 800m, 1500m, 3000m and 5000m distances (0.1% to 0.6%).

Discussion

The aim of this study was to determine the evolution in performances of French adult female runners for five official track distances over the last three Olympic cycles.

The results show that the number of performances in the 800m, 1500m, 3000m and 5000m events increased over the P₂₀₁₃₋₂₀₁₆ period for these runners. The increase might be explained by the rise in the number of athletics clubs in this period, from 2252 in 2013 to 2399 in 2016 (FFA, 2017), as well as a rise in the number of FFA women licence holders. Indeed, for women "senior/adult" (between 23 and 39 years), the FFA has indicated high growth at the end of each Olympic cycle, with 7622 members in 2004; 9965 in 2008 (+30.7%); 12,969 in 2012 (+30.1%); and 17,932 in 2016 (+38.3%). This increase in the number of FFA licence holders may also be explained by the development plan that the FFA has put into place over the last several years (FFA, 2017). For example, in P₂₀₁₃₋₂₀₁₆, the FFA underwent a federal reorganization and since then it has been organized in accord with the emergence of new businesses (with a gain of 22.6 M€ in 2015 vs 17.7 M€ in 2012) and the implementation of new means to meet specific objectives (e.g., territorial development and restructuring of athletics clubs, renovation of the FFA website, communications to promote athletics with a more attractive image ...) (FFA, 2017).

Although the number of performances in the 800m, 1500m, 3000m and 5000m events increased in P₂₀₁₃₋₂₀₁₆ for the French female athletes, it decreased in the 10,000m. This decline might be explained by runners' migration towards the more popular road races with the same running distance (e.g., 10km) as the 10,000m on track. Indeed, although the number of road races like the 10km (official competition) or trail running has substantially increased in the past few years, the number of 10,000m track races is limited in France (FFA, 2018). For example, 10,000m track races are organized much more often in the United States or Japan, where the density of good-level performances seems better (Bazin, 2018). Opportunities to enter 10,000m races remain limited in France, and the main opportunity is the 10,000m national championships.

The results show that the performances of both groups were improved for the 5000m and 10,000m in P₂₀₁₃₋₂₀₁₆ (Table 2). This improvement was probably linked to such factors as technological progress (e.g., running shoes.) and even the development plan put into place by the FFA (e.g., directives of the development of road-races) (FFA, 2012, 2016, 2017). With the evolution of shoes technology, some regional even national level athletes of longer distances such as 10,000m prefer to be equipped with shoes that promote the absorption of shocks (less traumatic/to decrease the risk of the medical events compared to spikes shoes) or that promote the restitution of energy with carbon plates to optimize their performance (e.g., the first marathon in less than 2 hours has been ran with the type of shoes). Moreover, the directives of the federation are, for example, to increase the popularity and the number of road-races (i.e., development of "hors-stade" sector) and to give athletes the possibility run with density and equally to qualify on 10,000m at the national championships (France) with a 10km road performance (only for the 10,000m, not possible for others distances on track).

The results also reveal an increase in the number of personal records in the 800 and 1500m for the Top10 (Table 3). One of the reasons for this rise might be the progressive decrease in the mean age of these Top10 athletes. Indeed, the average age for Top10 women in the 800m was 28.7 ± 3.7 years in P₂₀₀₅₋₂₀₀₈, 28.0 ± 3.6 years for P₂₀₀₉₋₂₀₁₂ and only 26.7 ± 2.7 years in P₂₀₁₃₋₂₀₁₆ (Table 4). The same trend was noted for the 1500m (27.7 years in P₂₀₁₃₋₂₀₁₆ vs 29.6 years for the other two periods). Several studies have identified the mean age for peak performance, generally observing that for world-class female athletes it is about 25–27 years old (Berthelot et al., 2012; Haugen et al., 2018; Hollings et al., 2014; Longo, Siffredi, Cardey, Aquilino, & Lentini, 2016). So, it seems possible to reach peak performance at older ages in the longest distances, which required a higher endurance capacity (e.g., 5000 and 10,000m) than in middle distances characterized by higher aerobic power (e.g., 800 and 1500m). This difference, in agreement with several studies (Allen & Hopkins, 2015; Longo et al., 2016) could be explained by the physiological parameters characterizing human capacities like the ability to produce aerobic power or to develop endurance capacity and also the knowledge or experience that evolve with ageing (Hollings et al., 2014; Nikolaidis & Knechtle, 2019; Schulz & Curnow, 1988; Tanaka & Seals, 2008).

The high number of personal records in the 800m and 1500m for the Top10 in P₂₀₁₃₋₂₀₁₆ may also be linked to the higher number of competitions in the same period (FFA (n.d.)). In contrast, the smaller of the total number of 3000m, 5000m and 10,000m events organized at national circuit French meets during each Olympic cycle (in P₂₀₀₉₋₂₀₁₂ and P₂₀₁₃₋₂₀₁₆, respectively, 13 and 8 for the 3000m, 10 and 14 for the 5000m, and 0 for the

10,000m in both periods) may be linked to lower opportunities to set new records for the Top10 athletes (FFA (n.d.)).

For all runners (global population including Top10), the number of personal records for the 800m was higher in P₂₀₀₅₋₂₀₀₈ than in the other two periods (41.6% records set in P₂₀₀₅₋₂₀₀₈ vs 34.9 and 36.7 for P₂₀₀₉₋₂₀₁₂ and P₂₀₁₃₋₂₀₁₆, respectively) (Table 3). Again, the French athletes may have benefited from more competitions. In addition, the changes in the qualifying procedures for the national championships may have encouraged them to break their personal records to achieve the minima. Indeed, as noted in the official FFA rules in 2012, it is possible that the change in qualifications for the national championships in 2011 (the abandonment of qualifying times vs the introduction of ranking) had a slight influence.

Last, the generally lower number of personal records for P₂₀₀₉₋₂₀₁₂ might be explained by a relative lack of competitions and less favourable environmental conditions. For example, precipitation, wind speed and direction, and even air temperature influence running performances (El Helou et al., 2012; Marc et al., 2014; Nikolaidis et al., 2019). Moreover, the best performances were often in the spring and autumn (Marc et al., 2014). For the middle running distances like the 1500m, it has been shown that the best performances are performed at 20-25°C as this temperature range seems to favour optimal human physiological capacities (Haïda et al., 2013). However, the official national data indicate that spring 2011 was the hottest ever recorded and that autumn 2011 was the second hottest autumn since 1900. These unaccustomed environmental conditions probably influenced the number of personal records in P₂₀₀₉₋₂₀₁₂.

The study shows that the longer the distances were, the older the athletes were, whatever the group, except for the 3000m and 5000m events, where there was no significant difference in ages (Table 4). According to Longo et al. (2016), this observation suggests that the age for peak running performance increases with the length of the race.

Our study confirmed this pattern, with, for example, women in the Top10 achieving their peak performances at a younger age in the 800m than their counterparts in the 10,000m (26.7 ± 2.7 vs 31.4 ± 5.4 , respectively). However, the age for peak performance in the French Top10 seems to be slightly higher than in the best world-class athletes (25.8 ± 4.8 years for running middle distance vs 28.9 ± 4.7 years for running long distance) (Longo et al., 2016). This younger age for the Top10 athletes may be explained by the FFA's policies (e.g., *Kid Stadium* in 1998 or even *Planètes Athlé Jeunes* in 2008 which aimed to (a) ensure that children/adolescents could discover the activities of jumping, running, throwing and developing athletic skills with an emphasis on fun so that they could continue to practice athletics after the age of 16 years, specializing according to their wishes and qualities, and (b) develop workspaces and equipments adapted to children and easily implant on various sites to train children/adolescents and detect young talents, to orient them towards an athletic speciality around the age of 16 years. Athletes who start specialized training at a young age (i.e., from 16 years, so U18 category characterized by an increase focus on performance and specific skills development) may be more likely to reach their peak performance at an earlier age than their counterparts who specialize somewhat later (Noble & Chapman, 2018; Onywera, 2009; Smith, 2003).

The Top10 athletes were younger than the others in all event races. Therefore, age seems to be a factor that limits peak performance due to age-related physiological modifications. For example, although the peak performance in heterogeneous long-distance runners depends on several physiological factors (such as aerobic endurance capacity and running energy cost), the main determinant of performance is maximal oxygen uptake (Bosquet, Léger, & Legros, 2002; Di Prampero, Atchou, Brückner, & Moia, 1986). Physiologists often consider that maximal oxygen uptake is bounded by a maximal limit in arteriovenous oxygen difference, stroke volume and heart rate (Levine, 2008), and it is widely acknowledged that maximal heart rate decreases with age (Nikolaidis et al., 2018). Thus, it is logical that the Top10 athletes, who often have higher maximal oxygen uptake than other runners, were younger.

The percentage of female athletes at each performance level was fairly constant across periods (Table 5). However, the descriptive analysis of these data shows a higher percentage of international competitors in the 10,000m compared with the other races, whatever the period (3.9–6.2% for 10,000m vs 0.1–0.6% for all other distances and all periods, Table 5). This higher percentage is perhaps explained by a better level of athletes specializing in this race or relatively easier qualifying procedures. In 2016, for example, the minima imposed by the FFA for the international level (Table 1) was 94.6, 93.8, 92.3, 91.7 and 89.6% of the best annual world performance in the 800m, 1500m, 3000m, 5000m and 10,000m events, respectively, suggesting that it is easier to reach a high level on long running distances than short distances. Therefore, some French athletes may have specifically trained for the 10,000m, hoping to qualify more easily for the French team.

Table 5. Evolution of number and percentage of French female performers in each performance level between 2005 and 2016.

Distance	Level	Period		
		P ₂₀₀₅₋₂₀₀₈	P ₂₀₀₉₋₂₀₁₂	P ₂₀₁₃₋₂₀₁₆
800m	International	13 (0.6)	8 (0.4)	11 (0.4)
	National	212 (9.4)	222 (10.1)	241 (9.7)
	Interregional	588 (26.1)	522 (23.6)	607 (24.6)
	Regional	896 (39.8)	905 (41.0)	1040 (42.1)
	Under-regional	542 (24.1)	551 (25.0)	573 (23.2)
1500m	International	12 (0.5)	9 (0.4)	8 (0.3)
	National	182 (7.8)	192 (8.3)	258 (10.3)
	Interregional	502 (21.5)	516 (22.3)	581 (23.1)
	Regional	973 (41.7)	921 (39.9)	1013 (40.3)
	Under-regional	662 (28.4)	672 (29.1)	653 (26.0)
3000m	International	8 (0.4)	2 (0.1)	3 (0.1)
	National	115 (5.3)	110 (5.0)	136 (5.5)
	Interregional	433 (19.9)	403 (18.5)	531 (21.5)
	Regional	893 (41.0)	930 (42.7)	919 (37.2)
	Under-regional	731 (33.5)	734 (33.7)	880 (35.6)
5000m	International	3 (0.6)	3 (0.6)	2 (0.3)
	National	84 (16.0)	97 (19.6)	115 (19.1)
	Interregional	109 (20.8)	94 (19.0)	154 (25.6)
	Regional	182 (34.7)	181 (36.6)	186 (30.9)
	Under-regional	147 (28.0)	119 (24.1)	145 (24.1)
10000m	International	7 (6.2)	4 (3.9)	5 (5.6)
	National	8 (7.1)	11 (10.8)	22 (24.4)
	Interregional	13 (11.5)	22 (21.6)	18 (20.0)
	Regional	52 (46.0)	40 (39.2)	20 (22.2)
	Under-regional	33 (29.2)	25 (24.5)	25 (27.8)

Conclusion

The main results show that the number of track performances has increased over the past few years (except for the 10,000m) and that performances have improved significantly for the long distances (*i.e.*, 5000 and 10,000m). These findings are probably explained by the development plan put into place by the FFA. Moreover, the longest distances were generally performed by older female athletes. Furthermore, a higher percentage of international athletes competed in the 10,000m (possibly because of lower minima). Therefore, the development plan put into place by the (FFA, 2017) seems to have contributed to encouraging French women to compete in track races and to improving running performances in long distances. Peak performances depend on physiological factors (some of which are age-dependent), as well as on psychological, technical and tactical factors and environmental conditions. Further studies are thus needed to better understand the conditions in which the peak performances are executed in France and the world.

Acknowledgments

The authors are grateful to the Institut de Recherche Interdisciplinaire Homme et Société (IRIHS) of Université de Rouen-Normandie, and the Commission de la Documentation et de l'Histoire (CDH) of the FFA.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Allen, S. V., & Hopkins, W. G. (2015). Age of peak competitive performance of elite athletes: A systematic review. *Sports Medicine (Auckland, N.Z.)*, 45(10), 1431–1441.
- Bazin, S. (2018, May 22). Le grand renouveau du 10 000 m européen. *Winder Mag*. <https://www.widermag.com/news-le-grand-renouveau-10-000-europeen>
- Berthelot, G., Len, S., Hellard, P., Tafflet, M., Guillaume, M., Vollmer, J.-C., ... Toussaint, J.-F. (2012). Exponential growth combined with exponential decline explains lifetime performance evolution in individual and human species. *Age (Dordrecht, Netherlands)*, 34(4), 1001–1009.
- Bosquet, L., Léger, L., & Legros, P. (2002). Methods to determine aerobic endurance. *Sports Medicine (Auckland, N.Z.)*, 32(11), 675–700.
- Cuk, I., Nikolaidis, P. T., & Knechtle, B. (2019). Sex differences in pacing during half-marathon and marathon race. *Research in Sports Medicine (Print)*, 1–10. doi:10.1080/15438627.2019.1593835
- Cushman, D. M., Markert, M., & Rho, M. (2014). Performance trends in large 10-km road running races in the United States. *Journal of Strength and Conditioning Research*, 28(4), 892–901.
- Di Prampero, P. E., Atchou, G., Brückner, J. C., & Moia, C. (1986). The energetics of endurance running. *European Journal of Applied Physiology and Occupational Physiology*, 55(3), 259–266.
- El Helou, N., Tafflet, M., Berthelot, G., Tolaini, J., Marc, A., Guillaume, M., ... Toussaint, J.-F. (2012). Impact of environmental parameters on marathon running performance. *PloS One*, 7(5), e37407.
- FFA. (2012). *Plan de développement 2013–2017* | Fédération Française d'Athlétisme. Paris, France: F. F. d'Athlétisme. Retrieved from http://www.athle.fr/pdf/docffa/Plan_Developpement_2013-2017.pdf

- FFA. (2016). *Règlement des compétitions nationales 2016* (pp. 61–62). Retrieved from https://www.athle.fr/Reglement/livret_2016.pdf
- FFA. (2017). *Rapport d'activités 2012–2016*. Paris, France: F. F. d'Athlétisme. www.athle.fr/pdf/AG2016/Rapport-activite-2012-2016.pdf
- FFA. (2018). *Plan de développement 2018–2024 | Fédération Française d'Athlétisme*. Paris, France: F. F. d'Athlétisme. Retrieved from <https://www.athle.fr/asp.net/main.html/html.aspx?htmlid=5763>
- FFA. (n.d.). *Tous les résultats 2019*. Retrieved from <http://www.athle.fr/asp.net/main.html/html.aspx?htmlid=5257>.
- Haïda, A., Dor, F., Guillaume, M., Quinquis, L., Marc, A., Marquet, L.-A., ... Toussaint, J.-F. (2013). Environment and scheduling effects on sprint and middle distance running performances. *PLoS One*, 8(11), e79548.
- Haugen, T. A., Solberg, P. A., Foster, C., Morán-Navarro, R., Breitschädel, F., & Hopkins, W. G. (2018). Peak age and performance progression in world-class track-and-field athletes. *International Journal of Sports Physiology and Performance*, 13(9), 1122–1129.
- Hollings, S. C., Hopkins, W. G., & Hume, P. A. (2014). Age at peak performance of successful track & field athletes. *International Journal of Sports Science & Coaching*, 9(4), 651–661.
- Levine, B. D. (2008). VO₂max: What do we know, and what do we still need to know? *The Journal of Physiology*, 586(1), 25–34.
- Longo, A. F., Siffredi, C. R., Cardey, M. L., Aquilino, G. D., & Lentini, N. A. (2016). Age of peak performance in Olympic sports: A comparative research among disciplines. *Journal of Human Sport and Exercise*, 11(1), 31–41.
- Marc, A., Sedeaud, A., Guillaume, M., Rizk, M., Schipman, J., Antero-Jacquemin, J., ... Toussaint, J.-F. (2014). Marathon progress: Demography, morphology and environment. *Journal of Sports Sciences*, 32(6), 524–532.
- Marc, A., Sedeaud, A., Schipman, J., Jacquemin, J. A., Saulière, G., Kryger, K. O., & Toussaint, J.-F. (2017). Geographic enrolment of the top 100 in athletics running events from 1996 to 2012. *The Journal of Sports Medicine and Physical Fitness*, 57(4), 418–425.
- Nikolaidis, P. T., Di Gangi, S., Chtourou, H., Rüst, C. A., Rosemann, T., & Knechtle, B. (2019). The role of environmental conditions on marathon running performance in men competing in Boston Marathon from 1897 to 2018. *International Journal of Environmental Research and Public Health*, 16, 4.
- Nikolaidis, P. T., & Knechtle, B. (2018). Pacing in age group marathoners in the “New York City Marathon.”. *Research in Sports Medicine (Print)*, 26(1), 86–99.
- Nikolaidis, P. T., & Knechtle, B. (2019). Force-velocity characteristics and maximal anaerobic power in male recreational marathon runners. *Research in Sports Medicine (Print)*, 1–12. doi:10.1080/15438627.2019.1608993
- Nikolaidis, P. T., Rosemann, T., & Knechtle, B. (2018). Age-predicted maximal heart rate in recreational marathon runners: A cross-sectional study on fox's and tanaka's equations. *frontiers in physiology*, 9, 226.
- Noble, T. J., & Chapman, R. F. (2018). Marathon specialization in elites: A head start for Africans. *International Journal of Sports Physiology and Performance*, 13(1), 102–106.
- Onywera, V. O. (2009). East African runners: Their genetics, lifestyle and athletic prowess. *Medicine and Sport Science*, 54, 102–109.
- Roeh, A., Bunse, T., Lembeck, M., Handrack, M., Pross, B., Schoenfeld, J., ... Scherr, J. (2019). Running effects on cognition and plasticity (ReCaP): Study protocol of a longitudinal examination of multimodal adaptations of marathon running. *Research in Sports Medicine (Print)*, 1–15. doi:10.1080/15438627.2019.1647205
- Schulz, R., & Curnow, C. (1988). Peak performance and age among superathletes: Track and field, swimming, baseball, tennis, and golf. *Journal of Gerontology*, 43(5), NaN–NaN.
- Smith, D. J. (2003). A framework for understanding the training process leading to elite performance. *Sports Medicine (Auckland, N.Z.)*, 33(15), 1103–1126.

- Tanaka, H., & Seals, D. R. (2008). Endurance exercise performance in Masters athletes: Age-associated changes and underlying physiological mechanisms. *The Journal of Physiology*, 586(1), 55–63.
- Vickers, A. J., & Vertosick, E. A. (2016). An empirical study of race times in recreational endurance runners. *BMC Sports Science, Medicine & Rehabilitation*, 8(1), 26.
- Vieluf, S., Scheer, V., Hasija, T., Schreier, P. J., & Reinsberger, C. (2019). Multimodal approach towards understanding the changes in the autonomic nervous system induced by an ultramarathon. *Research in Sports Medicine (Print)*, 1–10. doi:10.1080/15438627.2019.1665522

Annexe 3. Article 3 : *Prediction of middle-distance running performances of female runners using nomograms*

Manuscript submitted to editorial office



Prediction of middle-distance running performances of female runners using nomograms

Journal:	<i>International Journal of Sports Medicine</i>
Manuscript ID	Draft
Manuscript Type:	Training & Testing
Key word:	estimation, modelling, time to exhaustion, validity, female athletes
Abstract:	<p>Purpose: This study examined the validity, precision and accuracy of the predictions of middle-distance performances in female runners from three nomograms. Methods: Official rankings of French women for the 3000-m, 5000-m, and 10,000-m track-running events from 2005 to 2019 were examined. Only female runners who performed in the three distance events within the same year were included (n=158). Each performance over any distance was predicted using the three nomograms from the two other performances. Results: The 3000-m, 5000-m and 10,000-m performances were 11min17 s ± 1min20 s, 19min29 s ± 2min20 s, 41min18 s ± 5min7 s, respectively. No difference was found between the actual and predicted running performances regardless of the nomogram (p>0.05). All predicted running performances were significantly correlated with the actual ones, with a very high correlation coefficient (p<0.001; r>0.90). Bias and 95% limits of agreement were acceptable because, whatever the nomogram, they were less than or equal to -0.0±6.2% on the 3000-m, 0.0±3.7% on the 5000-m, and 0.1±9.3% on the 10,000-m. Conclusion: The study confirms the validity of the three nomograms to predict track-running performance with a high level of accuracy. The predictions from these nomograms are similar and may be used in training programs and competitions.</p>

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Original investigation

- 1
- 2
- 3 **1**
- 4
- 5 **2**
- 6
- 7 **3** **Prediction of middle-distance running performances of female runners using nomograms**
- 8
- 9 **4**
- 10
- 11 **5** **Running head:** Prediction of middle-distance running performances in women
- 12
- 13 **6**
- 14
- 15
- 16 **7** **Abstract word count:** 200 words
- 17
- 18 **8** **Text-only word count:** 3695 words (without references)
- 19
- 20 **9** **Number of tables/figures:** 2 tables and 4 figures
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**10 Abstract**

11 Purpose: This study examined the validity, precision and accuracy of the predictions of middle-
12 distance performances in female runners from three nomograms. Methods: Official rankings of
13 French women for the 3000-m, 5000-m, and 10,000-m track-running events from 2005 to 2019
14 were examined. Only female runners who performed in the three distance events within the same
15 year were included (n=158). Each performance over any distance was predicted using the three
16 nomograms from the two other performances. Results: The 3000-m, 5000-m and 10,000-m
17 performances were 11min17 s ± 1min20 s, 19min29 s ± 2min20 s, 41min18 s ± 5min7 s,
18 respectively. No difference was found between the actual and predicted running performances
19 regardless of the nomogram (p>0.05). All predicted running performances were significantly
20 correlated with the actual ones, with a very high correlation coefficient (p<0.001; r>0.90). Bias and
21 95% limits of agreement were acceptable because, whatever the nomogram, they were less than or
22 equal to -0.0±6.2% on the 3000-m, 0.0±3.7% on the 5000-m, and 0.1±9.3% on the 10,000-m.
23 Conclusion: The study confirms the validity of the three nomograms to predict track-running
24 performance with a high level of accuracy. The predictions from these nomograms are similar and
25 may be used in training programs and competitions.

26 Keywords: estimation, modelling, time to exhaustion, validity, female athletes.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**27 Introduction**

28

29 The ability to predict track-running performance can be of great interest for athletes and coaches.
30 It can help detect the best athletes to form a team, identify a specialty, prescribe speed during tempo
31 runs, and even determine competitive strategies during a race [1–3]. The relationships between
32 running distance (or speed) and time has long been used for this purpose [4–6]. Several studies
33 have investigated the prediction of track-running performances, but mainly in male runners [1–3,
34 7, 8], especially elite men [8–10], with power laws [3, 8, 10], critical speed [3, 7, 8, 10, 11] and
35 nomograms [1, 2, 7, 9]. However, the studies on predicting the track-running performances of
36 female athletes, particularly non-elite runners (*e.g.*, modest level), are scarcer.

37

38 Since the beginning of the 20th century, many authors have modeled world running records in order
39 to predict running performance [4]. The empirical and experimental analyses of these data suggest
40 that predicting running performance for one distance from the recent performances over two shorter
41 and/or longer distances is possible using simple practical methods. The nomograms of Mercier et
42 al. [12] and Vandewalle [9] are notable examples.

31 43

44 The nomogram of Mercier et al. [12] is an empirical tool that has numerous advantages as it can
45 be used to determine a runner's event specialty [12], estimate maximal oxygen uptake and
46 endurance capacity [13], and predict running performance. Indeed, Coquart et al. [14] showed that
47 this nomogram provides valid predictions of long-distance road-running performances (*i.e.*, 10-
48 km, 20-km and marathon) and further studies confirmed this finding for middle-distance track
49 performances (*i.e.*, 3000-m, 5000-m and 10,000-m events) [1, 2, 7]. The nomogram of Mercier et
50 al. [12] is thus considered a valid tool to predict track-running performance with accuracy in both
51 male and female runners.

46 52

53 More recently, Vandewalle [9] developed nomograms based on the logarithmic models of Péronnet
54 and Thibault [9, 15]. Two of them are two-dimensional diagrams designed to graphically obtain a
55 function without using a mathematical formula, and they can be used to estimate maximal aerobic
56 speed and endurance capacity and predict the individual performances of elite runners on a track
57 (1500-m, 3000-m, 5000-m and 10,000-m) [9]. However, although Vandewalle's two nomograms

3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86

are similarly constructed, the number of distances (*i.e.*, small and large range of distances included for the first and second nomograms, respectively) and the number of graduations used to predict the performances (for the same distance) vary and are more or less precise. For example, Vandewalle's first nomogram [9] contains graduations of 5 s, 10 s and 15 s, respectively, for 3000-m, 5000-m, and 10,000-m, whereas the second [9] includes variable graduations of 10 s (when the performance is between 7 and 14 min) or 20 s (when the performance is between 14 and 28 min) for 3000-m and 5000-m, graduations of 30 s (when the performance is between 28 and 55 min) for 5000-m and 10,000-m, and graduations of 1 min for 10,000-m (when the performance is between 55 and 97 min). Although Vandewalle [9] confirmed the validity of these nomograms, they were only tested using the performances of the best male runners in the world. Thus, their validity [9] remains to be confirmed for the track performances of non-elite female runners.

The purpose of the current study was therefore to test the validity and compare the accuracy (*i.e.*, the nearness of the actual performance to the predicted performance) and precision (*i.e.*, the closeness of the predicted performances) of three nomograms (*i.e.*, the nomogram of Mercier et al. [12] and two nomograms of Vandewalle [9]) in predicting the middle-distance performances of French female runners over the last 14 years.

Methods

1. Experimental approach

All official rankings of the French Athletics Federation (FFA: *Fédération Française d'Athlétisme*) for female runners between 23 and 39 years old (*i.e.*, age range for the "senior/adult" category in France) competing in the 3000-m, 5000-m and 10,000-m track events were scrutinized from 1 January 2005 to 31 December 2019. All performances were collected regardless of the performance level (*i.e.*, subregional, regional, interregional, national and international). Performances of female runners who competed over the three distances within a 12-month period were recorded. On-road performances in events like the 10-km were excluded, and only track performances were kept to control for any effect of the race profile (flat or hilly) on performance. Once the athletes had been identified, the actual velocity for each distance was calculated individually, and when the velocity was higher for a longer distance than for another shorter distance, the athlete was excluded. Such

1
2
3 87 a result may have been the consequence of a poor strategy, nonspecific preparation for bad weather
4
5 88 conditions during the competition, or an injury in the shorter distances.
6

7
8 89 This study was approved by the National Ethics Committee for Research in Sports Sciences
9
10 90 (CERSTAPS 2019-22-02-31). Moreover, the study protocol was legally declared in accordance
11
12 91 with the European General Data Protection Regulations.

13 14 92 2. Participants

15
16 93 The performances of 158 women were taken into account. The 3000-m, 5000-m and 10,000-m
17
18 94 performances were 11 min 17 s \pm 1 min 20 s (range: 8 min 53 s–15 min 56 s), 19 min 29 s \pm 2 min
19
20 95 20 s (range: 15 min 8 s–27 min 53 s) and 41 min 18 s \pm 5 min 7 s (range: 31 min 53 s–59 min 14
21
22 96 s), respectively (Table 1). These performances in the 3000-m, 5000-m and 10,000-m represented,
23
24 97 respectively, 60 \pm 17%, 62 \pm 17% and 59 \pm 18% of the current outdoor world records (*i.e.*, 8 min
25
26 98 6 s by Junxia Wang in 1993, 14 min 11 s by Tirunesh Dibaba in 2008, and 29 min 17 s by Almaz
27
28 99 Ayana in 2016).

29 30 100 3. Procedures & data treatment

31
32 101 For all middle-distance running (*i.e.*, 3000-m, 5000-m and 10,000-m), performances
33
34 102 were predicted individually using the nomogram of Mercier et al. [12] and the two nomograms of
35
36 103 Vandewalle [9].

37
38 104 To do so, a numerical procedure was implemented using custom-made MATLAB code (MATLAB
39
40 105 R2016b, The MathWorks Inc., Natick, MA, USA) to automatically predict the required
41
42 106 performance from each nomogram and avoid human errors. Intraclass correlation coefficients
43
44 107 (ICC) and the 95% confidence interval (95% CI) were assessed to confirm the internal consistency
45
46 108 of this method [16]. The results are presented in Table 2 and permit the conclusion that this
47
48 109 automated method is valid.

49 110 For example, for the nomogram of Mercier et al. [12], the MATLAB code loads a two-column
50
51 111 matrix for each prediction with the actual athletes' performances and digitizes the classic
52
53 112 nomogram, adding a coordinate system to the lower left corner of the image (*e.g.*, Step 1, Figure
54
55 113 1, for example). Then, the user clicks on multiple anchoring points on the vertical axis of all chosen
56
57 114 distances. This procedure is performed only once, and the results (coordinates of points and

5

1
2
3 115 corresponding times on the nomogram) are stored for all future predictions. Actual running
4
5 116 performances are then transformed into accurate coordinates that are computed from the closest
6
7 117 anchoring points (Step 2, Figure 1). The two coordinates are then used to compute the slope (a)
8
9 118 and intercept (b) of the line that passes through the actual running performances (Step 3, Figure 1).
10
11 119 As the x coordinate is known for all anchoring points, the y coordinate of the predicted running
12
13 120 performance is computed as $y = ax + b$. This final coordinate corresponds accurately to the
14
15 121 intersection point between the line that passes through the actual running performance and the
16
17 122 vertical axis of the desired distance. The intersection point (expressed in percentage of the distance
18
19 123 separating the coordinates of the two closest anchoring points) is finally transformed into the
20
21 124 predicted running performance. This process was automatized since the MATLAB code
22
23 125 implemented each line of the two-column matrix with the actual 158 performances. Automatization
24
25 126 avoids the errors that are possible during the manual process, as presented in Coquart et al. [1]. A
26
27 127 similar procedure was followed for the nomograms of Vandewalle [9, 15], with the exception that
28
29 128 the anchoring points had different x coordinates for each running distance.

30 129 31 130 4. Statistical Analysis

32
33 131 Standard statistical methods were used to calculate the means and standard deviations (SD) of the
34
35 132 performances. Normal Gaussian distribution was verified by the Shapiro-Wilk test. As the data did
36
37 133 not pass the test for normality, a Wilcoxon signed-rank test was used to compare actual and
38
39 134 predicted performances for the 3000-m, 5000-m and 10,000-m. The magnitude of the differences
40
41 135 was assessed by effect size (ES) and classified using the Cohen scale [17]: trivial for $ES < 0.2$,
42
43 136 small for $0.2 \leq ES < 0.5$, moderate for $0.5 \leq ES < 0.8$, and large for $ES \geq 0.8$. The association
44
45 137 between the actual and predicted running performances was tested with the Spearman rank order
46
47 138 correlation. We considered a coefficient of correlation $r \geq 0.90$ as very high, $0.70 \leq r \leq 0.89$ as
48
49 139 high, $0.50 \leq r \leq 0.69$ as moderate, and $0.26 \leq r \leq 0.49$ as low [18].

50
51 140 The bias, which assessed the accuracy of each nomogram in predicting the actual running
52
53 141 performances, and the 95% limits of agreement (95% LoA), which assessed the precision of the
54
55 142 three nomograms, were computed according to the method of Bland and Altman [19, 20]. This
56
57 143 method requires the calculation of the mean difference between two methods of measurement (the
58
59 144 bias) and $\pm 1.96 \times SD$ (the LoA).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

145 Statistical significance was set at $p < 0.05$, and all analyses were performed with the Statistical
146 Package for the Social Sciences (version 21.0, Chicago, USA).

147 **Results**

148 Means and SD of actual and predicted running performances for the 158 female runners according
149 to each nomogram are presented in Table 1.

150 No significant difference was found between the actual and predicted running performances
151 whatever the nomogram ($p > 0.05$, Table 1). All predicted performances were significantly
152 correlated with the actual ones, with a very high correlation coefficient ($p < 0.001$, $r \geq 0.94$, Table
153 1 and Figures 2–4).

154 The bias and 95% LoA are shown in Figures 2–4. The differences between actual and predicted
155 performances of the nomogram of Mercier et al. [12] were $0 \text{ s} \pm 21 \text{ s}$, $0 \text{ s} \pm 22 \text{ s}$, and $-2 \text{ s} \pm 1 \text{ min}$
156 51 s for the 3000-m, 5000-m and 10,000-m, respectively, and the lower and upper 95% LoA were
157 -42 s and 42 s , -42 s and 43 s , and $-3 \text{ min } 40 \text{ s}$ and $3 \text{ min } 36 \text{ s}$ on each distance, respectively (Table
158 1 and Figure 2). The differences between the actual and predicted performances of the first
159 nomogram of Vandewalle [9] (*i.e.*, the nomogram with the most accurate graduations) were only
160 $1 \text{ s} \pm 21 \text{ s}$, $-1 \text{ s} \pm 22 \text{ s}$, and $4 \text{ s} \pm 1 \text{ min } 58 \text{ s}$ for the 3000-m, 5000-m and 10,000-m, respectively,
161 and the lower and upper 95% LoA were -40 s and 41 s , -44 s and 41 s , and $-3 \text{ min } 46 \text{ s}$ and 3 min
162 55 s on each distance, respectively (Table 1 and Figure 3). The differences between the actual and
163 predicted performances of the second nomogram of Vandewalle [9] (*i.e.*, the one with the least
164 accurate graduations) were only $0 \text{ s} \pm 21 \text{ s}$, $0 \text{ s} \pm 22 \text{ s}$, and $2 \text{ s} \pm 1 \text{ min } 57 \text{ s}$ for the 3000-m, 5000-
165 m and 10000-m, respectively, and the lower and upper 95% LoA were -40 s and 41 s , -43 s and 43
166 s , and $-3 \text{ min } 48 \text{ s}$ and $3 \text{ min } 51 \text{ s}$ for each distance, respectively (Table 1 and Figure 4).

167 The bias and 95% LoA were very slightly lower with the nomogram of Mercier et al. [12] compared
168 to the nomograms of Vandewalle [9] for the 10,000-m, whereas the first nomogram of Vandewalle
169 [9] was very slightly more precise (*i.e.*, a lower distance between limits of agreement) compared
170 to the other nomograms in predicting performances for the 3000-m and 5000-m events (Table 1).
171 Moreover, a higher prediction precision (observed from a narrower 95% LoA) was noted for the

7

1
2
3 172 5000-m, whatever the nomogram (Table 1). The magnitude of the bias in these predicted
4
5 173 performances was systematically trivial ($ES \leq 0.013$; Table 1).
6
7

8 174 **Discussion**

9
10
11 175 The purpose of the current study was to test the validity and compare the accuracy (from the bias)
12 176 and precision (from 95% LoA) of the predictions of middle-distance running performances for
13 177 female runners using the nomogram of Mercier et al. [12] and the two nomograms of Vandewalle
14 178 [9]. Our main finding was that these three nomograms could be considered valid, accurate and
15 179 precise in predicting female athletes' performances. Indeed, in accordance with previous studies
16 180 [1, 2, 7, 9], all the results of the current study confirm the validity of the three nomograms to predict
17 181 the middle-distance track-running performances of women from their performances in two other
18 182 distances. The results revealed that in 95% of new predictions, the differences between the actual
19 183 and predicted performances of Mercier's nomogram [12], Vandewalle's first nomogram [9] (*i.e.*,
20 184 the nomogram with a small range of distances and the most accurate graduations) and Vandewalle's
21 185 second nomogram [9] (*i.e.*, the one with a large range of distances and the least accurate
22 186 graduations), respectively, would be less than or equal to $-0.0 \pm 6.2\%$, $0.1 \pm 6.0\%$ and $0.1 \pm 6.0\%$
23 187 for the 3000-m, $0.0 \pm 3.7\%$, $-0.1 \pm 3.6\%$ and $0.0 \pm 3.7\%$ for the 5000-m, and $-0.1 \pm 8.9\%$, $0.2 \pm$
24 188 9.3% and $0.1 \pm 9.3\%$ for the 10,000-m.
25
26
27
28
29
30
31
32
33
34
35

36 189 However, although the precision measurements (*i.e.*, 95% LoA values) of the predictions were
37 190 close between the three nomograms, we noted slight tendencies in the 95% LoA. Indeed, it seems
38 191 that Vandewalle's first nomogram [9] was slightly more precise (*i.e.*, a lower distance between
39 192 95% LoA) compared to the other two in predicting performances for the 3000-m and 5000-m,
40 193 whereas for the 10,000-m event, the nomogram of Mercier et al. was more precise [12]. These
41 194 findings could be explained by differences in design (*i.e.*, models of construction) between the
42 195 nomograms. The nomogram of Mercier et al. [12] was developed from the authors' empirical
43 196 knowledge, whereas the nomograms of Vandewalle were based on a logarithmic model [9].
44 197 Moreover, although Vandewalle's nomograms were constructed in a similar manner and based on
45 198 the same model of Péronnet and Thibault [9], the distances and graduations to predict the
46 199 performances varied (*i.e.*, distances that are the most often performed in international competitions
47 200 for Vandewalle's first nomogram [9]: 1500-m, 3000-m, 5000-m and 10,000-m vs a large range of
48
49
50
51
52
53
54
55
56
57

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

201 distances for the second nomogram [9]: 1000-m to 20000-m; and graduations of 5 s, 10 s and 15 s
202 for 3000-m, 5000-m and 10,000-m, respectively, for the first nomogram [9] vs graduations of 10 s,
203 20 s, 30 s and 1 min for 7–14 min (1000-m–5000-m), 14–28 min (2000-m–10,000-m), 28–55 min
204 (4000-m–20,000-m) and 55–97 min (10,000-m–20,000-m), respectively, for the second one [9]).
205 These variations might explain some of the tendencies in the precisions of the prediction results.
206 From the same perspective and given the tendencies obtained over 10,000-m, it is likely that the
207 scale to predict performance defined by Mercier et al. [12] (a scale that varied with time for a given
208 distance with, for example, a graduation of 15 s from 27 min to 31 min for 10,000-m, a graduation
209 of 20 s from 31 min to 38 min, etc.) is closer to the actual performance for the 10,000-m event than
210 those predicted by Vandewalle's two nomograms [9].

211 On the other hand, because of better correlations (*i.e.*, $r \geq 0.987$) and lower 95% LoA for the 5000-
212 m (*i.e.*, 95% LoA vary only ± 3.6 – 3.7%), it may be suggested that the prediction was better for this
213 distance (whatever the nomogram) compared to the 3000-m (*i.e.*, $0.962 \leq r \leq 0.967$; 95% LoA vary
214 ± 6.0 – 6.2%) and 10,000-m (*i.e.*, $0.706 \leq r \leq 0.895$; 95% LoA vary ± 8.9 – 9.3%). This result, for
215 example, seems to be in accordance with the study by Coquart and Bosquet [7], which showed that
216 the 95% LoA varied by $\pm 6.4\%$, $\pm 3.5\%$, and $\pm 10.4\%$ of the means of the actual and predicted
217 performances for the 3000-m, 5000-m, and 10,000-m performances, respectively, with the
218 nomogram of Mercier et al. [11]. One explanation for this result may be that predictions are more
219 accurate and precise when the predicted distance is made by interpolation (*i.e.*, when the distance
220 is bounded by shorter and longer distances; a predicted performance for the 5000-m from actual
221 performances for the 3-000 m and 10,000-m in the current study), as recommended by Mercier et
222 al. [12], instead of by extrapolation. Indeed, physiologically, running performances can be modeled
223 on the basis of parameters like maximal oxygen uptake (main predictor of performance in short
224 middle-distance races, such as the 3000-m), running economy and endurance capacity (which, as
225 a predictor, has a larger and larger share as the race distance increases, thus for long middle-
226 distance races like the 10,000-m) [21, 22]. Predicting the performance of intermediate distance
227 races by interpolation (like the 5000-m) rather than by extrapolation makes it easier to take maximal
228 oxygen uptake and endurance capacity into account and thus achieve more accurate and precise
229 predictions. Therefore, these results support this recommendation, which has already been
230 confirmed by other studies on improving prediction accuracy and precision [1, 2, 7, 10, 14].

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

231 In addition, data automatization using MATLAB code seems to have prevented more errors than
232 the manual process, as described in Coquart et al. [1]. The biases ($< 1\%$) and 95% LoA ($\leq 9.3\%$)
233 of the current study were slightly better than those of previous studies [1, 2], suggesting that the
234 automated MATLAB method was able to avoid human error. However, we note that the works of
235 Coquart et al. [1] and Tabben et al. [2] studied only male runners and took into account the specialty
236 and performance level, respectively, to validate and compare the precision and accuracy of the
237 performances predicted by Mercier's nomogram [12]. This might explain some of the differences
238 in the results and interpretations, given that Mercier et al. [12] suggested better predictions, for
239 example, when runners were specifically prepared for the distances from which the predictions
240 were made. This hypothesis tested by Coquart et al. [1] was nevertheless not confirmed. We also
241 note that the current study focused on a large sample of women's performances (whatever their
242 performance level), which have not been much studied until now, except by Coquart et al. [1] using
243 the nomogram of Mercier et al. [12] only. Moreover, as noted in the study by Lerebourg and
244 Coquart [23], women's performances are often less studied than men's and a deeper understanding
245 their performances might be of great interest for athletes and coaches.

246 The main limitation of this study was that the actual running performances were not executed in
247 similar environmental conditions (*e.g.*, same climate conditions and altitude), with identical
248 physical (*e.g.*, endurance capacity, running economy, running speed) and psychological (*e.g.*,
249 motivation and stress) conditions, and using the same opposition field and/or race strategy. Indeed,
250 as noted, we limited our analysis to track performances in order to control for the effect of race
251 profile, but it is possible that the above variables may have negatively affected the accuracy of the
252 predictions. Despite this potential limitation, however, we found that the predicted performances
253 using all three nomograms were valid and accurate. It is very interesting to note that these
254 nomograms and more particularly the nomogram of Mercier et al. [12], which was designed in the
255 1980s, are still to this day valid tools for predicting actual performances with an awareness of the
256 evolution of these performances [23].

257 One of the purposes of this study was to analyze all the performances of French female middle-
258 distance runners regardless of their performance level (*i.e.*, subregional, regional, interregional,
259 national and international) and therefore athletes of more modest levels were included. Although
260 it has been shown that Vandewalle's nomograms can predict the performances of elite track runners

10

1
2
3 261 (1500-m, 3000-m, 5000-m and 10,000-m) [9], the current study confirms these results in non-elite
4
5 262 runners. The performances in the present study could have been classified (for example) into four
6
7 263 levels: subregional, regional, national and international. However, had we done so, the sample sizes
8
9 264 would have been too weak for some groups ($n = 1, 1$ and 4 for international-level athletes,
10
11 265 respectively, for the 3000-m, 5000-m and 10,000-m events). Further studies, especially for women,
12
13 266 are needed to test whether the predictions are better when the performance level is high.

14
15 267 Last, due to the scarcity of studies on women and/or non-elite runners, comparisons of the predicted
16
17 268 performances for women might be made using nomograms with other predictive methods, such as
18
19 269 critical speed, the power law or even methods based on artificial intelligence, like neural networks.
20
21 270 Similarly, from a training or performance perspective, future studies might also be conducted on
22
23 271 the prediction of long-distance on-road performances (10-km, half-marathon and marathon) for
24
25 272 women, as it seems that only male performances have been studied over long distances with the
26
27 273 nomogram of Mercier et al. [14].

28
29
30
31 274

32 275 **Practical Applications**

33
34 276 Athletes and coaches need field measures that can provide them with accurate estimates of
35
36 277 performance. Our results showed that three nomograms are valid and accurate to predict track-
37
38 278 running performance from the performances of two other distances. Having an accurate idea of an
39
40 279 athlete's performance capability at a specific distance facilitates the detection of the best athletes
41
42 280 to form a team, the identification of a specialty, the prescription speed for tempo runs, and the
43
44 281 determination of competitive strategies during races. However, the prediction accuracy from the
45
46 282 nomograms is greater when interpolation is used rather than extrapolation.

47 283 Furthermore, nomograms can be used to estimate physical fitness indicators like maximal oxygen
48
49 284 uptake [13], aerobic speed [9] and capacity [24], and the index of endurance [9], which theoretically
50
51 285 would contribute to optimizing and adapting training programs for athletes and coaches.

52 53 286 **Conclusion**

54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

287 The results of the present study indicate the validity of the nomogram of Mercier et al. [12] and
288 Vandewalle's two nomograms [9]) to predict with high accuracy the individual track-running
289 performances of female runners for the 3000-m, 5000-m, and 10,000-m events from the
290 performances of two other distance events. Results show no significant difference between the
291 actual and predicted performances in female athletes, regardless of the nomogram used. The
292 predictions from these nomograms are similar and may be used by athletes or coaches.

293 References

- 294 [1] Coquart JB, Mercier D, Tabben M, et al. Influence of sex and specialty on the prediction
295 of middle-distance running performances using the Mercier et al.'s nomogram. *J Sports Sci* 2015;
296 33: 1124–1131.
297 [2] Tabben M, Bosquet L, Coquart JB. Effect of performance level on the prediction of
298 middle-distance-running performances using a nomogram. *Int J Sports Physiol Perform* 2016;
299 11: 623–626.
300 [3] Blythe DAJ, Király FJ. Prediction and quantification of individual athletic performance of
301 runners. *PLoS One* 2016; 11: e0157257.
302 [4] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. From empirical
303 models to physiological models. *Sports Med Auckl NZ* 1999; 27: 359–379.
304 [5] Ettema JH. Limits of human performance and energy-production. *Int Z Für Angew*
305 *Physiol Einschließlich Arbeitsphysiologie* 1966; 22: 45–54.
306 [6] Scherrer J, Monod H. Le travail musculaire local et la fatigue chez l'homme. *Presse Med*
307 1960; 68: 1717–1717.
308 [7] Coquart J, Bosquet L. Precision in the prediction of middle distance-running
309 performances using either a nomogram or the modeling of the distance-time relationship. *J*
310 *Strength Cond Res* 2010; 24: 2920–2926.
311 [8] Zinoubi B, Vandewalle H, Driss T. Modeling of running performances in humans:
312 comparison of power laws and critical speed. *J Strength Cond Res* 2017; 31: 1859–1867.
313 [9] Vandewalle H. A nomogram of performances in endurance running based on logarithmic
314 model of péronnet-thibault. *Am J Eng Res* 2017; 6: 78–85.
315 [10] Vandewalle H. Modelling of running performances: comparisons of power-law,
316 hyperbolic, logarithmic, and exponential models in elite endurance runners. *BioMed Res Int*
317 2018; 2018: 8203062.
318 [11] Gamelin FX, Coquart JM, Ferrari N, et al. Prediction of one-hour running performance
319 using constant duration tests. *J Strength Cond Res* 2006; 20: 735–739.
320 [12] Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le
321 VO₂max et l'endurance relative en course de fond. *Méd Sport* 1984; 58: 181–187.
322 [13] Tokmakidis SP, Léger L, Mercier D, et al. New approaches to predict VO₂max and
323 endurance from running performances. *J Sports Med Phys Fitness* 1987; 27: 401–409.
324 [14] Coquart J, Alberty M, Bosquet L. Validity of a nomogram to predict long distance
325 running performance. *J Strength Cond Res* 2009; 23: 2119–2123.
326 [15] Péronnet F, Thibault G. Mathematical analysis of running performance and world running
327 records. *J Appl Physiol Bethesda Md* 1985 1989; 67: 453–465.

12

- 1
2
3 328 [16] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients
4 329 for reliability research. *J Chiropr Med* 2016; 15: 155–163.
5 330 [17] Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1: 98–101.
6 331 [18] Munro BH. Correlation. *Stat Methods Health Care Res* 1997; 224–245.
7 332 [19] Bland JM, Altman D. Statistical methods for assessing agreement between two methods
8 333 of clinical measurement. *The lancet* 1986; 327: 307–310.
9 334 [20] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat*
10 335 *Methods Med Res* 1999; 8: 135–160.
11 336 [21] di Prampero PE, Atchou G, Brückner JC, et al. The energetics of endurance running. *Eur*
12 337 *J Appl Physiol* 1986; 55: 259–266.
13 338 [22] Péronnet F, Thibault G. [Physiological analysis of running performance: revision of the
14 339 hyperbolic model]. *J Physiol (Paris)* 1987; 82: 52–60.
15 340 [23] Lerebourg L, Coquart JB. Changes in performances/characteristics of French female
16 341 runners over the last 12 years. *Res Sports Med* 2020; 1–11.
17 342 [24] Bosquet L, Léger L, Legros P. Methods to determine aerobic endurance. *Sports Med*
18 343 *Auckl NZ* 2002; 32: 675–700.

345

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Mean values and standard deviation (SD) of actual and predicted performances from the nomogram of Mercier et al. [12] and the two nomograms of Vandewalle [9, 15] for 158 female runners; difference between actual and predicted performances (p), Spearman correlation with actual performance (r), bias and SD (min), bias and 95% LoA (min and %), and magnitude of the bias for the 3000-m, 5000-m and 10,000-m track events.

Note: Bias = difference between actual and predicted performances; r = coefficient of correlation; LoA = limits of agreement; ES = effect size. *Significantly correlated at $p < 0.001$

Table 2. Intra-class correlation coefficient (ICC) computations using single-rating, absolute-agreement, a two-way mixed-effects model between MATLAB code and manual methods from the nomogram of Mercier et al. [12] and the two nomograms of Vandewalle [9, 15] for 158 female runners in the 3000-m, 5000-m and 10,000-m track events.

Figure 1. Example of the automatic process used to predict running performance (illustration with the nomogram of Mercier et al. [12]).

The automatic process to predict 10,000-m performance based on the digitalizing of the nomogram of Mercier et al. [12]. *Step 1:* Digitalization of the nomogram, adjunction of a coordinate system and depiction of all anchoring points. *Step 2:* Close view of the interesting part of the nomogram: integration of the athlete's performances transformed into coordinates by using the closest anchoring points. *Step 3:* Computation of the slope and intercept of the line passing through the performance. The remaining performance (here for the 10,000-m) is predicted using the $y = a.x + b$ equation by transforming final coordinates into the corresponding running performance.

Figure 2. Validity of measurements with the Mercier et al. [12] nomogram to predict performance.

Left panel: association between actual and predicted performances from the Mercier et al. nomogram [12] in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The solid line is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman plots for the comparisons between the actual and predicted performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines are the bias, solid lines are the 95% limits of agreement.

Figure 3. Validity of measurements with the first Vandewalle et al. nomogram [9] to predict performance.

Left panel: association between actual and predicted performances from Vandewalle's first nomogram based on the logarithmic model of Péronnet and Thibault [9, 15] in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The solid line is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman plots for the comparison between the actual and predicted performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines are the bias, solid lines are the 95% limits of agreement.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

387 Figure 4. Validity of measurements with the second Vandewalle et al. nomogram [9] to
388 predict performance.

389
390 *Left panel:* association between actual and predicted performances from the second nomogram of
391 Vandewalle based on logarithmic model of Péronnet and Thibault [9, 15], in the 3,000 m (top),
392 5,000 m (middle), and 10,000 m (bottom) events for 158 female runners. The solid line is the linear
393 regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman plots for the
394 comparison between actual and predicted performances in the 3000-m (top), 5000-m (middle), and
395 10,000-m (bottom) events for 158 female runners. Dashed line is the bias, solid lines are the 95%
396 limits of agreement.

For Peer Review

Table 1. Mean values and standard deviation (SD) of actual and predicted performances from the nomogram of Mercier et al. [12] and the two nomograms of Vandewalle [9, 15] for 158 female runners; difference between actual and predicted performances (p), Spearman correlation with actual performance (r), bias and 95% LoA (min and %), and magnitude of the bias for the 3000-m, 5000-m and 10,000-m track events.

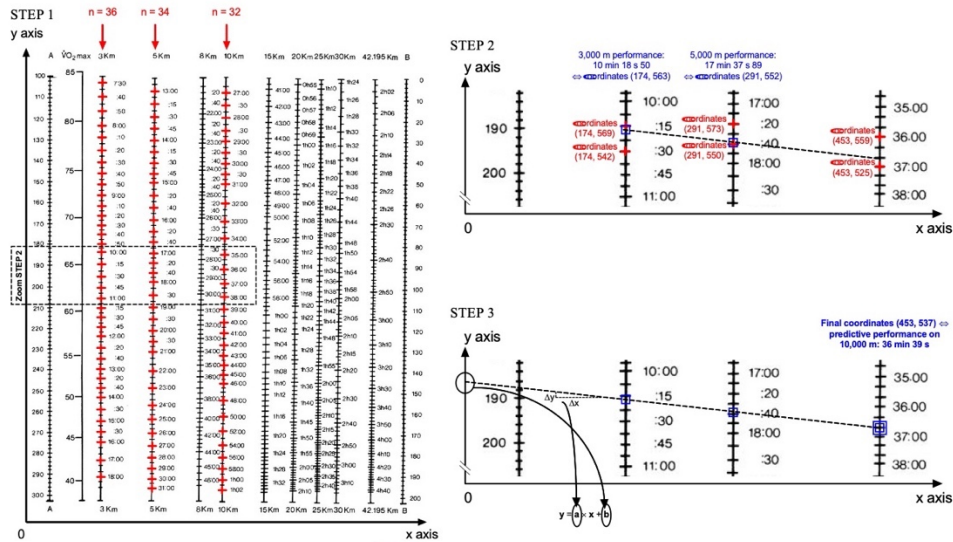
Distance (m)	Mean performance (min) \pm SD	p	Correlation			Bias (min) \pm 95% LoA	Bias (%) \pm 95% LoA	Magnitude of bias	
			r	Interpretation	ES			Interpretation	
3000	Actual	11.28 \pm 1.33							
	Predicted from the nomogram of Mercier	11.28 \pm 1.37	0.850	0.962*	Very high	-0.002 \pm 0.699	-0.0 \pm 6.2	-0.002	Trivial
	Predicted from the first nomogram of Vandewalle	11.29 \pm 1.37	0.888	0.967*	Very high	0.009 \pm 0.677	0.1 \pm 6.0	0.007	Trivial
	Predicted from the second nomogram of Vandewalle	11.28 \pm 1.37	0.897	0.966*	Very high	0.007 \pm 0.678	0.1 \pm 6.0	0.005	Trivial
5000	Actual	19.49 \pm 2.34							
	Predicted from the nomogram of Mercier	19.50 \pm 2.32	0.996	0.987*	Very high	0.005 \pm 0.712	0.0 \pm 3.7	0.002	Trivial
	Predicted from the first nomogram of Vandewalle	19.47 \pm 2.32	0.387	0.987*	Very high	-0.020 \pm 0.708	-0.1 \pm 3.6	-0.009	Trivial
	Predicted from the second nomogram of Vandewalle	19.49 \pm 2.32	0.951	0.987*	Very high	-0.000 \pm 0.716	0.0 \pm 3.7	-0.000	Trivial
10,000	Actual	41.03 \pm 5.12							
	Predicted from the nomogram of Mercier	41.01 \pm 5.25	0.837	0.946*	Very high	-0.026 \pm 3.634	-0.1 \pm 8.9	-0.005	Trivial
	Predicted from the first nomogram of Vandewalle	41.10 \pm 5.34	0.706	0.942*	Very high	0.070 \pm 3.842	0.2 \pm 9.3	0.013	Trivial
	Predicted from the second nomogram of Vandewalle	41.06 \pm 5.30	0.895	0.941*	Very high	0.026 \pm 3.828	0.1 \pm 9.3	0.005	Trivial

Note: Bias = difference between actual and predicted performances; r = coefficient of correlation; LoA = limits of agreement; ES = effect size. *Significantly correlated at $p < 0.001$

Table 2. Intra-class correlation coefficient (ICC) computations using single-rating, absolute-agreement, a two-way mixed-effects model between MATLAB code and manual methods from the nomogram of Mercier et al. [12] and the two nomograms of Vandewalle [9, 15] for 158 female runners in the 3000-m, 5000-m and 10,000-m track events.

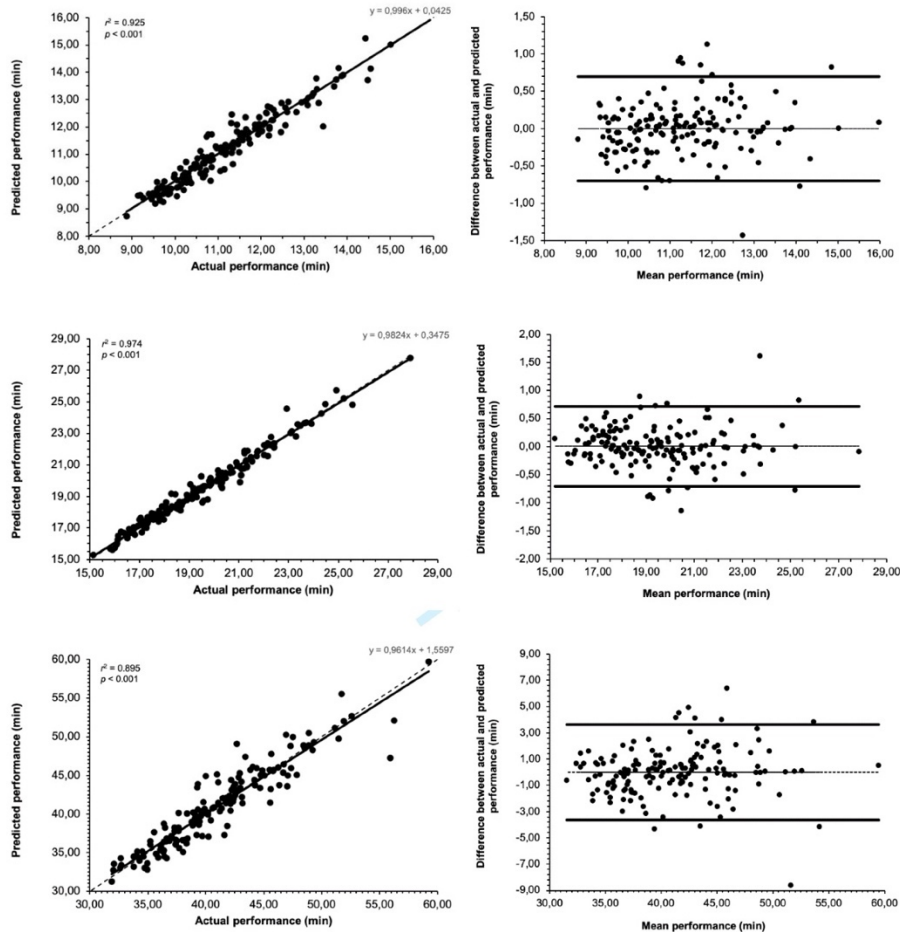
Distance (m)	Intra-class correlation coefficients	Interpretation of reliability [16]	95% Confidence Interval		
			Lower Bound	Upper Bound	
3000	Predicted from the nomogram of Mercier	0.991	Excellent	0.987	0.993
	Predicted from the first nomogram of Vandewalle	0.998	Excellent	0.997	0.998
	Predicted from the second nomogram of Vandewalle	0.876	Good	0.834	0.908
5000	Predicted from the nomogram of Mercier	0.944	Excellent	0.925	0.959
	Predicted from the first nomogram of Vandewalle	0.750	Good	0.673	0.811
	Predicted from the second nomogram of Vandewalle	0.726	Moderate	0.643	0.792
10,000	Predicted from the nomogram of Mercier	0.998	Excellent	0.997	0.999
	Predicted from the first nomogram of Vandewalle	0.820	Good	0.761	0.865
	Predicted from the second nomogram of Vandewalle	0.866	Good	0.821	0.901

Figure 1. Example of the automatic process used to predict running performance (illustration with the nomogram of Mercier et al. [12]).



The automatic process to predict 10,000-m performance based on the digitalizing of the nomogram of Mercier et al. [12]. *Step 1:* Digitalization of the nomogram, adjunction of a coordinate system and depiction of all anchoring points. *Step 2:* Close view of the interesting part of the nomogram: integration of the athlete's performances transformed into coordinates by using the closest anchoring points. *Step 3:* Computation of the slope and intercept of the line passing through the performance. The remaining performance (here for the 10,000-m) is predicted using the $y = a.x + b$ equation by transforming final coordinates into the corresponding running performance.

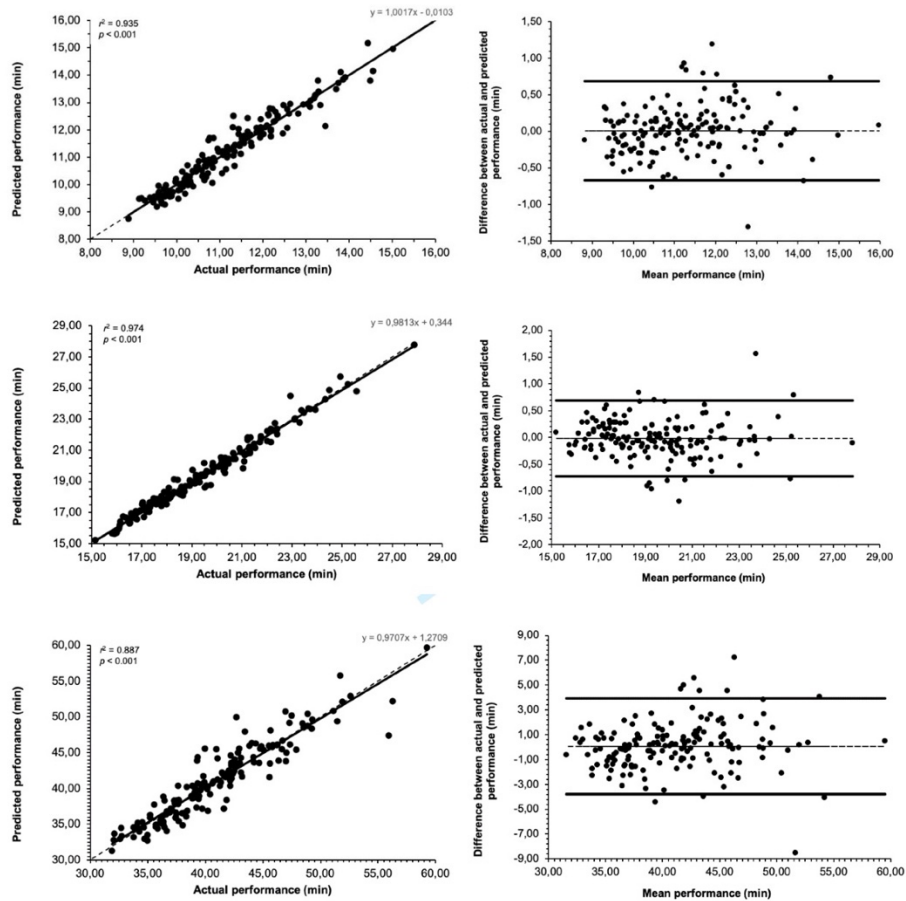
Figure 2. Validity of measurements with the Mercier et al. [12] nomogram to predict performance.



Left panel: association between actual and predicted performances from the Mercier et al. nomogram [12] in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The solid line is the linear regression. r^2 is the coefficient of determination.

Right panel: Bland and Altman plots for the comparisons between the actual and predicted performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines are the bias, solid lines are the 95% limits of agreement.

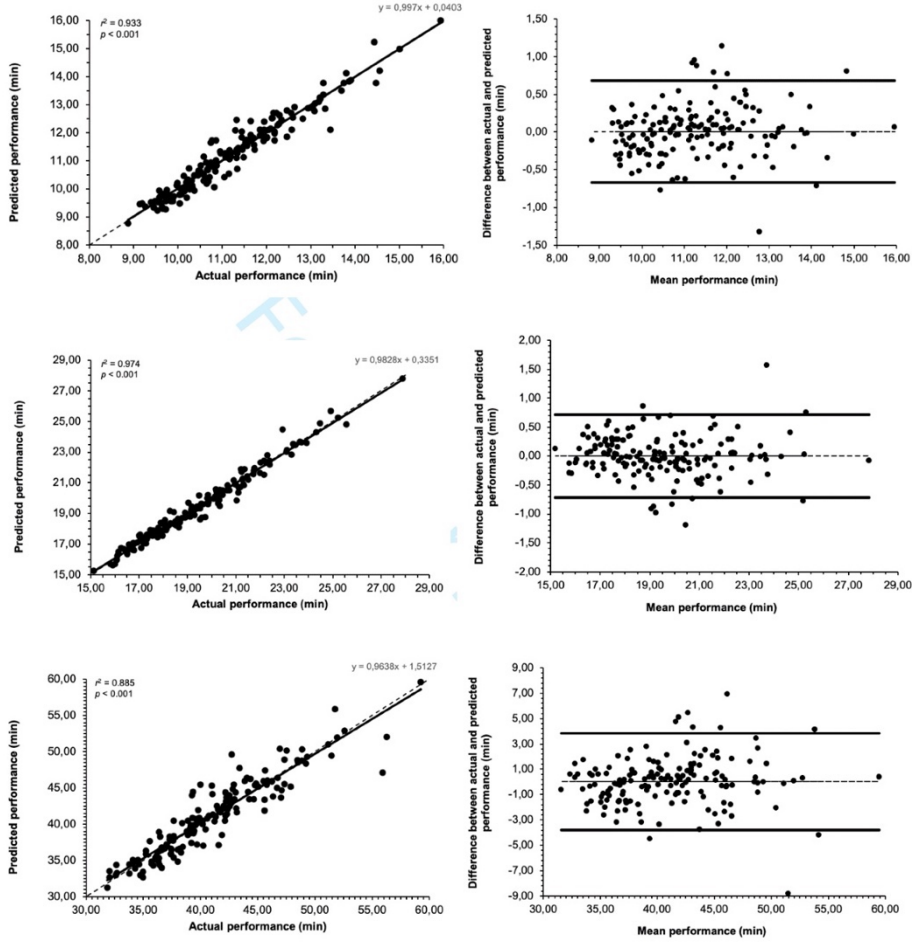
Figure 3. Validity of measurements with the first Vandewalle et al. nomogram [9] to predict performance.



Left panel: association between actual and predicted performances from Vandewalle's first nomogram based on the logarithmic model of Péronnet and Thibault [9, 15] in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The solid line is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman plots for the comparison between the actual and predicted performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines are the bias, solid lines are the 95% limits of agreement.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 4. Validity of measurements with the second Vandewalle et al. nomogram [9] to predict performance.



Left panel: association between actual and predicted performances from the second nomogram of Vandewalle based on logarithmic model of Péronnet and Thibault [9, 15], in the 3,000 m (top), 5,000 m (middle), and 10,000 m (bottom) events for 158 female runners. The solid line is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman plots for the comparison between actual and predicted performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed line is the bias, solid lines are the 95% limits of agreement.

Annexe 4. Article 4 : *Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram*

Journal of Sports Sciences

Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram
--Manuscript Draft--

Full Title:	Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram
Manuscript Number:	
Article Type:	Original Manuscript
Keywords:	estimation, modeling, time to exhaustion, validity, women athletes, track-and-field.
Abstract:	<p>Purpose: This study examined the validity and compared the precision and accuracy of a distance-time linear model (DTLM), a power law and a nomogram to predict the middle-distance performances of female runners. Methods: Official rankings of French women for the 3000-m, 5000-m, and 10,000-m track-running events from 2005 to 2019 were examined. Each performance was predicted from two other performances using the DTLM, power law or nomogram. Results: Between the actual and predicted performances, DTLM showed a difference ($p < 0.05$) but the power law and nomogram did not ($p > 0.05$). The magnitude of the differences in these predicted performances was small if not trivial. All predicted performances were significantly correlated with the actual ones, with a very high correlation coefficient ($p < 0.001$; $r > 0.90$), except for DTLM in the 3000-m, which showed a high correlation coefficient ($p < 0.001$; $r > 0.895$). Bias and 95% limits of agreement were acceptable because, whatever the method, they were $\leq -3.7 \pm 10.8\%$ on the 3000-m, $1.4 \pm 4.3\%$ on the 5000-m, and $-2.5 \pm 7.4\%$ on the 10,000-m. Conclusion: The study confirms the validity of the three methods to predict track-running performance and suggests that the most accurate and precise model was the nomogram followed by the power law, with the DTLM being the least accurate.</p>
Order of Authors:	<p>Lucie LEREBOURG, Ph.D student</p> <p>Brice Guignard, Senior lecturer</p> <p>Maxime L'Hermette, Senior lecturer</p> <p>Jérémy Coquart, Full Professor</p>

Manuscript - with author details

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Original investigation

3 Predictions of the middle-distance performances of female runners using a distance-time linear model, a power law and a nomogram

6 Running head: Predictions of the middle-distance performances in women runners

8 Authors: Lucie Lerebourg^{1,2}, Brice Guignard¹, Maxime L'Hermette¹, and Jeremy B. Coquart¹

10 Affiliations:

1. *Normandie Univ, UNIROUEN, Centre d'Études des Transformations des Activités Physiques et Sportives Rouen, France.*

2. *Orthodynamica Center, Mathilde Hospital 2, Rouen, France*

15 Corresponding Author:

Lucie Lerebourg, Centre d'Études des Transformations des Activités Physiques et Sportives, Boulevard Siegfried, 76 821 Mont Saint Aignan Cedex, France. Phone: (+33)648157388. E-mail: lucie.lerebourg@orthodynamica.com

20 Abstract word count: 200 words

21 Text-only word count: 3795 words (without references)

22 Number of tables/figures: 2 tables and 3 figures

24 **ABSTRACT**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

25 **Purpose:** This study examined the validity and compared the precision and accuracy of a
26 distance-time linear model (DTLM), a power law and a nomogram to predict the middle-
27 distance performances of female runners. **Methods:** Official rankings of French women for the
28 3000-m, 5000-m, and 10,000-m track-running events from 2005 to 2019 were examined. Each
29 performance was predicted from two other performances using the DTLM, power law or
30 nomogram. **Results:** Between the actual and predicted performances, DTLM showed a
31 difference ($p<0.05$) but the power law and nomogram did not ($p>0.05$). The magnitude of the
32 differences in these predicted performances was small if not trivial. All predicted performances
33 were significantly correlated with the actual ones, with a very high correlation coefficient
34 ($p<0.001$; $r>0.90$), except for DTLM in the 3000-m, which showed a high correlation
35 coefficient ($p<0.001$; $r>0.895$). Bias and 95% limits of agreement were acceptable because,
36 whatever the method, they were $\leq -3.7\pm 10.8\%$ on the 3000-m, $1.4\pm 4.3\%$ on the 5000-m, and
37 $-2.5\pm 7.4\%$ on the 10,000-m. **Conclusion:** The study confirms the validity of the three methods
38 to predict track-running performance and suggests that the most accurate and precise model
39 was the nomogram followed by the power law, with the DTLM being the least accurate.

40 **Keywords:** estimation, modeling, time to exhaustion, validity, women athletes, track-and-field.

41 INTRODUCTION

1
2
3 42 The prediction of track-running performance is useful to detect the best athletes, identify a
4
5 43 specialty, prescribe speeds for tempo runs, and even determine competitive strategies [1–3]. To
6
7 44 do so, several tools have been proposed [3–8]. The concept of critical speed [4–7, 9], the power
8
9 45 law [3, 5, 7, 10, 11] and some nomograms [1, 2, 4, 12] are notable examples. These empirical
10
11 46 or theoretical tools show that it is possible to predict running performance for one event simply
12
13 47 from the running performances in two other events.

14 48 Modeling the relationship between time and distance or speed is a common and easy way to
15
16 49 predict performance [7, 13–15]. Several models based on the concept of critical speed have
17
18 50 been proposed [14–18], but the original model, the distance-time linear model (DTLM), uses
19
20 51 the following equation [19]:

21 52

$$23 \quad D_{lim} = S_{crit} \cdot T_{lim} + ARC \quad (1)$$

24 54

25
26 55 where D_{lim} is the maximal distance (m) that can be run using both aerobic and anaerobic
27
28 56 metabolisms, S_{crit} is the critical speed ($\text{m}\cdot\text{min}^{-1}$) [20] and is interpreted as the maximal rate of
29
30 57 synthesis of the reserves by aerobic metabolism, ARC (*i.e.*, anaerobic running capacity, which
31
32 58 corresponds to the original y-intercept of the model) is the maximal distance (m) that can be
33
34 59 performed with anaerobic metabolism, and T_{lim} is the time to exhaustion (min) for a given
35
36 60 distance (D_{lim}).

37 61 Therefore, to predict performance the final equation is:

38 62

39 63

$$43 \quad T_{lim} = \frac{(D_{lim} - ARC)}{S_{crit}} \quad (2)$$

44
45
46
47 65 However, the relationship between D_{lim} and T_{lim} is not linear, as the slope of this relationship
48
49 66 depends on the range of T_{lim} (*i.e.*, higher for sprint races than for long duration races) [18]. This
50
51 67 relationship can also be modeled from individual power laws. Indeed, many years ago, in 1906,
52
53 68 Kennelly [21] proposed the power law based on the relationship between speed (S), D_{lim} and
54
55 69 T_{lim} in a descriptive model from world running records. In other words, the relationship
56
57 70 between two quantities is such that one is proportional to a fixed power of the other:

58

59 71

$$60 \quad D_{lim} = k \cdot T_{lim}^g \quad (3)$$

61

62

63

64

65

72 where k is a constant that should, in theory, be correlated with the highest running speed reached
73 during a maximal sprint [5, 7, 11, 22]. Exponent g is a dimensionless parameter that can be
74 considered an endurance index (*i.e.*, g is close to 1 when endurance runners have good
75 endurance capacity and 0 when capacity is low; and if g is equal to 1, the relationship of
76 T_{lim} - D_{lim} would be perfectly linear) [5, 7, 11, 22]. When these two parameters are unknown
77 but datasets are available for variables T_{lim} and D_{lim} , non-linear least squares analysis may be
78 performed to determine the values for k and g using repeated iterations through an algorithm
79 that finds the best fit parameters for the given data.

80 Kennelly's [21] model accurately describes the relationships between T_{lim} and D_{lim} or S ,
81 especially in elite endurance runners [5, 7], and has the advantage of being independent on the
82 range of exhaustion times. However, the validity of applying Kennelly's power law model [21]
83 to predict running performances in non-elite runners has remained unverified [22].

84 Other practical tools such as nomograms can also be used to easily estimate running
85 performance. Although the nomogram of Mercier et al. [23] is an empirical tool that has many
86 advantages (*e.g.*, determining a runner's event specialty [23], estimating maximal oxygen
87 uptake and endurance capacity [24]), it is often used to predict running performance without
88 using a mathematical formula. Indeed, Coquart et al. [25] showed that this nomogram provides
89 valid predictions of long-distance road-running performances and further studies confirmed this
90 finding for middle-distance track performances [1, 2, 4]. However, although the nomogram of
91 Mercier et al. [23] is considered a valid tool to accurately predict track-running performance in
92 both male and female runners, it seems necessary to compare it again with other prediction
93 methods, this time using a numerical process (*i.e.*, not a manual method) to avoid potential
94 human errors in prediction. Indeed, in previous studies, the authors manually drew a line to
95 predict one performance from the performances for two other distances [1, 2, 4, 25]. In the
96 current study, this process was automatized with MATLAB code (numerical process), which
97 digitized the nomogram of Mercier et al. [23] and drew anchoring points for the two known
98 performances to estimate the third one using the equation of the affine function.

99 Both athletes and coaches would benefit from the determination of the best model to optimally
100 calibrate training sessions (*e.g.*, prescribe speeds for tempo runs) and/or achieve the best
101 performance (*e.g.*, detect the best athletes, identify specialties, determine competitive
102 strategies) [1–3]. A number of studies have investigated the prediction of track-running
103 performances, but mainly in male runners [1–6], especially elite men [5, 7, 10, 22]. The studies
104 on predicting the track-running performances of non-elite female runners remain quite limited.

105 The purpose of the current study was therefore to test the validity and compare the accuracy
106 (*i.e.*, the nearness of the actual performance to the predicted performance) and the precision
107 (*i.e.*, the closeness of the predicted performances) of three methods (*i.e.*, DTLM, power law and
108 nomogram of Mercier et al. [23]) in predicting the middle-distance performances of female
109 runners regardless of their performance level. We hypothesized that all three models would be
110 valid and that the model dependent on the range of the exhaustion times (*i.e.*, DTLM) would be
111 the least accurate and precise of the three.

112

113 **METHODS**

114 *1. Experimental approach*

115 All official rankings of the French Athletics Federation (FFA: *Fédération Française*
116 *d’Athlétisme*) for female runners between 23 and 39 years old (*i.e.*, age range for the “senior”
117 category in France) competing in the 3000-m, 5000-m and 10,000-m track events between 1
118 January 2005 and 31 December 2019 were scrutinized. All performances were collected
119 regardless of the performance level. Performances of female runners who competed over the
120 three distances within a 12-month period were recorded. On-road performances in events like
121 the 10-km were excluded, and only track performances were kept to control for any effect of
122 the race profile (flat or hilly) on performance. The athlete was excluded if the speed was higher
123 for a longer distance than for another shorter distance.

124 This study was approved by the National Ethics Committee for Research in Sports Sciences
125 (CERSTAPS 2019-22-02-31).

126 *2. Participants*

127 The performances of 158 women were considered. The 3000-m, 5000-m and 10,000-m
128 performances were 11 min 17 s \pm 1 min 20 s (range: 8 min 53 s–15 min 56 s), 19 min 29 s \pm 2
129 min 20 s (range: 15 min 8 s–27 min 53 s) and 41 min 18 s \pm 5 min 7 s (range: 31 min 53 s–59
130 min 14 s), respectively. These performances in the 3000-m, 5000-m and 10,000-m represented,
131 respectively, 60 \pm 17%, 62 \pm 17% and 59 \pm 18% of the current outdoor world records.

132 *3. Procedures & data treatment*

133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165

133 For all middle-distance running events (*i.e.*, 3000-m, 5000-m and 10,000-m), performances
 134 were predicted individually using the DTLM [15], the power law [7, 21] and the nomogram of
 135 Mercier et al. [23].

136 In the first method, the distance–time relationship was modeled using the DTLM [15]. Two
 137 running performances were plotted, equation 2 being used to determine the third performance.

138 For example, a runner covering the 3000-m in 9.46 min (*i.e.*, 9 min 28 s) and the 5000-m in

139 16.11 min (*i.e.*, 16 min 7 s) has a $S_{Crit} = \frac{5000-3000}{16.11-9.46}$ of 151 m.min⁻¹ and an *ARC* of 301 m.

140 Therefore, she should be able to run the 10,000-m in $T_{lim} = \frac{10000-151}{301} = 32.71$ min (*i.e.*, 32
 141 min 43 s).

142

143 In the second method, the power law was computed by non-linear least-square regressions
 144 between the logarithm of T_{lim} and the logarithm of D_{lim} .

145

146 Therefore, if $D_{lim} = k \cdot T_{lim}^g$, the logarithm of D_{lim} is equal to:

147

$$148 \quad \ln(D_{lim}) = \ln(k) + g \ln(T_{lim}) \quad (4)$$

149

150 where $T_{lim} = e^{\ln(T_{lim})}$

$$151 \quad T_{lim} = e^{\frac{\ln(\frac{D_{lim}}{k})}{g}} \quad (5)$$

152 The values of g and k of the individual power law were computed from the values of D_{lim}
 153 corresponding to T_{lim} of the first and second distances run, without considering D_{lim} and T_{lim}
 154 of the third distance run.

155 The values of g and k were generated using MATLAB (MATLAB R2016b, The MathWorks
 156 Inc., Natick, MA, USA). A one-term power series model as given by $y = a \cdot x^b$ was provided
 157 by the Curve Fitting toolbox, and a fit function to fit the power series models to the data was
 158 used (*i.e.*, fit(time,distance,'power1')). The toolbox calculated optimized start points for the
 159 power series models, based on the current data set.

160

161 Finally, to predict performance with the nomogram of Mercier et al. [23], a numerical process
 162 was implemented using another custom-made MATLAB code to avoid human errors. Intra-class
 163 correlation coefficients (ICC) and the 95% confidence interval (95% CI) were assessed to
 164 confirm the internal consistency of this method [26]. The results are presented in Table 2 and

62
63
64
65

165 the validity of this automated method was confirmed. For example, the MATLAB code loads
166 a two-column matrix for each prediction with the athletes' actual performances and digitizes
167 the classic nomogram, adding a coordinate system to the lower left corner of the image. Then,
168 the user clicks on multiple anchoring points on the vertical axis of all chosen distances. This
169 process is performed only once, and the results (coordinates of points and corresponding times
170 on the nomogram) are stored for all future predictions. Actual running performances are then
171 transformed into accurate coordinates that are computed from the closest anchoring points. The
172 two coordinates are then used to compute the slope (a) and intercept (b) of the line that passes
173 through the actual running performances. As the x coordinate is known for all anchoring points,
174 the y coordinate of the predicted running performance is computed as $y = a \cdot x + b$. This final
175 coordinate corresponds accurately to the intersection point between the line that passes through
176 the actual running performance and the vertical axis of the desired distance. The intersection
177 point (expressed in percentage of the distance separating the coordinates of the two closest
178 anchoring points) is finally transformed into the predicted running performance. This process
179 was automatized since the MATLAB code implemented each line of the two-column matrix
180 with the actual 158 performances. Automatization avoids the errors that are possible during the
181 manual process, as reported in Coquart et al. [1].

182

183 4. *Statistical Analysis*

184 Standard statistical methods were used to calculate the means and standard deviations (SD) of
185 the performances. Normal Gaussian distribution was verified by the Shapiro-Wilk test. As the
186 data did not pass the test for normality, a Wilcoxon signed-rank test was used to compare actual
187 and predicted performances for the 3000-m, 5000-m and 10,000-m. The magnitude of the
188 differences was assessed by effect size (ES) and classified using the Cohen scale [27]. The
189 association between the actual and predicted running performances was tested with the
190 Spearman rank order correlation. We considered a coefficient of correlation $r \geq 0.90$ as very
191 high, $0.70 \leq r \leq 0.89$ as high, $0.50 \leq r \leq 0.69$ as moderate, and $0.26 \leq r \leq 0.49$ as low [28].

192 The bias, which assessed the accuracy of each approach in predicting the actual running
193 performances, and the 95% limits of agreement (95% LoA), which assessed the precision of
194 the three approaches, were computed according to the method of Bland and Altman [29]. The
195 95% of CI for the bias of the three predicted methods illustrated the magnitude of the systematic
196 difference.

60
61
62
63
64
65

197 We assumed the model would be valid if the magnitude of the differences, as assessed by ES,
198 was acceptable and the biases and approval limits were accurate and precise. The best model
199 was selected based on the highest accuracy and precision assessed from the method of Bland
200 and Altman [29].

201 Statistical significance was set at $p < 0.05$, and all analyses were performed with Jeffrey's
202 Amazing Statistics Program (JASP Team, version 0.14.1, computer software).

203

204 RESULTS

205 Means and SD of the actual and predicted running performances for each method are presented
206 in Table 1 for the 158 female runners.

207 No significant difference was found between the actual and predicted performances with the
208 power law or the nomogram ($p > 0.05$, Table 1), but this was not the case for the DTLM ($p <$
209 0.01 , Table 1). However, whatever the method and distance, the magnitude of the bias was
210 trivial or small ($ES \leq 0.293$, Table 1).

211 All predicted performances were significantly correlated with the actual ones, with high or very
212 high correlation coefficients ($p < 0.001$, $r \geq 0.895$, Table 1 and Figures 1–3).

213 The bias \pm 95% LoA of the DTLM were $-24 \text{ s} \pm 1 \text{ min } 11$, $17 \text{ s} \pm 51 \text{ s}$, and $-1 \text{ min } 00 \text{ s} \pm 2 \text{ min}$
214 57 s for the 3000-m, 5000-m and 10,000-m, respectively (Table 1 and Figure 1). The bias \pm
215 95% LoA of the power law were $0 \text{ s} \pm 43 \text{ s}$, $1 \text{ s} \pm 43 \text{ s}$, and $-2 \text{ s} \pm 3 \text{ min } 38 \text{ s}$ for the 3000-m,
216 5000-m and 10,000-m, respectively (Table 1 and Figure 2). The bias \pm 95% LoA of the
217 nomogram of Mercier et al. [23] were $0 \text{ s} \pm 42 \text{ s}$, $0 \text{ s} \pm 43 \text{ s}$, and $-2 \text{ s} \pm 3 \text{ min } 38 \text{ s}$ for the 3000-
218 m, 5000-m and 10,000-m, respectively (Table 1 and Figure 3).

219 On the whole, the accuracy (observed from the bias) and precision (observed from the 95%
220 LoA) were very slightly higher with the nomogram of Mercier et al. [23] compared to the power
221 law, and were higher with the nomogram of Mercier et al. [23] compared to the DTLM (Table
222 1). Moreover, a higher prediction precision (observed from a narrower 95% LoA) was noted
223 for the 5000-m, whatever the method (Table 1).

224 The 95% of CI for the bias of the three predicted methods are presented in Table 1. Compared
 1 225 to the nomogram of Mercier et al. [23] and the power law, the predictions with the DTLM
 2 226 method seemed to be less precise because the line of equality was not comprised within the
 3
 4 227 interval.
 5
 6
 7

8 228
 9

10 229 **DISCUSSION**

11
 12
 13
 14 230 The purpose of the current study was to test the validity and compare the accuracy (from the
 15
 16 231 bias) and precision (from 95% LoA) of the predictions of middle-distance running
 17
 18 232 performances for female runners using the DTLM, the power law and the nomogram of Mercier
 19
 20 233 et al. [23]. The main finding was that these three methods can be considered as valid in
 21
 22 234 predicting female athletes' performances. Indeed, as shown in Table 1, the magnitude of these
 23
 24 235 differences was considered as small if not trivial. Moreover, the results revealed that for 95%
 25
 26 236 of the predictions, the differences between the actual and predicted performances of the DTLM,
 27
 28 237 the power law and the nomogram of Mercier et al. [23], respectively, were less than or equal to
 29
 30 238 $-3.7 \pm 10.8\%$, $-0.1 \pm 6.3\%$ and $0.0 \pm 6.2\%$ for the 3000-m, $1.4 \pm 4.3\%$, $0.1 \pm 3.7\%$ and $0.0 \pm$
 31
 32 239 3.7% for the 5000-m, and $-2.5 \pm 7.4\%$, $-0.1 \pm 8.9\%$ and $-0.1 \pm 8.9\%$ for the 10,000-m. The
 33
 34 240 comparison of these differences suggests that the most accurate and precise model was the
 35
 36 241 nomogram of Mercier et al. [23] (*i.e.*, lower biases and smaller distances between 95% LoA)
 37
 38 242 followed by the power law and in third position the DTLM, whereas this last is most often cited
 39
 40 243 in the literature [7]. Indeed, although the predictions of the nomogram of Mercier et al. [23] and
 41
 42 244 the power law were close, the gaps were larger in comparisons with the DTLM. The 95% CI
 43
 44 245 for the bias [30] also showed these tendencies, with the DTLM being the least accurate. In
 45
 46 246 accordance with previous studies, all the results therefore confirmed the validity of the three
 47
 48 247 methods [1, 2, 4, 5, 7] and the lowest accuracy of the DTLM to predict the middle-distance
 49
 50 248 track-running performances of women. Indeed, although no study has compared the power law
 51
 52 249 *vs* the nomogram, previous studies have shown that the power law was more accurate than the
 53
 54 250 DTLM in predicting performance [5, 11], and the nomogram was better than the DTLM, as
 55
 56 251 well [4]. These findings might be explained by differences in design (*i.e.*, models of
 57
 58 252 construction) between the three methods. Indeed, the nomogram of Mercier et al. [23] was
 59
 60 253 developed from the authors' empirical knowledge. The power law was modeled as $y = a \cdot x^b$,
 61
 62 254 a relationship between the logarithm of T_{lim} and the logarithm of D_{lim} or S and computed by
 63
 64 255 non-linear least-square regressions [7, 21]. The DTLM was modeled as $y = a \cdot x + b$ (affine
 65

256 function), *i.e.*, a linear relationship between D_{lim} and T_{lim} [15] especially when the values of
1 257 T_{lim} are between 3.5 to 35 min [31, 32]. Beyond this range of T_{lim} , the relationship between
2 258 T_{lim} and D_{lim} is not strictly linear. Consequently, the time prediction is overestimated when
3 259 supramaximal distances are used and this would explain the lower accuracy and precision of
4 260 predictions for 10,000-m performances in this study. The results (*i.e.*, validity, accuracy and
5 261 precision) to predict one performance from two other performances presented here seem in
6 262 accordance with the results of previous studies [5, 11, 22]. However, the accuracy or precision
7 263 of predicted performances should vary as a function of the number of distance performances
8 264 taken into account in the prediction models [7].

17 265 In addition, the predictions were better for the 5000-m (whatever the method), as the 95% LoA
18 266 were narrower compared to those of the 3000-m and 10,000-m (Table 1). This observation
19 267 agrees with previous studies, which showed that the precision of predictions is higher when the
20 268 performance is interpolated rather than obtained by extrapolation [1, 2, 4, 7, 25]. Indeed, the
21 269 predictions are more accurate and precise when the distance is bounded by shorter and longer
22 270 distances (*e.g.*, a predicted performance for the 5000-m from actual performances for the 3000-
23 271 m and 10,000-m), as recommended by Mercier et al. [23]. Physiologically, predicting the
24 272 performance of intermediate distance races by interpolation rather than by extrapolation permits
25 273 a better consideration of running performance factors. In short running distances (*e.g.*, 3000-
26 274 m), maximal oxygen uptake is one of the main factors, whereas aerobic endurance is a key
27 275 factor in long running distances (*e.g.*, 10,000-m). Therefore, performance predictions of middle
28 276 running distances (*i.e.*, 5000-m) from both short (*i.e.*, 3000-m) and long distances (*e.g.*, 10,000-
29 277 m) have an advantage in that they take into consideration more of runners' essential qualities
30 278 (*e.g.*, maximal oxygen uptake, aerobic endurance...) to thus achieve more accurate and precise
31 279 predictions [33, 34].

46 280 Compared to the findings in previous studies [1, 2], the better precision and accuracy of the
47 281 predictions with the nomogram of Mercier et al. [23] (*i.e.*, the biases < 1%, and 95% LoA ≤
48 282 8.9%) could be explained by the numerical process used in the current study. Indeed, data
49 283 automatization using MATLAB code seems to have prevented a larger number of errors than
50 284 the manual process described in Coquart et al. [1]. For all the compared methods, predictions
51 285 were made using computer-assisted mathematical programs or formulas, thus eliminating the
52 286 potential risk of human error. Therefore, the three tested methods could be compared since they
53
54
55
56
57
58
59
60
61
62
63
64
65

287 all integrated a computerized process and none was dependent on a manual approach. This data
 1 288 automatization could be considered as one of the strongest points of this study.
 2
 3
 4
 5 289 Some of the differences in the results and interpretations may be explained by the decisions
 6
 7 290 made about taking into account the specialty and performance level [1, 2]. Indeed, Mercier et
 8
 9 291 al. [23] suggested that predictions would be better if the runners were specifically prepared for
 10 292 the distances from which the predictions would be made, but this hypothesis was not confirmed
 11
 12 293 in the study of Coquart et al. [1]. Moreover, the current study focused on a large sample of
 13
 14 294 women's performances regardless of their performance level. The performances might have
 15
 16 295 been classified into levels (*e.g.*, from regional to national), but the sample size of athletes with
 17
 18 296 a national level would have been too small to draw correct conclusions. Further studies,
 19
 20 297 especially for women, are needed to test whether the predictions vary according to the
 21
 22 298 performance level (*e.g.*, if the predictions are better when the performance level is high). With
 23
 24 299 the exception of the study of Coquart et al. [1] using the nomogram of Mercier et al. [23],
 25
 26 300 women's performances have not been much studied, as noted by Lerebourg and Coquart [35],
 27
 28 301 and a deeper understanding of their performances might be of great interest for both the athletes
 29
 30 302 and coaches. As pacing strategies may differ between the sexes during races [36], coaches and
 31
 32 303 federations need to understand such sex-related data in order to best prepare their athletes for
 33
 34 304 major sporting events (*e.g.*, National or Continental Championships, Olympic Games).
 35
 36 305 The main limitation of this study was that the actual running performances were probably not
 37
 38 306 executed in similar environmental conditions or with identical physical and psychological
 39
 40 307 conditions. Indeed, as noted, the analysis was limited to track performances in order to control
 41
 42 308 for the effect of the race profile, but it is possible that other variables (*e.g.*, climate conditions,
 43
 44 309 moment in the training program, motivation...) may have affected these performances. Despite
 45
 46 310 this potential limitation, we found that all three approaches were valid for predicting the
 47
 48 311 performances. It is very interesting to note particularly that the nomogram of Mercier et al. [23],
 49
 50 312 which was designed in the 1980s, is still a valid and accurate tool for predicting actual
 51
 52 313 performances [35].
 53
 54 314 Last, comparisons of predicted performances might be made using other predictive methods
 55
 56 315 based on artificial intelligence because supervised learning has recently shown promising
 57
 58 316 results in the prediction of sports results [37]. Furthermore, future studies might also be
 59
 60 317 conducted on the prediction of long-distance on-road performances (10-km, half-marathon and
 61
 62 318 marathon) for women and/or non-elite runners.
 63
 64
 65

319 **PRACTICAL APPLICATIONS**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

320 In order to prescribe speeds for tempo runs, detect the best athletes, identify a specialty and
321 even determine competitive strategies, athletes and coaches need easy tools that provide them
322 with accurate estimates of performance. The results of the current study (with an advantage for
323 the nomogram) showed that it is possible and easy to predict track-running performance from
324 the performances of two other distances. Furthermore, in addition to being the easiest method
325 to use (no calculations to perform), the nomogram of Mercier can also be used to estimate
326 physical fitness indicators (*e.g.*, maximal oxygen uptake [24], aerobic speed [12] and capacity
327 [38], index of endurance [12]) to optimize and adapt training programs for athletes.

328

329 **CONCLUSION**

330 The results of the present study confirm the validity of the three approaches (*i.e.*, DTLM, power
331 law and nomogram of Mercier et al. [23]) to predict the individual track-running performances
332 of female runners for the 3000-m, 5000-m, and 10,000-m events from the performances of two
333 other distance events. The DTLM seems to be the least accurate and precise model. Although
334 the 95% LoA of the nomogram and the power law could be considered similar, the slightly
335 lower bias of the nomogram between the actual and predicted performances for all distances
336 give it a slight practical advantage for calibrating training sessions and defining the potential
337 limits of an athlete's speed to achieve the best performance.

338

339 **ACKNOWLEDGMENTS**

340 The authors are grateful to the French Athletics Federation (Fédération Française d'Athlétisme)
341 for data diffusion, and the Orthodynamica Center at Mathilde Hospital 2 for funding.

342

343

344 **REFERENCES**

- 1
2
3 345 [1] Coquart JB, Mercier D, Tabben M, et al. Influence of sex and specialty on the
4
5 346 prediction of middle-distance running performances using the Mercier et al.'s nomogram. *J*
6
7 347 *Sports Sci* 2015; 33: 1124–1131.
8
9 348 [2] Tabben M, Bosquet L, Coquart JB. Effect of performance level on the prediction of
10 349 middle-distance-running performances using a nomogram. *Int J Sports Physiol Perform* 2016;
11
12 350 11: 623–626.
13
14 351 [3] Blythe DAJ, Király FJ. Prediction and quantification of individual athletic
15
16 352 performance of runners. *PloS One* 2016; 11: e0157257.
17
18 353 [4] Coquart J, Bosquet L. Precision in the prediction of middle distance-running
19 354 performances using either a nomogram or the modeling of the distance-time relationship. *J*
20
21 355 *Strength Cond Res* 2010; 24: 2920–2926.
22
23 356 [5] Zinoubi B, Vandewalle H, Driss T. Modeling of running performances in humans:
24
25 357 comparison of power laws and critical speed. *J Strength Cond Res* 2017; 31: 1859–1867.
26
27 358 [6] Nimmerichter A, Novak N, Triska C, et al. Validity of treadmill-derived critical speed
28 359 on predicting 5000-meter track-running performance. *J Strength Cond Res* 2017; 31: 706–
29
30 360 714.
31
32 361 [7] Vandewalle H. Modelling of running performances: comparisons of power-law,
33
34 362 hyperbolic, logarithmic, and exponential models in elite endurance runners. *BioMed Res Int*
35
36 363 2018; 2018: 8203062.
37
38 364 [8] Gamelin FX, Coquart JM, Ferrari N, et al. Prediction of one-hour running performance
39 365 using constant duration tests. *J Strength Cond Res* 2006; 20: 735–739.
40
41 366 [9] Hill DW, Vingren JL, Nakamura FY, et al. Relationship between speed and time in
42
43 367 running. *Int J Sports Med* 2011; 32: 519–522.
44
45 368 [10] Sylvan Katz J, Katz L. Power laws and athletic performance. *J Sports Sci* 1999; 17:
46
47 369 467–476.
48
49 370 [11] Zinoubi B, Vandewalle H, Zbidi S, et al. Estimation of running endurance by means of
50 371 empirical models: A preliminary study. *Sci Sports* 2019; 34: 24–29.
51
52 372 [12] Vandewalle H. A nomogram of performances in endurance running based on
53
54 373 logarithmic model of péronnet-thibault. *Am J Eng Res* 2017; 6: 78–85.
55
56 374 [13] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. From
57
58 375 empirical models to physiological models. *Sports Med Auckl NZ* 1999; 27: 359–379.
59
60 376 [14] Morton RH, Hodgson DJ. The relationship between power output and endurance: a
61
62
63
64
65

- 377 brief review. *Eur J Appl Physiol* 1996; 73: 491–502.
- 1
2 378 [15] Hughson RL, Orok CJ, Staudt LE. A high velocity treadmill running test to assess
3
4 379 endurance running potential. *Int J Sports Med* 1984; 5: 23–25.
- 5 380 [16] Hill AV. Muscular movement in man: the factors governing speed and recovery from
6
7 381 fatigue. *Muscular Mov Man Factors Gov Speed Recovery Fatigue*; 1927.
- 8
9 382 [17] Scherrer J, Monod H. Le travail musculaire local et la fatigue chez l'homme. *Presse*
10
11 383 *Med* 1960; 68: 1717–1717.
- 12 384 [18] Ettema JH. Limits of human performance and energy-production. *Int Z Für Angew*
13
14 385 *Physiol Einschließlich Arbeitsphysiologie* 1966; 22: 45–54.
- 15
16 386 [19] Monod H, Scherrer J. The work capacity of a synergic muscular group. *Ergonomics*
17
18 387 1965; 8: 329–338.
- 19
20 388 [20] di Prampero PE. The concept of critical velocity: a brief analysis. *Eur J Appl Physiol*
21
22 389 1999; 80: 162–164.
- 23 390 [21] Kennelly AE. An approximate law of fatigue in the speeds of racing animals. In:
24
25 391 *Proceedings of the American Academy of Arts and Sciences*. JSTOR, 1906, pp. 275–331.
- 26
27 392 [22] Vandewalle H, Zinoubi B, Driss T. Modelling of running performances: comparison
28
29 393 of power laws and Endurance Index. In: *Communication to the 14th Annual Conference of the*
30
31 394 *Society of Chinese Scholars on Exercise Physiology and Fitness (SCSEPF) Macau*. 2015, pp.
32
33 395 22–23.
- 34 396 [23] Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le
35
36 397 VO₂max et l'endurance relative en course de fond. *Méd Sport* 1984; 58: 181–187.
- 37
38 398 [24] Tokmakidis SP, Léger L, Mercier D, et al. New approaches to predict VO₂max and
39
40 399 endurance from running performances. *J Sports Med Phys Fitness* 1987; 27: 401–409.
- 41
42 400 [25] Coquart J, Alberty M, Bosquet L. Validity of a nomogram to predict long distance
43
44 401 running performance. *J Strength Cond Res* 2009; 23: 2119–2123.
- 45 402 [26] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation
46
47 403 coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
- 48
49 404 [27] Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1: 98–101.
- 50
51 405 [28] Munro BH. Correlation. *Stat Methods Health Care Res* 1997; 224–245.
- 52
53 406 [29] Bland JM, Altman D. Statistical methods for assessing agreement between two
54
55 407 methods of clinical measurement. *The lancet* 1986; 327: 307–310.
- 56
57 408 [30] Giavarina D. Understanding bland altman analysis. *Biochem Medica* 2015; 25: 141–
58
59 409 151.
- 60 410 [31] Vandewalle H, Vautier JF, Kachouri M, et al. Work-exhaustion time relationships and
61
62
63
64
65

- 411 the critical power concept. *Crit Rev J Sports Med Phys Fit* 1997; 37: 89–102.
- 1
2 412 [32] Hopkins WG, Edmond IM, Hamilton BH, et al. Relation between power and
3
4 413 endurance for treadmill running of short duration. *Ergonomics* 1989; 32: 1565–1571.
- 5 414 [33] di Prampero PE, Atchou G, Brückner JC, et al. The energetics of endurance running.
6
7 415 *Eur J Appl Physiol* 1986; 55: 259–266.
- 8
9 416 [34] Péronnet F, Thibault G. [Physiological analysis of running performance: revision of
10
11 417 the hyperbolic model]. *J Physiol (Paris)* 1987; 82: 52–60.
- 12 418 [35] Lerebourg L, Coquart JB. Changes in performances/characteristics of French female
13
14 419 runners over the last 12 years. *Res Sports Med* 2020; 1–11.
- 15 420 [36] Hanley B, Williams EL. Successful Pacing Profiles of Olympic Men and Women
16
17 421 3,000 m Steeplechasers. *Front Sports Act Living* 2020; 2: 21.
- 18 422 [37] Bunker RP, Thabtah F. A machine learning framework for sport result prediction.
19
20 423 *Appl Comput Inform* 2019; 15: 27–33.
- 21
22 424 [38] Bosquet L, Léger L, Legros P. Methods to determine aerobic endurance. *Sports Med*
23
24 425 *Auckl NZ* 2002; 32: 675–700.
- 25
26
27
28 426
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Prédiction de la performance en course à pied : demi-fond et fond

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Mean values and standard deviation (SD) of actual and predicted performances from the distance-time linear model (DTLM), the power law and the nomogram of Mercier et al. [23] for 158 female runners; difference between actual and predicted performances (p), Spearman correlation with actual performance (r), bias and 95% LoA (min and %), 95% CI of the bias (min), and magnitude of the bias for the 3000-m, 5000-

Distance (m)	Mean performance (min) ± SD	p	Correlation			Bias (min) ± 95% LoA	Bias (%) ± 95% LoA	95% CI for the bias (min)		Magnitude of bias	
			r	Interpretation	Lower Bound			Upper Bound	ES	Interpretation	
	Actual	11.28 ± 1.33									
3000	Predicted from the DTLM	10.88 ± 1.40	<.001	0.895*	High	-0.399 ± 1.179	-3.7 ± 10.8	-0.445	-0.290	-0.293	Small
	Predicted from the power law	11.27 ± 1.37	0.782	0.962*	Very high	-0.007 ± 0.715	-0.1 ± 6.3	-0.060	0.450	-0.005	Trivial
	Predicted from the nomogram of Mercier	11.28 ± 1.37	0.845	0.962*	Very high	-0.002 ± 0.699	0.0 ± 6.2	-0.550	0.045	-0.002	Trivial
	Actual	19.49 ± 2.34									
5000	Predicted from the DTLM	19.78 ± 2.38	<.001	0.984*	Very high	0.285 ± 0.842	1.4 ± 4.3	0.210	0.320	0.121	Trivial
	Predicted from the power law	19.50 ± 2.32	0.752	0.987*	Very high	0.010 ± 0.722	0.1 ± 3.7	-0.045	0.060	0.005	Trivial
	Predicted from the nomogram of Mercier	19.50 ± 2.32	0.988	0.987*	Very high	0.005 ± 0.712	0.0 ± 3.7	-0.050	0.050	0.002	Trivial
	Actual	41.03 ± 5.12									
10,000	Predicted from the DTLM	40.03 ± 4.94	<.001	0.962*	Very high	-0.997 ± 2.948	-2.5 ± 7.4	-1.105	-0.720	-0.199	Trivial
	Predicted from the power law	40.99 ± 5.26	0.783	0.946*	Very high	-0.040 ± 3.636	-0.1 ± 8.9	-0.280	0.210	-0.008	Trivial
	Predicted from the nomogram of Mercier	41.01 ± 5.25	0.837	0.946*	Very high	-0.026 ± 3.634	-0.1 ± 8.9	-0.270	0.210	-0.005	Trivial

m and 10,000-m track events.

Note: Bias = mean difference between actual and predicted performances; r = coefficient of correlation; LoA = limits of agreement; CI = confidence intervals; ES = effect size.

*Significantly correlated at $p < 0.001$

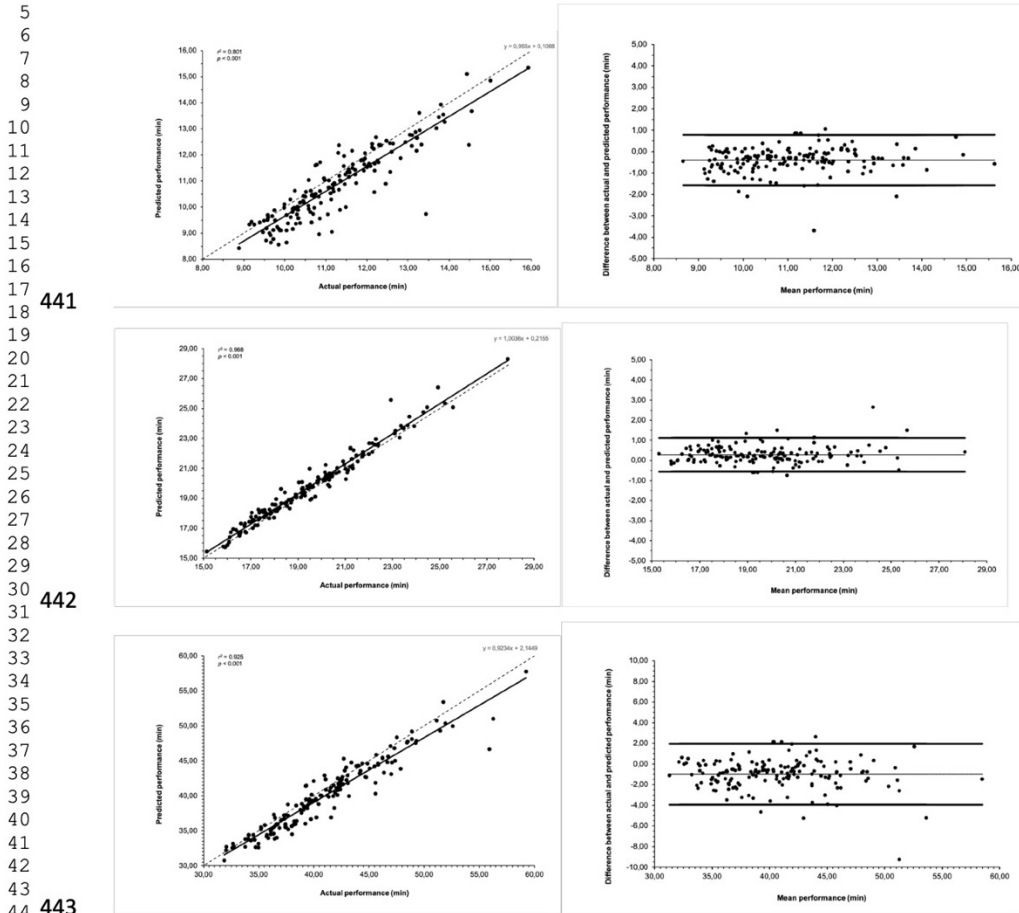
433 **Table 2.** Intra-class correlation coefficient (ICC) computations using single-rating, absolute-
 1 434 agreement, a two-way mixed-effects model between MATLAB code and manual methods to
 2
 3 435 predict performances from the nomogram of Mercier et al. [23] for 158 female runners in the
 4
 5 436 3000-m, 5000-m and 10,000-m track events.
 6

Distance (m)	Intra-class correlation coefficients	Interpretation of reliability [26]	95% Confidence Interval	
			Lower Bound	Upper Bound
3000	0.991	Excellent	0.987	0.993
5000	0.944	Excellent	0.925	0.959
10,000	0.998	Excellent	0.997	0.999

17 437
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65

438 **Figure 1.** Validity of measurements with distance-time linear model (DTLM) to predict
 439 performance.

440



441

442

443

444 *Left panel:* association between actual and predicted performances from the DTLM in the 3000-
 445 m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The solid line
 446 is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and Altman
 447 plots for the comparisons between the actual and predicted performances in the 3000-m (top),
 448 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines are the
 449 bias, solid lines are the 95% limits of agreement.

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

450 **Figure 2.** Validity of measurements with the power law to predict performance.

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

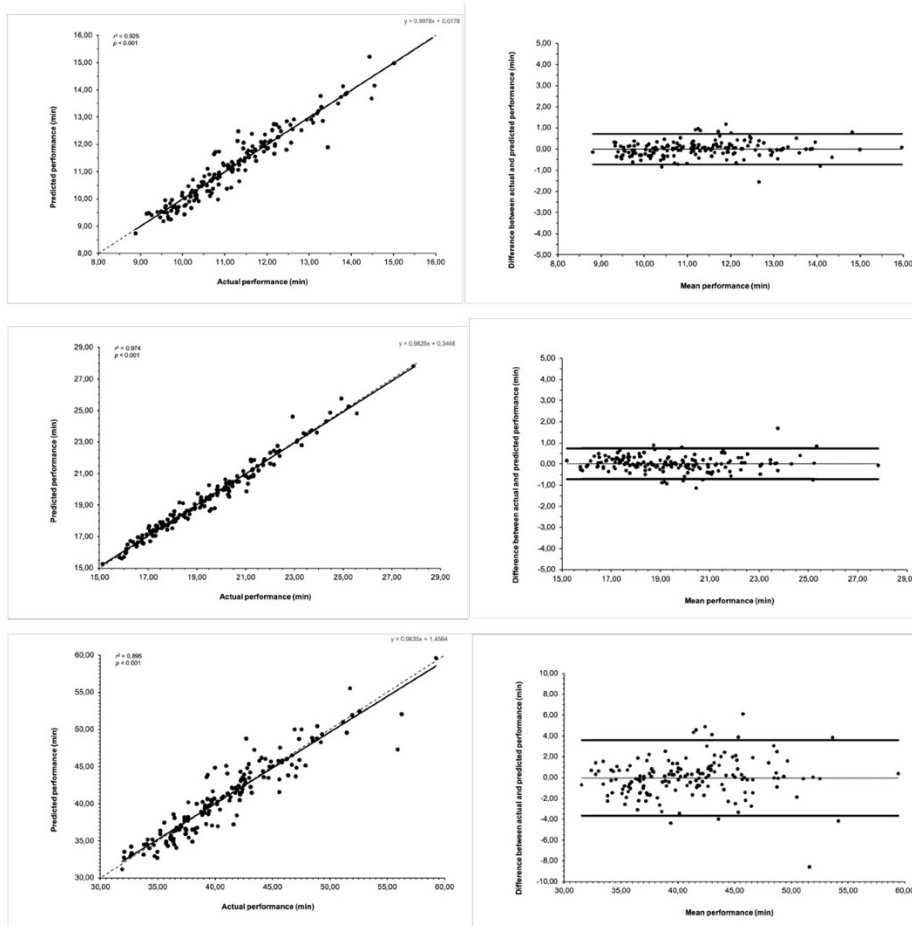
491

492

493

494

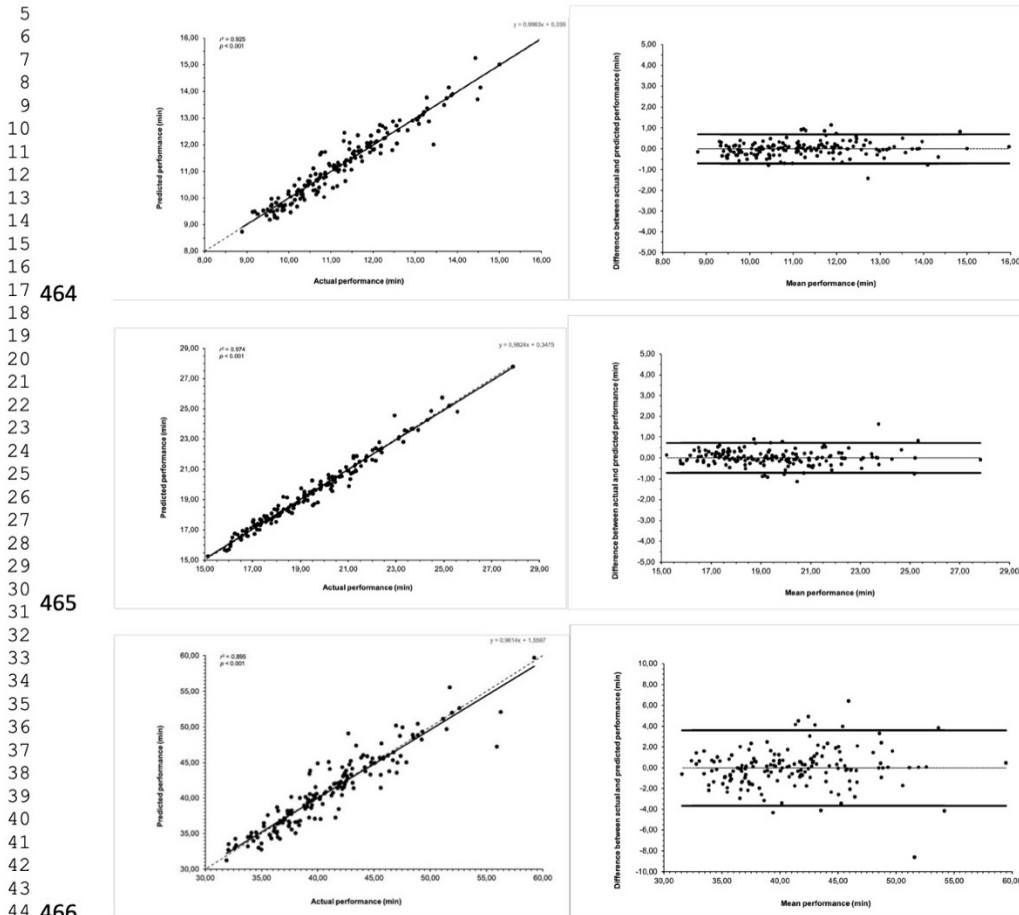
495



455 *Left panel:* association between actual and predicted performances from the power law in the
 456 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. The
 457 solid line is the linear regression. r^2 is the coefficient of determination. *Right panel:* Bland and
 458 Altman plots for the comparisons between the actual and predicted performances in the 3000-
 459 m (top), 5000-m (middle), and 10,000-m (bottom) events for 158 female runners. Dashed lines
 460 are the bias, solid lines are the 95% limits of agreement.

461 **Figure 3.** Validity of measurements with the Mercier et al. [23] nomogram to predict
 462 performance.

463



467 *Left panel:* association between actual and predicted performances from the Mercier et al.
 468 nomogram [23] in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158
 469 female runners. The solid line is the linear regression. r^2 is the coefficient of determination.
 470 *Right panel:* Bland and Altman plots for the comparisons between the actual and predicted
 471 performances in the 3000-m (top), 5000-m (middle), and 10,000-m (bottom) events for 158
 472 female runners. Dashed lines are the bias, solid lines are the 95% limits of agreement.

473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

Annexe 5. Article 5 : *Prediction of 3000-m steeplechase performance using a simple equation*

Research Quarterly for Exercise and Sport

**Research Quarterly
for Exercise and Sport**

Prediction of 3000-m steeplechase performance using a simple equation

Journal:	<i>Research Quarterly for Exercise and Sport</i>
Manuscript ID	Draft
Manuscript Type:	Original Article
Manuscript Category:	Exercise Physiology
Keyword:	Time to exhaustion, Validity, Running, Hurdling

SCHOLARONE™
Manuscripts

URL: <http://mc.manuscriptcentral.com/rqes> Email: URQE-peerreview@journals.tandf.co.uk

1 **Abstract**

2 **Introduction:** This study aimed to develop and test the validity and accuracy of a simple
3 equation for predicting 3000-m steeplechase performance ($\text{Perf}_{\text{Steeple}}$).

4 **Methods:** The official rankings of French runners for the 3000-m track-running and 3000-m
5 steeplechase events were examined. Sex, age, height, and body mass index (BMI) of each
6 runner were collected. From 171 included athletes, two groups were randomly composed: one
7 comprising 80% of the sample ($n = 137$, including 20 women) to develop a simple equation to
8 predict $\text{Perf}_{\text{Steeple}}$ (*i.e.*, development group) and the other comprising the remaining 20% ($n =$
9 34, including 5 women) to test the validity and accuracy of the developed prediction equation
10 (*i.e.*, cross-validation group).

11 **Results:** The simple prediction equation included 3000-m running performance, age and sex.
12 No significant difference was noted between the actual and predicted performances ($p = 0.387$).
13 Predicted performances were significantly correlated with the actual ones, with a very high
14 correlation coefficient ($p < 0.001$; $r = 0.934$). Bias and 95% limits of agreement were -3 ± 22 s,
15 *i.e.*, $-0.5 \pm 7.1\%$.

16 **Conclusion:** The study confirms the validity and accuracy of the equation for predicting
17 $\text{Perf}_{\text{Steeple}}$. Predictions using this simple equation may be used in training programs and
18 competitions for athletes and coaches.

19 **Keywords:** time to exhaustion, validity, running, hurdling, athlete.

20

22 Introduction

23

24 The prediction of running performances is very important for athletes and trainers.¹⁻⁴ Indeed,
25 predictions might be used to detect talents (*i.e.*, those who have the potential for high-level
26 performances and/or the potential to break records), select a speed for specific pace training
27 sessions, choose splitting times during competition, and even evaluate the effect of a training
28 program.²⁻⁵ Moreover, predictions circumvent the need for additional tests requiring volitional
29 exhaustion, which may result in inadequate recovery and impaired performances on subsequent
30 days.^{6,7}

31 Sports researchers have therefore long sought ways to accurately predict running performances
32 and have developed certain theories and tools, such as the power law,^{8,9} the concept of critical
33 speed^{10,11} and some nomograms.^{12,13} However, these different tools/methods are used to predict
34 middle- and long-distance track-running performances, but only in non-obstacle races, and the
35 studies on predicting the track-running performances of the steeplechase remain very limited.

36 The 3000-m steeplechase is a technical middle-distance running event which takes place on a
37 track with the athletes crossing 28 barriers and 7 barriers followed by a water pit, called water-
38 jumps.¹⁴ The main technical/regulatory difference between the sexes is the height of the
39 barriers, which is 0.914 m and 0.762 m for men and women, respectively.¹⁴

40 The predictive variables of a 3000-m steeplechase might be the same variables as for flat
41 middle- and long-distance running, such as maximal oxygen uptake ($\dot{V}O_2\text{max}$), endurance
42 capacity and running economy,^{15,16} which are key physiological factors of running known to be
43 dependent on individual characteristics like age, sex, height and body mass index.^{17,18} Few
44 studies have investigated the relationship between performance and age among elite
45 steeplechase runners.^{19,20} The authors have shown that the age of elite athletes at peak
46 performance on the 3000-m steeplechase is close to 25 years for men and between 25 and 27
47 years for women. Sex also appears also to be a variable to consider in predicting performance
48 in a technical race like the 3000-m steeplechase. Indeed, beyond the fact that men typically

URL: <http://mc.manuscriptcentral.com/rqes> Email: URQE-peerreview@journals.tandf.co.uk

49 perform better than women in middle- and long-distance events,^{21,22} differences in the running
50 technique (*e.g.*, step lengths and body positioning in hurdling) used to cross the steeplechase
51 water-jumps and barriers have been identified between the sexes,²³ probably due to the lower
52 official barrier height for women.¹⁴ These differences in hurdling technique may also be due to
53 the approach velocity and body height.²⁴ This is perhaps why Sánchez Muñoz et al.²⁵ recently
54 showed that runners of long-distance events (including the 3000-m steeplechase) are taller than
55 runners of shorter distances like the 800-m and 1500-m. Being taller is often associated with
56 longer legs, which would facilitate barrier crossing. Moreover, it is well established that a lower
57 body mass index (BMI) positively impacts middle- and long-distance running performance.²⁶
58 Consequently, due to the technical aspect and the energy cost of crossing barriers,²⁷ height and
59 BMI appear to be other predictive variables of steeplechase performance.
60 The purpose of the current study was to develop and test the validity and accuracy of a simple
61 equation in predicting 3000-m steeplechase performance from individual variables (*e.g.*, sex,
62 age, height, and BMI).

63

64 **Methods**

65 All official [REDACTED] rankings from the [REDACTED] Athletics Federation [REDACTED]
66 [REDACTED] in 2019 for the 3000-m track ($n = 5,385$) and 3000-m steeplechase (n
67 = 1,333) events were retrospectively scrutinized. From these rankings, all adult athletes (≥ 18
68 years) who had participated in the two races were retained ($n = 387$). Runners who had not self-
69 reported their body mass and/or height were then removed from the analysis ($n = 211$). Thus,
70 176 athletes were included in this stage. However, runners who had maintained a higher speed
71 in the 3000-m steeplechase than in the 3000-m race were also eliminated ($n = 5$). From the 171
72 remaining athletes, two groups were randomly composed using Statistical Package for the
73 Social Sciences (SPSS) software, one comprising 80% of the sample ($n = 137$, including 20
74 women) to develop a simple prediction equation (*i.e.*, development group) and a second

75 comprising the remaining 20% ($n = 34$, including 5 women) to test the validity and accuracy of
76 the developed simple prediction equation (*i.e.*, cross-validation group).

77 For each athlete, the sex (*i.e.*, woman vs man), birth date (to calculate the age), height and body
78 mass (to calculate the BMI), as well as the race times on the 3000-m steeplechase (*i.e.*,
79 $Perf_{\text{Steeple}}$) and 3000-m race (*i.e.*, $Perf_{3000}$) were recorded.

80 This study was approved by the National Ethics Committee for Research in Sports Sciences
81 (CERSTAPS 2019-22-02-31). Moreover, the protocol for this study was legally declared, in
82 accordance with the European General Data Protection Regulations.

83 *Statistical Analysis*

84 Standard statistical methods were used to calculate the means and standard deviations (SD).

85 The normal Gaussian distribution was verified by the Shapiro-Wilk test.

86 *Development of the simple prediction equation*

87 To investigate whether an equation could be developed to predict $Perf_{\text{Steeple}}$, a Pearson product
88 moment correlation was used to evaluate the association between the dependent (*i.e.*, $Perf_{\text{Steeple}}$)
89 and several independent variables (*i.e.*, $Perf_{3000}$, sex, age, height, body mass and BMI).

90 A simple prediction equation was then developed from the significantly correlated variables
91 using stepwise multiple linear regression analysis.

92 The variance inflation factor (VIF) was used to detect the severity of multicollinearity among
93 the independent variables in the regression model.

94 Fisher's tests were used to examine the contribution of each variable to the model, and
95 confirmation was based on the analysis of standardized β coefficients.

96 Moreover, the relationship between $Perf_{\text{Steeple}}$ estimated by the simple prediction equation and
97 actual $Perf_{\text{Steeple}}$ was analyzed using the Bravais-Pearson method and quantified using Pearson's
98 correlation coefficient.

99 The autocorrelation in the residuals was examined with the Durbin-Watson test.

100 *Cross-validation of the simple prediction equation*

101 Student's paired samples t-test was used to compare the actual and predicted Perf_{SteepLe}. The
102 magnitude of the differences was assessed by the effect size (ES).

103 The association between the actual and predicted Perf_{SteepLe} was tested using the Bravais-
104 Pearson method and quantified by Pearson's correlation coefficient.

105 The bias (*i.e.*, difference between actual and predicted Perf_{SteepLe}) and 95% limits of agreement
106 (95% LoA, *i.e.*, ± 1.96 SD) were computed according to the Bland-Altman method.

107 Statistical significance was set at $p < 0.05$ and all analyses were performed with the SPSS
108 (release 20.0, Chicago, IL, USA).

109 **Results**

110 The characteristics of the athletes in both groups are presented in Table 1.

111 *Development of the simple prediction equation*

112 Significant bivariate correlations were found between Perf₃₀₀₀ ($p < 0.001$, $r = 0.892$), age ($p <$
113 0.001 , $r = 0.428$), sex ($p = 0.021$, $r = -0.173$), height ($p = 0.050$, $r = -0.141$), BMI ($p = 0.003$, r
114 $= 0.230$) and Perf_{SteepLe}.

115 Stepwise multiple linear regression analysis using Perf₃₀₀₀, sex and age as potential independent
116 variables yielded the simple prediction equation.

117 This equation was as follows:

$$118 \text{ Perf}_{\text{SteepLe}} = -89.740 + 1.153 \times \text{Perf}_{3000} + 0.952 \times \text{age} + 14.810 \times \text{sex}$$

119 with Perf_{SteepLe} and Perf₃₀₀₀ in seconds, sex = 1 for women and 2 for men, and age in years.

120 VIF was 1.341, 1.241 and 1.233 for Perf₃₀₀₀, age and sex, respectively. Therefore,
121 multicollinearity was very low (because VIF < 5).

122 The increases in r^2 by adding a second (*i.e.*, age) and third predictor (*i.e.*, sex) into the prediction
123 equation were significant with $F(1,134) = 14.912$ ($p < 0.001$) and $F(1,133) = 4.465$ ($p = 0.036$),
124 respectively. Moreover, Fisher's test revealed a $p < 0.001$.

125 The performance estimated by the simple prediction equation (including the 3 independent
126 variables: Perf₃₀₀₀, age and sex) was significantly correlated with the actual Perf_{Steeple} ($r = 0.906$
127 and $r^2 = 0.822$).

128 The standardized β coefficients and p values on the Student's t -test were 0.878 ($p < 0.001$),
129 0.125 ($p = 0.003$) and 0.086 ($p < 0.036$) for Perf₃₀₀₀, age and sex, respectively.

130 The Durbin-Watson test indicated no autocorrelation in the residuals, with a value of 1.737.
131 Moreover, the standardized residuals followed a straight line, showing that no residual was too
132 high, so the prediction was valid for all performances.

133 *Cross-validation of the simple prediction equation*

134 No significant difference was noted between the actual and predicted Perf_{Steeple} ($p = 0.387$), and
135 the magnitude of the difference was considered small (ES = 0.36), as proposed by Cohen.²⁸

136 Predicted Perf_{Steeple} was significantly correlated with actual Perf_{Steeple} ($p < 0.001$), and a very
137 high correlation coefficient was found according to Munro's scale²⁹ ($r = 0.934$, Figure 1).

138 The bias and 95% LoA were -3 ± 22 s, *i.e.*, $-0.5 \pm 7.1\%$ of mean Perf_{Steeple} (Figure 1).

139

140 **Discussion**

141 The purpose of this study was to develop and test the validity and accuracy of a simple equation
142 to predict Perf_{Steeple} from individual variables (*e.g.*, Perf₃₀₀₀, sex, age, height and BMI). The
143 main result suggests that a simple equation including Perf₃₀₀₀, age and sex (*i.e.*, Perf_{Steeple} = -
144 $89.740 + 1.153 \times \text{Perf}_{3000} + 0.952 \times \text{age} + 14.810 \times \text{sex}$) is a valid and accurate method for
145 predicting athletes' 3000-m steeplechase performances. The results revealed that for 95% of the
146 predictions, the difference between the actual and predicted performances would be $\leq -0.5 \pm$
147 7.1% for the 3000-m steeplechase. Although no study has focused on predicting steeplechase
148 performances from a simple equation, a few studies have validated predictive methods (*e.g.*,
149 nomogram and critical speed) for the flat 3000-m with results comparable (*i.e.*, from low bias
150 and 95% LoA) to our own;^{4,30,31} thus confirming the validity of the present simple equation for
151 the 3000-m steeplechase.

152 The present results suggest that the main predictor of Perf_{Steeple} is Perf₃₀₀₀. This result might be
153 explained by the observation that the intensity and duration of effort in these two events are
154 relatively close (identical running distance with only about 3-5% prolonged effort in the 3000-
155 m steeplechase, probably having little impact on the importance of the key physiological
156 predictors of performance, such as $\dot{V}O_2\text{max}$ and endurance capacity).^{27,32,33} However, unlike
157 the 3000-m event, the 3000-m steeplechase generally produces a deceleration during the takeoff
158 step for both hurdling and water-jumps even in elite steeplechasers.³⁴ Therefore, in addition to
159 requiring a high level of $\dot{V}O_2\text{max}$ (approximately 95%) during the race,³³ 3000-m steeplechase
160 performance also requires athletes to have high technical abilities to cross the barriers and
161 water-jumps without losing speed.

162 The technical characteristics of the 3000-m steeplechase, such as the height of the barriers with
163 a 0.152-m difference between men and women,¹⁴ can lead to differences in performance
164 capacities between the sexes. Indeed, the results showed that the sex variable slightly enhanced
165 the precision of Perf_{Steeple}. This result might be explained by sex-related differences in the
166 running-hurdling technique (*e.g.*, approach velocity, landing distance and crouching height)
167 used to cross the barriers and water-jumps in the steeplechase.²³ For example, the lower barrier
168 height relative to body height for women (45.9 ± 1.4 vs $51.6 \pm 1.9\%$ of their body height for
169 women and men, respectively, in the current study) might help them to maintain race pace when
170 crossing the barriers or water-jumps.²⁴ However, although women's race paces may be less
171 affected when crossing barriers,²⁴ they are more affected by the water-jump (*e.g.*, lower
172 approach velocity, shorter jump, uphill step out of the water pit following landing) than those
173 of the men²³ probably because they need a longer stride to cross the water due to their generally
174 shorter leg length, and this is accentuated if they lean on the lower barriers.

175 Although the present study shows that age contributes significantly to performance prediction
176 in the 3000-m steeplechase, the inclusion of this variable in our simple equation provides only
177 a small gain in accuracy. This result can be explained by the observation that most

178 steeplechasers (95% of the population included in the current study) are between 20 and 35
179 years old (Table 1). This is not surprising, given that a peak age of around 25 and 27 years old
180 has been noted, respectively, for men and women in this event.¹⁹ It has been well established
181 that age can limit peak performance due to the age-related physiological modifications that
182 characterize human capacities.³⁵ Indeed, a progressive reduction in $\dot{V}O_2\text{max}$ appears to be the
183 primary mechanism associated with performance declines with adult aging.³⁶

184 Although height and BMI might have been predictive variables of $\text{Perf}_{\text{Steeple}}$ given the technical
185 aspect and energy cost of crossing barriers,²⁷ the current study suggests the opposite. This result
186 might be explained by our use of Perf_{3000} for the equation to predict $\text{Perf}_{\text{Steeple}}$. Indeed, although
187 height is only slightly associated with speed in running events, BMI seems to be a better
188 indicator of speed on the 3000-m.²⁶ Consequently, the impact of height and BMI on Perf_{3000}
189 would explain why these variables were not included in the equation for predicting $\text{Perf}_{\text{Steeple}}$
190 (because they are already included in Perf_{3000} and would therefore be redundant).

191 The main limitation of this study was that the actual running performances were not executed
192 in similar environmental conditions (*e.g.*, temperature, barometric pressure, precipitation...),
193 with identical physical and psychological (*e.g.*, motivation and stress) conditions, and using the
194 same opposition field and/or race strategy. Moreover, another limitation was the self-
195 declaration of body height and mass and thus the calculation of the athletes' BMI. Height and
196 body mass were not measured, with the athletes only reporting their anthropometric data.
197 However, runners are known to self-report accurately for this type of data.³⁷

198 Athletes and coaches need field measures that can provide them with accurate predictions of
199 performance. The results showed that this simple equation is valid to predict $\text{Perf}_{\text{Steeple}}$ from
200 individual variables (*i.e.*, $\text{Perf}_{3000\text{m}}$, age and sex). Having an accurate idea of an athlete's
201 performance capability at a specific competition like the steeplechase event may facilitate the
202 detection of the best performers, the design of training programs, the prescription speed for
203 pacing, and the determination of strategies during races.

204

205 **Perspective**

206 The results of the present study indicate the validity and accuracy of a simple equation to predict
207 Perf_{Steeple} from individual variables like Perf₃₀₀₀, age and sex. Results show no significant
208 difference between the actual and predicted performances in steeplechasers. The predictions
209 from this simple equation may be used by athletes and coaches.

210

211 **Acknowledgements**

212 The authors are grateful to the [REDACTED] for data collection, the
213 *Institut de Recherche Interdisciplinaire Homme et Société (IRIHS)* and the *Orthodynamica*
214 *Center at Mathilde Hospital 2*, as well as the *Institut des Rencontres de la Forme*, for funding.

215

216 **Declarations of interest**

217 None

218

219

220 **References**

- 221 1. Blythe DAJ, Király FJ. Prediction and quantification of individual athletic performance of
222 runners. *PLoS One*. 2016;11(6):e0157257. doi:10.1371/journal.pone.0157257
- 223 2. Coquart JB, Eston RG, Noakes TD, et al. Estimated time limit: a brief review of a
224 perceptually based scale. *Sports Med Auckl NZ*. 2012;42(10):845-855.
225 doi:10.2165/11635370-000000000-00000
- 226 3. Coquart JBJ, Alberty M, Bosquet L. Validity of a nomogram to predict long distance
227 running performance. *J Strength Cond Res*. 2009;23(7):2119-2123.
228 doi:10.1519/JSC.0b013e3181b3dccc
- 229 4. Coquart JBJ, Bosquet L. Precision in the prediction of middle distance-running
230 performances using either a nomogram or the modeling of the distance-time relationship.
231 *J Strength Cond Res*. 2010;24(11):2920-2926. doi:10.1519/JSC.0b013e3181c69278
- 232 5. Berndsen J, Smyth B, Lawlor A. Pace my race: recommendations for marathon running.
233 In: *Proceedings of the 13th ACM Conference on Recommender Systems*. ; 2019:246-250.
- 234 6. Coquart J, Tabben M, Farooq A, Tourny C, Eston R. Submaximal, Perceptually Regulated
235 Exercise Testing Predicts Maximal Oxygen Uptake: A Meta-Analysis Study. *Sports Med*
236 *Auckl NZ*. 2016;46(6):885-897. doi:10.1007/s40279-015-0465-x
- 237 7. Sassi A, Marcora SM, Rampinini E, Mognoni P, Impellizzeri FM. Prediction of time to
238 exhaustion from blood lactate response during submaximal exercise in competitive
239 cyclists. *Eur J Appl Physiol*. 2006;97(2):174-180. doi:10.1007/s00421-006-0157-1
- 240 8. Kennelly AE. An Approximate Law of Fatigue in the Speeds of Racing Animals. *Proc*
241 *Am Acad Arts Sci*. 1906;42(15):275-331. doi:10.2307/20022230
- 242 9. Zinoubi B, Vandewalle H, Driss T. Modeling of running performances in humans:
243 comparison of power laws and critical speed. *J Strength Cond Res*. 2017;31(7):1859-1867.
244 doi:10.1519/JSC.0000000000001542
- 245 10. Gamelin FX, Coquart JM, Ferrari N, et al. Prediction of one-hour running performance
246 using constant duration tests. *J Strength Cond Res*. 2006;20(4):735-739. doi:10.1519/R-
247 17905.1
- 248 11. Hughson RL, Orok CJ, Staudt LE. A high velocity treadmill running test to assess
249 endurance running potential. *Int J Sports Med*. 1984;5(1):23-25. doi:10.1055/s-2008-
250 1025875
- 251 12. Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le
252 VO₂max et l'endurance relative en course de fond. *Médecine Sport*. 1984;58:181-187.
- 253 13. Vandewalle H. A nomogram of performances in endurance running based on logarithmic
254 model of péronnet-thibault. *Am J Eng Res*. 2017;6:78-85.
- 255 14. IAAF. *Les Règles Des Compétitions 2018-2019*.; 2017. <http://www.iaaf.org>
- 256 15. Di Prampero PE, Atchou G, Brückner JC, Moia C. The energetics of endurance running.
257 *Eur J Appl Physiol*. 1986;55(3):259-266. doi:10.1007/bf02343797

URL: <http://mc.manuscriptcentral.com/rqes> Email: URQE-peerreview@journals.tandf.co.uk

- 258 16. Kenney WL, Hodgson JL. Variables predictive of performance in elite middle-distance
259 runners. *Br J Sports Med.* 1985;19(4):207. doi:10.1136/bjism.19.4.207
- 260 17. Trappe SW, Costill DL, Vukovich MD, Jones J, Melham T. Aging among elite distance
261 runners: a 22-yr longitudinal study. *J Appl Physiol Bethesda Md 1985.* 1996;80(1):285-
262 290. doi:10.1152/jappl.1996.80.1.285
- 263 18. Vickers AJ, Vertosick EA. An empirical study of race times in recreational endurance
264 runners. *BMC Sports Sci Med Rehabil.* 2016;8(1):26. doi:10.1186/s13102-016-0052-y
- 265 19. Haugen TA, Solberg PA, Foster C, Morán-Navarro R, Breitschädel F, Hopkins WG. Peak
266 Age and Performance Progression in World-Class Track-and-Field Athletes. *Int J Sports
267 Physiol Perform.* 2018;13(9):1122-1129. doi:10.1123/ijsp.2017-0682
- 268 20. Hollings SC, Hopkins WG, Hume PA. Age at Peak Performance of Successful Track &
269 Field Athletes. *Int J Sports Sci Coach.* 2014;9(4):651-661. doi:10.1260/1747-
270 9541.9.4.651
- 271 21. Chevront SN, Carter R, Deruisseau KC, Moffatt RJ. Running performance differences
272 between men and women:an update. *Sports Med Auckl NZ.* 2005;35(12):1017-1024.
273 doi:10.2165/00007256-200535120-00002
- 274 22. Coast JR, Blevins JS, Wilson BA. Do gender differences in running performance
275 disappear with distance? *Can J Appl Physiol Rev Can Physiol Appl.* 2004;29(2):139-145.
276 doi:10.1139/h04-010
- 277 23. Hunter I, Lindsay BK, Andersen KR. Gender Differences and Biomechanics in the 3000M
278 Steeplechase Water Jump. *J Sports Sci Med.* 2008;7(2):218-222.
- 279 24. Hunter I, Bushnell TD. Steeplechase barriers affect women less than men. *J Sports Sci
280 Med.* 2006;5(2):318-322.
- 281 25. Sánchez Muñoz C, Muros JJ, López Belmonte Ó, Zabala M. Anthropometric
282 Characteristics, Body Composition and Somatotype of Elite Male Young Runners. *Int J
283 Environ Res Public Health.* 2020;17(2). doi:10.3390/ijerph17020674
- 284 26. Sedeaud A, Marc A, Marck A, et al. BMI, a performance parameter for speed
285 improvement. *PloS One.* 2014;9(2):e90183. doi:10.1371/journal.pone.0090183
- 286 27. Earl S, Hunter I, Mack GW, Seeley M. The relationship between steeplechase hurdle
287 economy, mechanics, and performance. *J Sport Health Sci.* 2015;4(4):353-356.
288 doi:10.1016/j.jshs.2015.03.009
- 289 28. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* L. Erlbaum Associates;
290 1988.
- 291 29. Munro BH. *Statistical Methods for Health Care Research.*; 1997.
- 292 30. Coquart JB, Mercier D, Tabben M, Bosquet L. Influence of sex and specialty on the
293 prediction of middle-distance running performances using the Mercier et al.'s nomogram.
294 *J Sports Sci.* 2015;33(11):1124-1131. doi:10.1080/02640414.2014.986499

- 295 31. Tabben M, Bosquet L, Coquart JB. Effect of Performance Level on the Prediction of
296 Middle-Distance-Running Performances Using a Nomogram. *Int J Sports Physiol*
297 *Perform.* 2016;11(5):623-626. doi:10.1123/ijsp.2015-0471
- 298 32. Popov T. Hurdling in the steeplechase. *Modern Athlete and coach.* 1983:17-18.
- 299 33. Slawinski J, Chiron F, Millot B, Taoouji A, Brocherie F. Effect of a 16-Day Altitude
300 Training Camp on 3,000-m Steeplechase Running Energetics and Biomechanics: A Case
301 Study. *Front Sports Act Living.* 2019;1. doi:10.3389/fspor.2019.00063
- 302 34. Kipp S, Taboga P, Kram R. Ground reaction forces during steeplechase hurdling and
303 waterjumps. *Sports Biomech.* 2017;16(2):152-165. doi:10.1080/14763141.2016.1212917
- 304 35. Lerebourg L, Coquart JB. Changes in performances/characteristics of French female
305 runners over the last 12 years. *Res Sports Med.* Published online 2020:1-11.
- 306 36. Tanaka H, Seals DR. Endurance exercise performance in Masters athletes: age-associated
307 changes and underlying physiological mechanisms. *J Physiol.* 2008;586(1):55-63.
308 doi:10.1113/jphysiol.2007.141879
- 309 37. Nikolaidis PT, Knechtle B. Validity of Recreational Marathon Runners' Self-Reported
310 Anthropometric Data. *Percept Mot Skills.* 2020;127(6):1068-1078.
311 doi:10.1177/0031512520930159

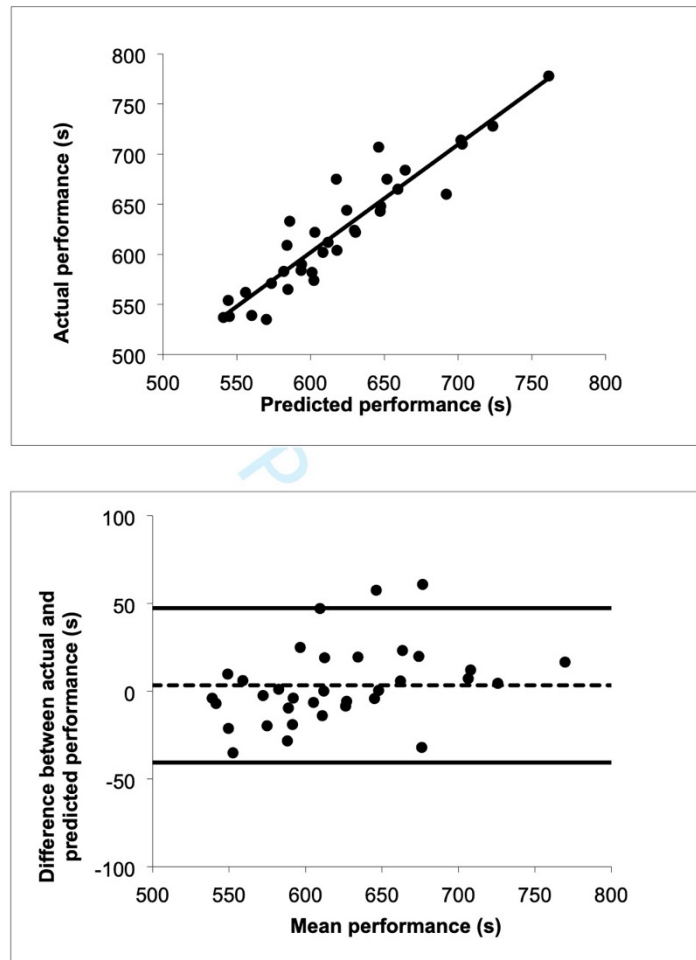
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340

URL: <http://mc.manuscriptcentral.com/rqes> Email: URQE-peerreview@journals.tandf.co.uk

Table 1. Descriptive data and variables used to develop the simple equation of prediction (mean values \pm SD).

	Group for the development of the prediction equation	Group for the cross-validation of the prediction equation
Sample size	137	34
Men (%)	85.4	85.3
Age (y)	27.8 \pm 8.0	27.3 \pm 7.3
Body mass (kg)	61.2 \pm 8.0	61.8 \pm 7.0
Height (cm)	175 \pm 8	176 \pm 5
Body mass index (kg.m ⁻²)	19.8 \pm 1.6	19.8 \pm 1.5
Perf _{Steep} le (min and s)	10 min 26 s \pm 1 min 01 s	10 min 23 s \pm 1 min 02 s
Perf _{3000m} (min and s)	9 min 34 s \pm 47 s	9 min 29 s \pm 46 s

Figure 1. Validity of measurements with the simple equation to predict performance.



Annexe 6. Article 6 : *Prediction of marathon performance using artificial intelligence*

Artificial Intelligence In Medicine
Prediction of marathon performance using artificial intelligence
 --Manuscript Draft--

Manuscript Number:	
Article Type:	Research Paper
Keywords:	artificial neural networks; k-nearest neighbor; machine learning; endurance running; modeling
Corresponding Author:	Lucie LEREBOURG, Ph.D student Université de Rouen UFR STAPS FRANCE
First Author:	Lucie LEREBOURG, Ph.D student
Order of Authors:	Lucie LEREBOURG, Ph.D student Damien Saboul Michel Cléménçon Jeremy B. Coquart
Manuscript Region of Origin:	Europe
Abstract:	<p>Purpose</p> <p>The purpose of this study was to test the validity and compare the precision and accuracy of predictions of marathon performance using two artificial intelligence techniques: artificial neural network (ANN) and k-nearest neighbor (KNN).</p> <p>Methods</p> <p>Official French rankings for the 10-km road and marathon events in 2019 were scrutinized. Only athletes (women and men) who performed both a 10-km road race and a marathon in the same year (i.e., 2019) were examined. Analysis was performed with a dataset of 820 athletes. For the two supervised machine learning algorithms, the same inputs (10-km race time, body mass index, age and sex) were used to solve a linear regression problem which consists in estimating the value of the same continuous output (marathon race time).</p> <p>Results</p> <p>No difference was found between the actual and predicted marathon performances for either method ($p > 0.05$). All predicted performances were significantly correlated with the actual ones, with very high correlation coefficients ($p < 0.001$; $r > 0.90$). KNN outperformed ANN with a mean absolute error of 2.4 vs 5.6%.</p> <p>Conclusion</p> <p>The study confirms the validity of the two supervised methods of artificial intelligence, with better accuracy for KNN in predicting marathon performance. Consequently, the predictions from these artificial intelligence methods may be used in training programs and competitions.</p>

Original investigation

Prediction of marathon performance using artificial intelligence

Running head: Prediction of marathon performance using artificial intelligence

Authors: Lucie Lerebourg¹, Damien Saboul^{2,3}, Michel Cléménçon¹, Jeremy B. Coquart¹

Affiliations:

1. *Normandie Univ, UNIROUEN, Centre d'Etudes des Transformations des Activités Physiques et Sportives, Rouen, France.*
2. *Be-Ys Research, a Brand of Be-Ys Group, Genève, Suisse.*
3. *Univ Lyon, Université Claude Bernard Lyon 1, Laboratoire Inter Universitaire de Biologie de la Motricité, Lyon, France.*

Corresponding Author:

Lucie Lerebourg, Centre d'Etudes des Transformations par les Activités Physiques et Sportives, Boulevard Siegfried, 76 821 Mont Saint Aignan, France. Phone: (+33)648157388. E-mail: lulu.lerebourg@gmail.com

Abstract word count: 213 words

Text-only word count: 3341 words (without references)

Number of tables/figures: 1 table and 3 figures

Introduction

The marathon, an athletic endurance event of 42.195 km, was created for the first modern Olympic Games in Athens in 1896. Since the first "urban tour" marathon in New York City in 1976 [1], the marathon has been gaining in popularity and has evolved from an Olympic event to a worldwide social phenomenon [2]. As enthusiasm for this event has increased [1–3], race times have steadily improved [2–4]. From recreational runners of all ages to elite athletes, the objectives may differ widely, from being a finisher, to running the race as fast as possible, to winning it and/or breaking records (*e.g.*, personal, national, world records). Although long-distance performances, as in the marathon, can be influenced by factors beyond the athlete's control (*e.g.*, climate conditions, seasonal characteristics...), they mainly depend on personal characteristics (physical qualities, psychological traits and states...) and training variables (tactics, pacing strategy...) [4–7]. Athletes and coaches need to be aware of these parameters and should focus on developing appropriate training programs, with particular emphasis on setting speeds for tempo runs and building competitive or optimal pace strategies to optimize performance [8–12]. For these reasons, the ability to predict marathon performance can be of great interest in the calibration of training sessions and the definition of the athlete's potential speed limits in order to achieve the best performance.

The relationships between running distance (or speed) and time have long been used for this purpose [13–15]. Several studies have sought to predict long-distance running performances for events like the marathon with equations (*e.g.*, logarithmic, hyperbolic, exponential, multiple regression models...) [12, 16–18], including concepts of critical speed [19, 20] or power laws [19, 21], and machine learning algorithms (*i.e.*, artificial intelligence: AI) [8, 10, 11].

Recently, there has been growing interest in machine learning algorithms, notably with supervised learning, one of the intelligent methodologies that have shown promising results in

the prediction of continuous variables in many areas such as weather [22], health [23] and sports [24]. The literature indicates that sport is one of the expanding areas requiring good predictive accuracy [10, 24, 25]. However, although machine learning regression models like artificial neural networks (ANN) [23, 25–27] or k-nearest neighbors (KNN) [8] have been used to predict performances in some sports activities, the validity and accurate prediction of individual or team performances using AI merits further exploration [28–31].

ANN is a powerful black-box supervised learning algorithm capable of producing nonlinear input-output mapping [28, 32, 33]. The model consists of one input layer, one or more hidden layers and one output layer. The interconnected components (*i.e.*, neurons) transform a set of inputs into a desired output [32, 33]. The accuracy of this type of model is typically improved by using additional data (*i.e.*, weights associated with interconnected components are continuously changing) during the ANN training process [24, 32]. On the other hand, the KNN model uses one of the simplest types of supervised machine learning algorithms based on learning by analogy, that is, by comparing a given test example with training examples that are similar to it [23, 32]. The basic KNN algorithm has two steps: find the k training examples that are closest (“closeness” is defined in terms of a distance metric, such as the Euclidean distance) to the unseen example and take the average of these k label values [28, 32]. This machine learning model is also noted for not requiring learning (*i.e.*, the computation of the algorithm occurs during runtime) as it memorizes the training dataset [28, 30, 32].

Although studies have focused on the use of machine learning algorithms (bagging, local matrix completion...) to predict marathon performance [8, 11, 16] and future slowdowns during the race [10], no study to the best of our knowledge has used and validated ANN and KNN supervised machine learning in this running discipline and compared the accuracy (*i.e.*, the nearness of the actual performance to the predicted performance, and thus a lower mean

absolute error or a lower bias, meaning higher accuracy) or precision (*i.e.*, the closeness of the predicted performances, and thus a smaller distance between limits of agreement, meaning higher precision).

The objectives of the current study were therefore to test the validity of two supervised machine learning methods (ANN and KNN) and to compare the accuracy and precision of the marathon performance predictions to determine which one performed best.

Methods

All French official rankings of the French Athletics Federation (FFA for *Fédération Française d'Athlétisme*) for the 10-km road race ($n = 217,669$) and the marathon ($n = 92,813$), both performed in 2019, were retrospectively scrutinized. In France, the marathon is not open to younger categories of athletes, so only athletes over the age of 21 years were selected for both races ($n = 201,990$ on the 10-km and $n = 92,813$ on the marathon). If the athletes had not self-reported their body mass and/or height, they were removed from the analysis. Thus, 7,716 athletes with a 10-km performance, and 4,130 in the marathon were included. Then, only those athletes (women and men) who performed the 10-km and the marathon in the same year (*i.e.*, 2019) were retained. Thus, 1,728 performances were collected. However, as the aim was to predict marathon performance based on 10-km road performance, athletes who ran a marathon before their 10-km were removed ($n = 833$). Moreover, athletes who maintained a higher speed in the marathon than in the 10-km race were also eliminated ($n = 11$). Finally, among the 884 remaining athletes, those with a performance in the 10-km below the lowest ranking of the FFA were eliminated (*i.e.*, performance > 50 and 60 min, respectively, for men and women). The analysis was thus performed with a dataset of 820 athletes. For each athlete, the sex (*i.e.*, woman vs man), date of birth (to calculate age), body mass and height (to calculate the body mass index: BMI), and race times (*i.e.*, the performances on the 10-km and marathon) were recorded.

Two supervised machines learning regression algorithms were used: ANN and KNN [24, 25, 31–33]. Both algorithms were implemented in R language. R software (version R X64 3.6.1 - R Development Core Team, Vienna, Austria) was used for our analysis, and the following R packages for machine learning approaches were used: dplyr version 0.8.3, neural net version 1.44.2 [34, 35].

All data were normalized to meet the requirement of the sigmoid transfer function (for ANN) and to remove the scale differences between the input variables (for KNN).

The data from the 820 athletes were randomly separated into a random train/test split for training and testing processes, respectively. The 90:10 ratio was used, meaning that 90% of the data was randomly selected for the training process (*i.e.*, 738 out of the 820 performances), while 10% of the data was randomly selected for the testing process (*i.e.*, 82 performances) [36, 37].

For both supervised machine learning algorithms, the same inputs (10-km race time, BMI, age and sex) were used to solve the linear regression problem, which consisted in estimating the value of the same continuous output (marathon race time). We also specified that exactly the same training ($n = 738$) and testing ($n = 82$) data were used for the two algorithms.

In ANN, a multilayer perceptron was used with four inputs and one output (Figure 1) [31]. In this network, the computing units are arranged into three layers, which are conveniently ordered. The information flows forward from the four neurons of the input layer to the two connecting neurons of the hidden layer and, finally, to the single neuron of the output layer using no backward connection. The first layer (the input layer) corresponds to the independent variables (*i.e.*, performance on 10-km, BMI, age and sex), while the third layer (the output layer) corresponds to the dependent variable score (marathon performance). The intermediate layer, which is the hidden layer, consists of all possible connections between the input and output layers and allows for the combined impact of a multiple set of independent variables on

the output layer. This ANN makes use of Rprop, which is short for resilient back propagation, a training technique without weight backtracking for supervised learning, [35]. The training stopping point (*i.e.*, threshold) was set at 0.01.

* Figure 1 here*

To test the advantage of the neural network algorithm, KNN was applied using the same four input variables and the same output variable for comparison with ANN. In this study, the KNN algorithm was tested by selecting the closest neighbors ($k = 3$). In other words, for each athlete of the testing dataset, we retained the three athletes of the training dataset having the smallest Euclidean distance (sum of the differences between their four respective inputs). The estimated output (marathon time) was calculated using the average of the marathon times of the three closest neighbors (athletes) weighted by the inverse of their respective Euclidean distances to the testing athlete.

This study was approved by the National Ethics Committee for Research in Sports Sciences (CERSTAPS 2019-22-02-31). The protocol for this study was legally declared, in accordance with the European General Data Protection Regulations.

Statistical Analysis

Standard statistical methods were used to calculate the means and standard deviations (SD).

The normal Gaussian distribution was verified by the Shapiro-Wilk test. A Student paired samples t-test was used for normally distributed data to compare the actual and predicted marathon performances for each machine learning algorithm. When these data did not pass the test for normality, a Wilcoxon signed-rank test was used. The magnitude of the differences was assessed by the effect size (ES), which was classified using the Cohen scale [38].

The association between the actual and predicted performances was tested with Pearson's product-moment or Spearman's rank order correlations, depending on whether the data

followed a normal Gaussian distribution. We considered a correlation of $r = 0.90$ or more as very high, between 0.70 and 0.89 as high, between 0.50 and 0.69 as moderate, and between 0.26 and 0.49 as low.

The coefficient of determination (r^2) and the mean absolute error (MAE) criteria based on a common 90:10 training/test data split were chosen to evaluate the success of the regression analysis via the two modeling regression processes. MAE is the average of absolute errors, with lower error values typically meaning the model is more accurate and the predictions closely match the actual values [8, 23]; r^2 determines the precision of the predictions and how well the model fits the data [39]. This also makes it easier to compare and evaluate the results [25, 37].

Moreover, the bias (*i.e.*, difference between actual and predicted performances, to assess accuracy) and 95% limits of agreement (95% LoA, *i.e.*, ± 1.96 SD, to assess precision) were computed according to the Bland-Altman method.

Finally, the KNN and ANN models were compared to determine which one performed better. The outputs of the predicted marathon performances were verified with actual marathon performances. The model was considered valid if the MAE was less than 5% and if the biases and approval limits were acceptable. The best model was selected based the lowest MAE and the highest accuracy (*i.e.*, from bias) and precision (*i.e.*, from LoA).

Statistical significance was set at $p < 0.05$ and all analyses were performed with the Statistical Package for the Social Sciences (release 20.0, Chicago, IL, USA).

Results

Means and SD of actual and predicted marathon performances are presented in Table 1.

* Table 1 here*

No difference was found between the actual and predicted performances for either algorithm ($p > 0.05$, Table 1). Moreover, the magnitude of the bias in these predicted performances was systematically trivial ($ES \leq -0.091$; Table 1).

All predicted running performances were correlated with the actual ones, with a very high correlation coefficient ($p < 0.001$, $r \geq 0.918$, Table 1).

The MAE, presented in Table 1, were, respectively, 11 min 16 s (*i.e.*, 5.6%) and 4 min 48 s (*i.e.*, 2.4%) for ANN and KNN.

The bias \pm 95% LoA are shown in Figures 2-3. The bias \pm 95% LoA of ANN and KNN were, respectively: 3 min 14 s \pm 28 min 30 s (*i.e.*, $1.5 \pm 14.1\%$) and -47 s \pm 14 min 21 s (*i.e.*, $-0.4 \pm 7.2\%$) (Table 1 and Figures 2-3).

Figures 2 & 3 here

Discussion

The objectives of this study were to test the validity of two supervised machine learning algorithms (ANN and KNN) and to compare them to determine which one was better at predicting marathon performances in terms of accuracy (from the MAE and the bias) and precision (from 95% LoA).

One of the main findings was that the two algorithms can indeed be considered valid, accurate and precise for predicting marathon performances, as all the results confirmed their validity for predicting performances from independent variables (*i.e.*, performance on 10-km, body mass index, age and sex), with both of them demonstrating a prediction accuracy above 94%. These results were comparable to those of another study [8] that showed 97% accuracy in elite runners with a local matrix completion machine learning technique.

The MAE was lower with KNN than with ANN, meaning that KNN was more accurate (up to 98% instead of 94%) and the predictions matched more closely to actual performances. The

fields of application differed (*i.e.*, classification algorithm), and it should be noted that the results were in accordance with those of the studies of Mustafa et al. [40] and Tamilarasi and Porkodi [41], who showed that KNN demonstrated better accuracy than ANN, and in disagreement with those of Peace et al. [32], Musa et al. [42] and Anyama et al. [43, 44], who showed better predictions with ANN than KNN. Whatever the domain, there is no real consensus on which algorithm model is best in terms of regression or classification, and each model has advantages and limitations [23, 26, 45]. Indeed, unlike ANN, KNN performs better on small data samples and does not risk falling into overlearning. Therefore, even though the data sample size (*i.e.*, 820 athletes) was relatively large in this study, ANN might have presented some limitations (*i.e.*, generalizing or the risk of over-fitting of the data), which would explain the greater accuracy of KNN in predicting marathon performances. In other words, the prediction results may be partly explained not only by the breadth of the data (*e.g.*, size of the dataset [46]) but also by all its parameters (*e.g.*, ratio used for training and testing datasets, number of hidden layers or learning rate for training a neural network, k in KNN, distance type in KNN...) and the practitioner's method (*i.e.*, the modeling procedure [24, 32, 33, 47], which defines the algorithms).

Moreover, the results revealed that the coefficients of correlation were, respectively, 0.918 and 0.982 for ANN and KNN and also that in 95% of the dependent variable score (*i.e.*, marathon performance), the biases from ANN and KNN were $3 \text{ min } 14 \text{ s} \pm 28 \text{ min } 30 \text{ s}$ (*i.e.*, $1.5 \pm 14.1\%$) and $-47 \text{ s} \pm 14 \text{ min } 21 \text{ s}$ (*i.e.*, $-0.4 \pm 7.2\%$), respectively. Therefore, the coefficient of correlation was slightly higher (*i.e.*, higher prediction precision) for KNN than ANN, and the bias and 95% LoA were lower with KNN (*i.e.*, higher accuracy).

Coquart et al. [9] showed that marathon performance could be predicted from the nomogram of Mercier et al. [48] with acceptable levels of accuracy and precision. Indeed, the authors found

a bias and 95% LoA of -1 min 25 s \pm 27 min 3 s (*i.e.*, - 0.7 \pm 13.2%) [9], which is comparable to ANN in the current study. However, for the same level of accuracy and precision, Coquart et al. [9] had the athletes perform two long-distance maximal performances (*i.e.*, 10-km and 20-km), while the predictions from ANN (and KNN, which is more accurate and precise) were obtained from only one performance (*i.e.*, 10-km). More recently, Vickers and Vertosick [12] explored several methods (*i.e.*, the Riegel formula [49] and two models based on one or two prior races, respectively) for predicting the marathon race times of recreational runners. Between the predicted and observed marathon times, they showed a mean square error (MSE) of 6 min 21 s for the Riegel formula [49], 3 min 48 s for the model based on one prior race, and 3 min 28 s for the model based on two prior races. It thus seems that in addition to the number of performances a model includes to predict performance, the association of certain variables (*e.g.*, sex, age, BMI) with race velocity also improves the accuracy of predictions. Moreover, it is especially interesting to note that in the current study, which integrated similar factors in the algorithms (*i.e.*, sex, age, BMI), the MSEs were lower, with an MSE of 2 min 39 s for ANN and an MSE of 39 s for KNN, by taking only one performance (*i.e.*, 10-km time). Therefore, it might be interesting to add another performance (*e.g.*, on the half-marathon) to ANN and KNN to see whether prediction accuracy/precision is significantly better than prediction methods not relying on AI. To limit athlete fatigue, ANN and especially KNN should nevertheless be preferred over other prediction methods that require two performances.

The main potential limitations of this study were the sample size and the content of our dataset. As mentioned in the discussion, KNN has the advantage of performing better on small data samples (no risk of falling into overlearning) than the ANN algorithm. Moreover, the choice and inclusion of input variables (limited, with only four variables, *i.e.*, 10-km performance, BMI, age and sex) structured the algorithms and influenced the prediction results. In other

words, the possibility of adjusting the algorithm to better model the problem domain will always be conceivable. Indeed, although marathon performance can be affected by a multitude of factors since the run involves several performance elements, physiological (*e.g.*, maximal oxygen uptake, running economy, anaerobic threshold... [16]), psychological (*e.g.*, motivation, stress) and environmental variables can also be determinant in long-distance running performance [4–6]. Thus, the other limitation of this study concerned the quality of the recovered data. For example, it is likely that the performance data (*i.e.*, 10-km time) was influenced by the race profile (*e.g.*, course profile: uphill, downhill...), meteorological conditions (*e.g.*, temperature, wind, rain...), opposition field or race strategy (*e.g.*, run to win or run for a time), but this information was not available.

The applications from the current study would be further extended by performing validation studies in race (*i.e.*, middle- or long-distance running on track or road), with other inputs that influence running performance (*e.g.*, maximal oxygen uptake, running economy and anaerobic threshold) in specific levels of runners (*i.e.*, subregional, regional, inter-regional, national and international). It would also be interesting to use machine learning techniques to identify the determinant factors of the marathon (*i.e.*, the use of classification techniques like random forest or naive Bayes) in order to help athletes, staff and professional sport analysts to design training programs and detect athletic talent, in addition to predicting results (marathon time).

Conclusion

Few studies on the prediction of running performance, especially for the marathon, have used artificial intelligence with ANN and/or KNN algorithms. The results of the current study demonstrated that both KNN and ANN were able to predict the performances of marathon runners with an acceptable level. Both models were valid and able to attain a prediction

accuracy above 94%, although KNN appears to be superior to ANN as it accurately predicted marathon performance above 98%. These approaches can therefore be used to predict performances over the course of appropriate training programs, with particular emphasis on prescribing speeds for tempo runs and determining competitive strategies. Future studies should be directed toward the use of machine learning techniques to gain insight into other parameters that impact marathon performance by means of classification techniques in order to detect talented athletes, for example, and not only to predict marathon performance.

Disclosure of interest

None of the authors had any conflict of interest.

Acknowledgments

The authors are grateful to the French Athletics Federation (Fédération Française d’Athlétisme) for data collection, and the Orthodynamica Center at Mathilde Hospital 2 for funding.

References

- [1] Burfoot A. The history of the marathon. *Sports Med* 2007; 37: 284–287.
- [2] Vitti A, Nikolaidis PT, Villiger E, et al. The “New York City Marathon”: participation and performance trends of 1.2M runners during half-century. *Res Sports Med* 2019; 0: 1–17.
- [3] Knechtle B, Di Gangi S, Rüst CA, et al. Men’s Participation and Performance in the Boston Marathon from 1897 to 2017. *Int J Sports Med* 2018; 39: 1018–1027.
- [4] Marc A, Sedeaud A, Guillaume M, et al. Marathon progress: demography, morphology and environment. *J Sports Sci* 2014; 32: 524–532.
- [5] El Helou N, Tafflet M, Berthelot G, et al. Impact of environmental parameters on marathon running performance. *PloS One* 2012; 7: e37407.
- [6] Nikolaidis PT, Di Gangi S, Chtourou H, et al. The Role of Environmental Conditions on Marathon Running Performance in Men Competing in Boston Marathon from 1897 to 2018. *Int J Environ Res Public Health*; 16. Epub ahead of print 20 2019. DOI: 10.3390/ijerph16040614.
- [7] Boullosa D, Esteve-Lanao J, Casado A, et al. Factors Affecting Training and Physical

Performance in Recreational Endurance Runners. *Sports* 2020; 8: 35.

[8] Blythe DAJ, Király FJ. Prediction and quantification of individual athletic performance of runners. *PloS One* 2016; 11: e0157257.

[9] Coquart J, Alberty M, Bosquet L. Validity of a nomogram to predict long distance running performance. *J Strength Cond Res* 2009; 23: 2119–2123.

[10] Berndsen J, Smyth B, Lawlor A. Pace my race: recommendations for marathon running. In: *Proceedings of the 13th ACM Conference on Recommender Systems*. 2019, pp. 246–250.

[11] Ruiz-Mayo D, Pulido E, Martınoz G. Marathon performance prediction of amateur runners based on training session data. *Proc Mach Learn Data Min Sports Anal* 2016; 8.

[12] Vickers AJ, Vertosick EA. An empirical study of race times in recreational endurance runners. *BMC Sports Sci Med Rehabil* 2016; 8: 26.

[13] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. From empirical models to physiological models. *Sports Med Auckl NZ* 1999; 27: 359–379.

[14] Ettema JH. Limits of human performance and energy-production. *Int Z Für Angew Physiol Einschließlich Arbeitsphysiologie* 1966; 22: 45–54.

[15] Scherrer J, Monod H. Le travail musculaire local et la fatigue chez l'homme. *Presse Med* 1960; 68: 1717–1717.

[16] Alvero-Cruz JR, Carnero EA, García MAG, et al. Predictive Performance Models in Long-Distance Runners: A Narrative Review. *Int J Environ Res Public Health*; 17. Epub ahead of print 9 November 2020. DOI: 10.3390/ijerph17218289.

[17] Péronnet F, Thibault G. [Physiological analysis of running performance: revision of the hyperbolic model]. *J Physiol (Paris)* 1987; 82: 52–60.

[18] Tanda G. Prediction of marathon performance time on the basis of training indices. *J Hum Sport Exerc* 2011; Volume 6: pag. 521-520.

[19] Vandewalle H. Modelling of running performances: comparisons of power-law, hyperbolic, logarithmic, and exponential models in elite endurance runners. *BioMed Res Int* 2018; 2018: 8203062.

[20] Florence SL, Weir JP. Relationship of Critical Velocity to Marathon Running Performance.: 43. *Med Sci Sports Exerc* 1995; 27: S8.

[21] Mulligan M, Adam G, Emig T. A minimal power model for human running performance. *PloS One* 2018; 13: e0206645.

[22] Huang Z-Q, Chen Y-C, Wen C-Y. Real-Time Weather Monitoring and Prediction Using City Buses and Machine Learning. *Sensors*; 20. Epub ahead of print 10 September 2020. DOI: 10.3390/s20185173.

[23] Uddin S, Khan A, Hossain ME, et al. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak* 2019; 19: 281.

- [24] Bunker RP, Thabtah F. A machine learning framework for sport result prediction. *Appl Comput Inform* 2019; 15: 27–33.
- [25] Wiseman O. *Using Machine Learning to Predict the Winning Score of Professional Golf Events on the PGA Tour*. National College of Ireland, 2016.
- [26] Bunker R, Susnjak T. The Application of Machine Learning Techniques for Predicting Results in Team Sport: A Review. *Prepr Submitt ArXiv* 2019; 1–48.
- [27] McCabe A, Trevathan J. Artificial intelligence in sports prediction. In: *Fifth International Conference on Information Technology: New Generations (itng 2008)*. IEEE, 2008, pp. 1194–1197.
- [28] Claudino JG, de Oliveira Capanema D, de Souza TV, et al. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. *Sports Med-Open* 2019; 5: 28.
- [29] Velichkov B, Koychev I, Boytcheva S. Deep learning contextual models for prediction of sport event outcome from sportsman’s interviews. In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*. 2019, pp. 1240–1246.
- [30] Dhanabal S, Chandramathi S. A review of various k-nearest neighbor query processing techniques. *Int J Comput Appl* 2011; 31: 14–22.
- [31] Maszyk A, Golaś A, Pietraszewski P, et al. Application of Neural and Regression Models in Sports Results Prediction. *Procedia - Soc Behav Sci* 2014; 117: 482–487.
- [32] Peace IC, Uzoma AO, Ita SA, et al. A Comparative Analysis of K-NN and ANN Techniques in Machine Learning. *Int J Eng Res Technol* 2015; 4: 420–425.
- [33] Ruiz JR, Ramirez-Lechuga J, Ortega FB, et al. Artificial neural network-based equation for estimating VO₂max from the 20 m shuttle run test in adolescents. *Artif Intell Med* 2008; 44: 233–245.
- [34] Renganathan V. Overview of artificial neural network models in the biomedical domain. *Bratisl Lek Listy* 2019; 120: 536–540.
- [35] Fritsch S, Guenther F, Suling M, et al. *Package ‘neuralnet’: Training of Neural Networks, 2016*.
- [36] Malone C, Fennell L, Folliard T, et al. Using a neural network to predict deviations in mean heart dose during the treatment of left-sided deep inspiration breath hold patients. *Phys Medica PM Int J Devoted Appl Phys Med Biol Off J Ital Assoc Biomed Phys AIFB* 2019; 65: 137–142.
- [37] Smith PF, Ganesh S, Liu P. A comparison of random forest regression and multiple linear regression for prediction in neuroscience. *J Neurosci Methods* 2013; 220: 85–91.
- [38] Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1: 98–101.
- [39] Lin L, Torbeck LD. Coefficient of accuracy and concordance correlation coefficient: new statistics for methods comparison. *PDA J Pharm Sci Technol* 1998; 52: 55–59.

- [40] Mustafa M, Taib MN, Murat Z, et al. Comparison between KNN and ANN Classification in Brain Balancing Application via Spectrogram Image. *J Comput Sci Comput Math* 2012; 2: 17–22.
- [41] Tamilarasi R, Porkodi DR. A study and analysis of disease prediction techniques in data mining for healthcare. *Int J Emerg Res Manag Technoly ISSN* 2015; 1: 2278–9359.
- [42] Musa RM, Majeed AA, Taha Z, et al. The application of Artificial Neural Network and k-Nearest Neighbour classification models in the scouting of high-performance archers from a selected fitness and motor skill performance parameters. *Sci Sports* 2019; 34: e241–e249.
- [43] Anyama OU, Nwachukwu EO. A hybrid prediction system for american NFL results. *Int J Comput Appl Technol Res* 2015; 4: 42–47.
- [44] Anyama OU, Igiri CP. An application of linear regression & artificial neural network model in the NFL result prediction. *Int J Eng Res Technol* 2015; 4: 457–461.
- [45] Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod* 2019; 20: 41.
- [46] Ajiboye AR, Abdullah-Arshah R, Hongwu Q. Evaluating the effect of dataset size on predictive model using supervised learning technique. *Int J Softw Eng Comput Sci IJSECS* 2015; 1: 75–84.
- [47] Yang L, Shami A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* 2020; 415: 295–316.
- [48] Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le VO₂max et l'endurance relative en course de fond. *Méd Sport* 1984; 58: 181–187.
- [49] Riegel PS. Athletic records and human endurance: A time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. *Am Sci* 1981; 69: 285–290.

Prédiction de la performance en course à pied : demi-fond et fond

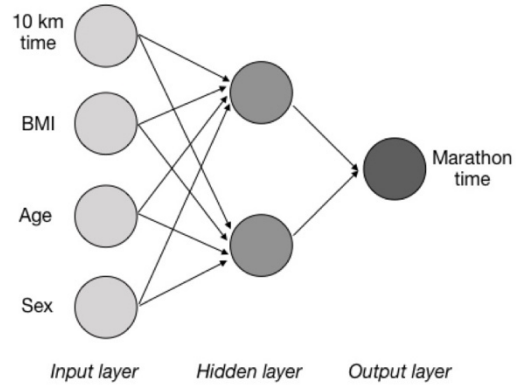
Table 1. Mean values and standard deviation (SD) of actual and predicted performances on marathon from each algorithm (*i.e.*, artificial neuronal network (ANN) and k-nearest neighbors (KNN)) in 82 performances athletes (*i.e.*, 10% of the data is selected for testing process), difference between actual and predicted performances (p), magnitude of the difference, Pearson's product-moment with actual performance (r), bias and SD (min), bias and 95% LoA (min and %) and mean absolute error (MAE) (%).

Distance		Mean performance (min) \pm SD	p	Magnitude of difference		Correlation		Bias (min) \pm 95% LoA	Bias (%) \pm 95% LoA	MAE (%)
				ES	Interpretation	r	Interpretation			
	Actual	199.75 \pm 36.01								
Marathon	Predicted from the ANN	202.77 \pm 30.24	0.063	-0.091	Trivial	0.918*	Very high	3.023 \pm 28.492	1.5 \pm 14.1	5.6
	Predicted from the KNN	198.96 \pm 32.74	0.333	-0.024	Trivial	0.982*	Very high	-0.79 \pm 14.35	-0.4 \pm 7.2	2.4

Note: Bias = difference between actual and predicted performances; r = coefficient of correlation; LoA = limits of agreement; ES = effect size.

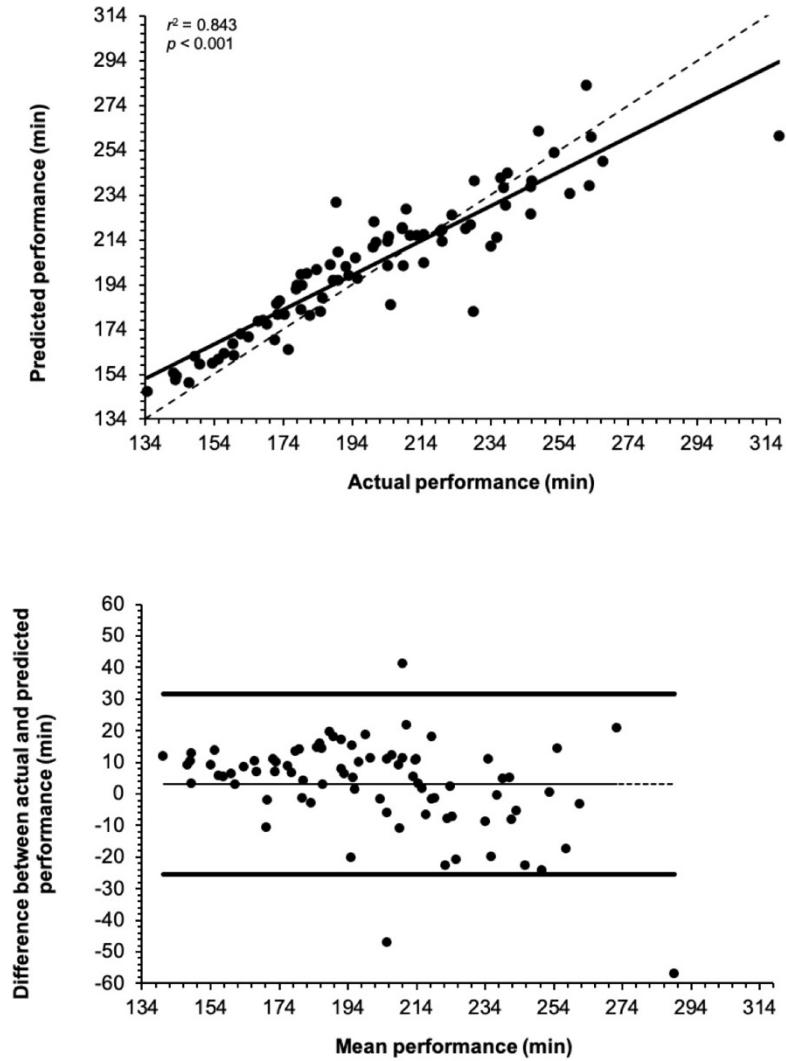
*Significantly correlated at $p < 0.001$

Figure 1. Neural network architecture



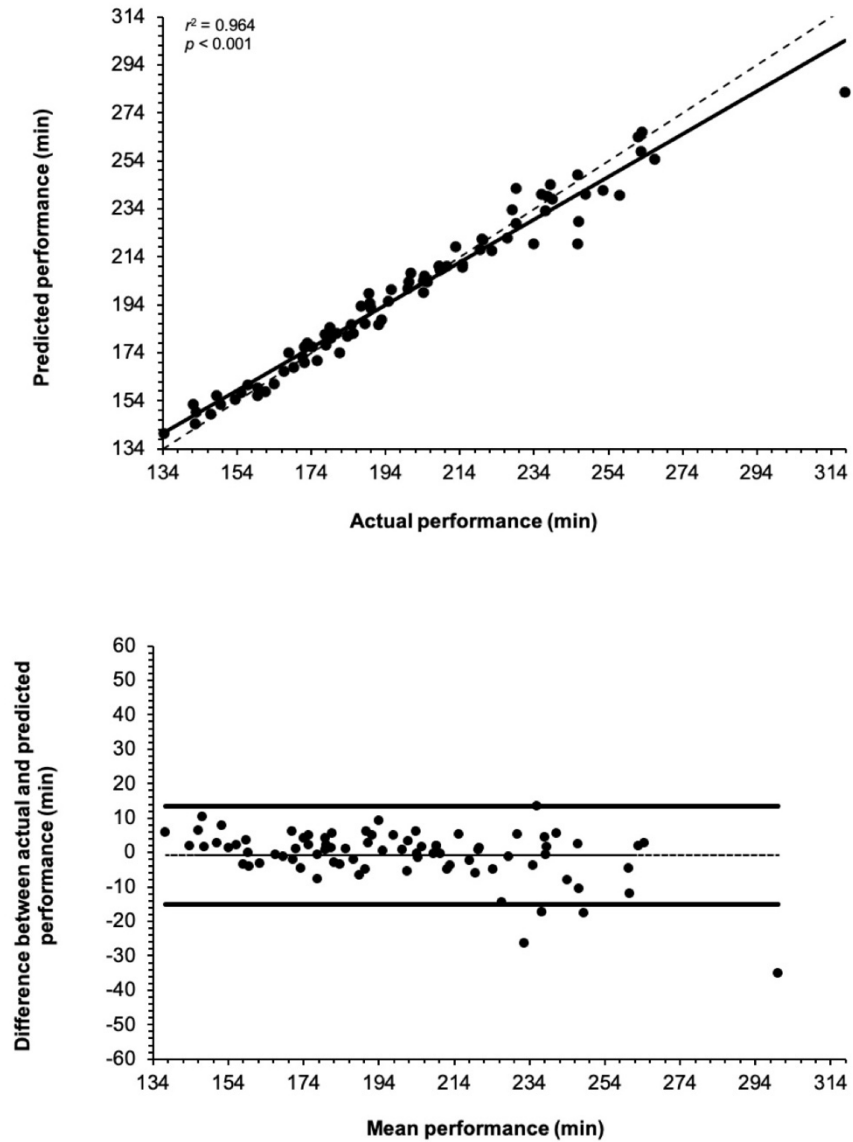
Note: BMI: body mass index

Figure 2. Validity of measurements with the ANN algorithm to predict performance.



Top panel: association between actual and predicted performance from the ANN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel:* Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95% limits of agreement.

Figure 3. Validity of measurements with the KNN algorithm to predict performance.

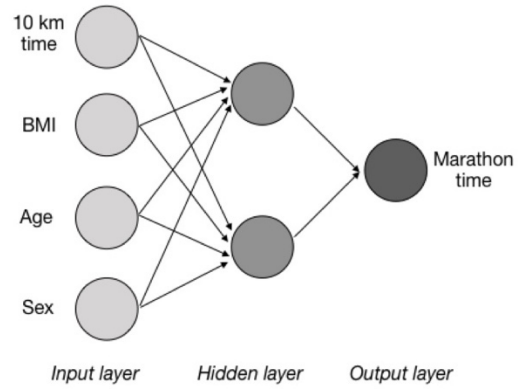


Top panel: association between actual and predicted performance from the KNN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel:* Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95% limits of agreement.

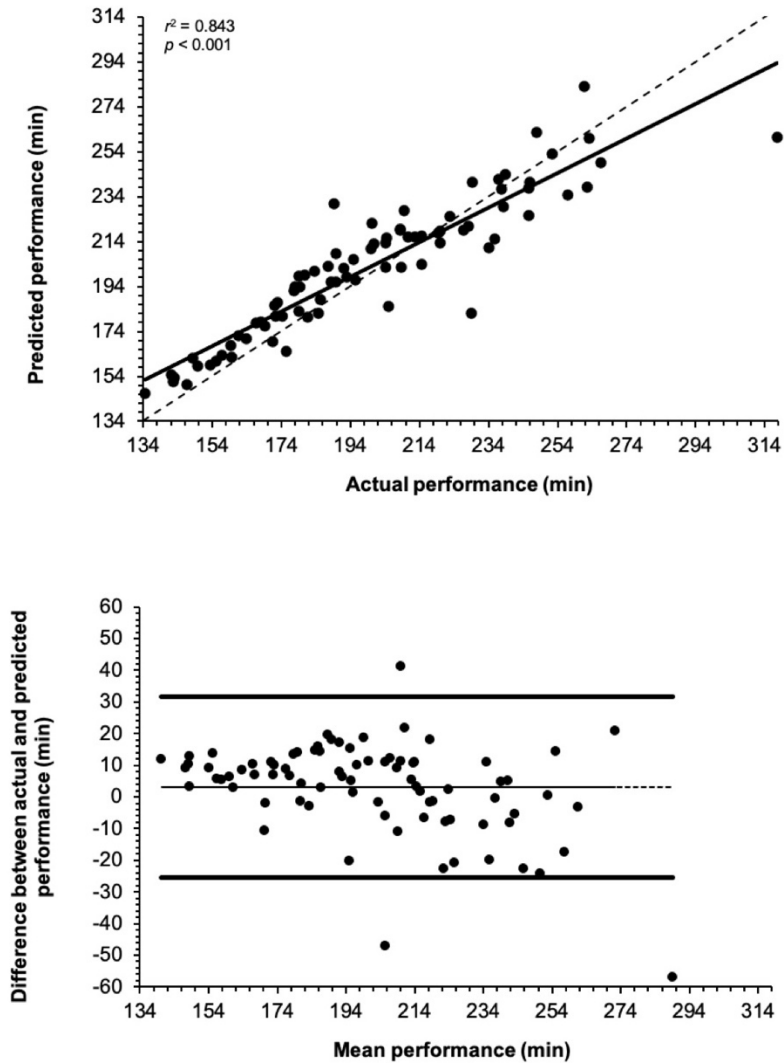
Figure1

[Click here to access/download;Figure;Figure 1.docx](#) 

Figure 1. Neural network architecture

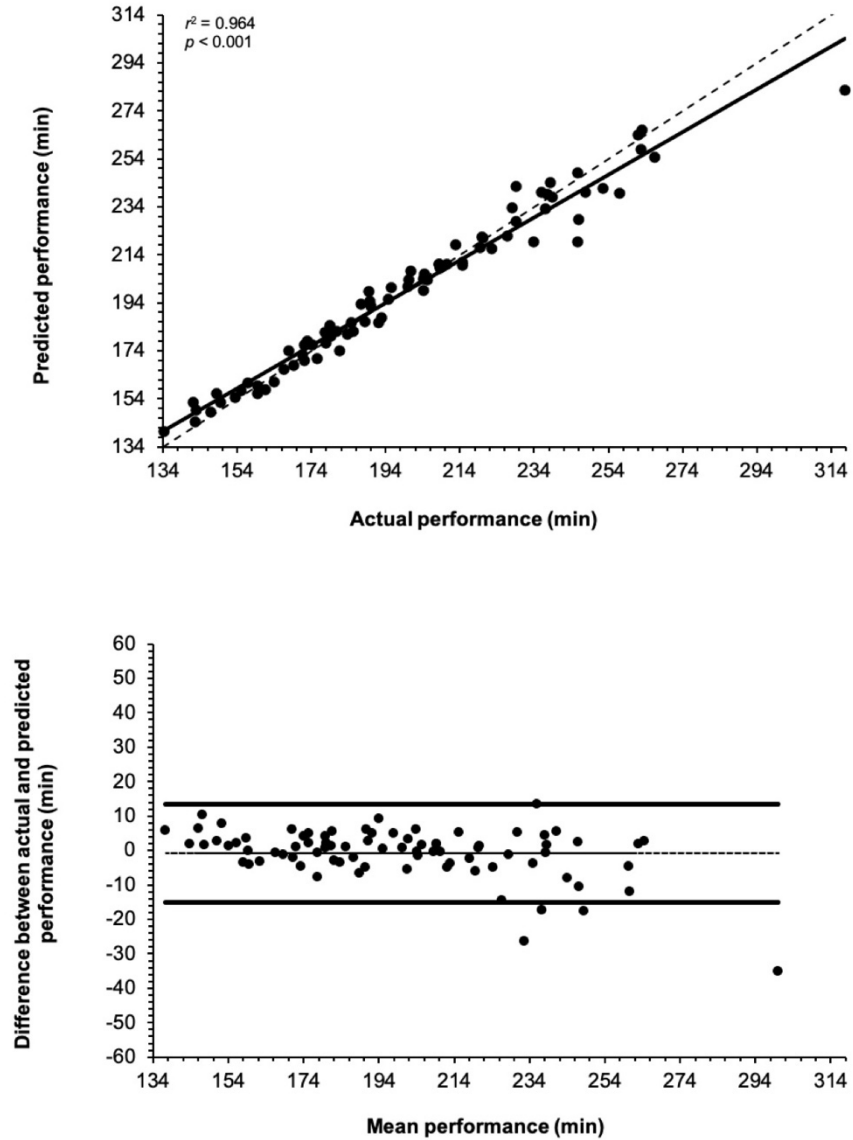


Note: BMI: body mass index

Figure 2. Validity of measurements with the ANN algorithm to predict performance.

Top panel: association between actual and predicted performance from the ANN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel:* Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95% limits of agreement.

Figure 3. Validity of measurements with the KNN algorithm to predict performance.



Top panel: association between actual and predicted performance from the KNN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel:* Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95% limits of agreement.

Prédiction de la performance en course à pied : demi-fond et fond

Table1

[Click here to access/download;Table;Table 1.docx](#)

Table 1. Mean values and standard deviation (SD) of actual and predicted performances on marathon from each algorithm (*i.e.*, artificial neuronal network (ANN) and k-nearest neighbors (KNN)) in 82 performances athletes (*i.e.*, 10% of the data is selected for testing process), difference between actual and predicted performances (p), magnitude of the difference, Pearson's product-moment with actual performance (r), bias and SD (min), bias and 95% LoA (min and %) and mean absolute error (MAE) (%).

Distance		Mean performance (min) \pm SD	p	Magnitude of difference		Correlation		Bias (min) \pm 95% LoA	Bias (%) \pm 95% LoA	MAE (%)
				ES	Interpretation	r	Interpretation			
	Actual	199.75 \pm 36.01								
Marathon	Predicted from the ANN	202.77 \pm 30.24	0.063	-0.091	Trivial	0.918*	Very high	3.023 \pm 28.492	1.5 \pm 14.1	5.6
	Predicted from the KNN	198.96 \pm 32.74	0.333	-0.024	Trivial	0.982*	Very high	-0.79 \pm 14.35	-0.4 \pm 7.2	2.4

Note: Bias = difference between actual and predicted performances; r = coefficient of correlation; LoA = limits of agreement; ES = effect size.
*Significantly correlated at $p < 0.001$

Annexe 7. Article 7 : *Predicting the potential of middle- and long- distance track runners by the evaluation of performance improvement rate*

1

Original investigation

2

3 **Predicting the potential of middle- and long-distance track runners by the evaluation of**
4 **performance improvement rate**

5

6 **Running head:** Predicting the potential of the middle- and long-distance track runners

7

8 **Authors:** Lucie Lerebourg^{1,2}, Brice Guignard¹, Maxime L'Hermette¹, and Jeremy B. Coquart^{1,3}

9

10 **Affiliations:**

11 1. *Normandie Univ, UNIROUEN, Centre d'Études des Transformations des Activités*
12 *Physiques et Sportives, F-76000, Rouen, France*

13 2. *Orthodynamica Center, Mathilde Hospital 2, F-76000, Rouen, France*

14 3. *Univ. Lille, Univ. Artois, Univ. Littoral Côte d'Opale, ULR 7369 - URePSSS - Unité de*
15 *Recherche Pluridisciplinaire Sport Santé Société, F-59000 Lille, France*

16

17 **Abstract word count:** 226words

18 **Text-only word count:** 4406 words (main text)

19 **Number of tables/figures:** 3 tables and 3 figures

20

21 **Corresponding Author:**

22 Lucie Lerebourg, Centre d'Études des Transformations des Activités Physiques et Sportives,

23 Boulevard Siegfried, 76 821 Mont Saint Aignan Cedex, France. Phone: (+33)648157388. E-

24 mail: lucie.lerebourg@orthodynamica.com

25 **Abstract**

26 **Purpose:** The aims of this study were to highlight the understanding of personal best
27 achievement through an analysis of the multiple performances per season of the athletes
28 composing the national-top 10 in middle- and long-distance track events and to propose an
29 index to assess this optimal performance improvement rate that would be used as a tool for elite
30 coaches and athletes willing to evaluate the progression of athletes, predict potential, and
31 possibly detect high potential athletes. **Methods:** The top 10 National athletes in the *Fédération*
32 *Française d'Athlétisme* (FFA) ranking of 800-m, 1500-m and 5000-m events for each sex were
33 assessed for their history of race times before achieving their BRT. For each athlete, the
34 difference between the actual BRT and BRT in the last season was defined as the 1st season
35 improvement rate (1-SIR), whereas the average improvement rate in the last and preceding
36 seasons was the multi-season improvement rate (M-SIR). **Result:** There were sex differences
37 in the 5000-m only for the M-SIR (%) ($p < 0.05$) and significant higher rate of improvement
38 for 1-SIR (min) and M-SIR for the longer distance mainly (5000-m) than the shorter distances
39 (800-m and 1500-m). **Conclusions:** Prediction using this index may be used to detect talents
40 (those who have the potential for high-level performances and/or the potential to break national
41 records) and to evaluate and optimize training strategy and competitions.

42 **Keywords:** prediction, running, sports detection, athletes, estimation.

43

44 **Introduction**

45 Analyze improvements and predict sport performance may be of great interest for athletes and
46 coaches (Coquart et al., 2015; Lerebourg & Coquart, 2020; Tabben et al., 2016). Running
47 performance is evolving and athletes are in constant search of records (Berthelot et al., 2015;
48 Gembris et al., 2007; Lerebourg & Coquart, 2020). The analysis and prediction of track-running
49 performance, and therefore the potential performance, is useful to detect a future champion and
50 prescribing them speeds for tempo runs to progress or to break a personal record or the race
51 record (national, meeting, world or Olympic record) in competition (Blythe & Király, 2016;
52 Coquart et al., 2015; Tabben et al., 2016). It could also allow to confirm the athletes' specialty
53 or to help them to direct to another specialty as a function of the degree of expected annual
54 progress over a chosen distance.

55 Several authors have studied the progression in performances and identify the trainable
56 variables that can be improved through training (*i.e.*, maximal oxygen uptake, endurance,
57 strength, psychological states, tactics, pacing strategy) and other factors that are beyond the
58 athlete's control (*e.g.*, demographics, climate conditions, seasonal characteristics) that might
59 influence successful the performances (Berthelot et al., 2015; S. N. Cheuvront et al., 2005;
60 Haugen et al., 2018). Some authors have also sought to examine the progression of human
61 performance over the decades, comparing sexes (Berthelot et al., 2015; S. N. Cheuvront et al.,
62 2005; Coast et al., 2004; Haugen et al., 2018) and/or distances of race events in performance
63 (Berthelot et al., 2015; Coast et al., 2004; Haugen et al., 2018; Lerebourg & Coquart, 2020).
64 Haugen et al. (Haugen et al., 2018) observed, for example, a greater improvement in women
65 than men corresponding in middle- and long-distance track running at $0.4, \pm 0.2\%$ between
66 2002 to 2016. In general, while the results show a greater increase in performance for women
67 compared to men in recent years, the results also show that running performances have reached
68 a peak in numerous disciplines (Berthelot et al., 2015; S. N. Cheuvront et al., 2005). However,
69 the most studies that could help to better understand the performances and changes of
70 performances between the sexes and distances in middle- and long-distance track running
71 concerned mainly world-class elite and men.

72 The middle- and long-distance track running, including distances from 800-m to 5000-m,
73 requires a high metabolic contribution, that evolves according to the distance or the duration,
74 with a mix of anaerobic/aerobic contributions (Duffield et al., 2005a, 2005b; Kenney &
75 Hodgson, 1985; Spencer et al., 1996). Indeed, longer distances such as the 5000-m require a

76 higher contribution of oxidative phosphorylation pathway while shorter distances, such as the
77 800-m and 1500-m, are characterized by higher preponderance of the “glycolytic pathway” to
78 maintain the speed (Chamari & Padulo, 2015; Duffield et al., 2005a, 2005b; Kenney &
79 Hodgson, 1985; Spencer et al., 1996). Moreover, beyond the physiological parameters
80 characterizing human capacities, with training and over the long-term, it seems possible to reach
81 peak performance at older ages in the longest distances (e.g., 5000-m) than in the shortest
82 distances (e.g., 800-m and 1500-m) (Haugen et al., 2018; Hollings et al., 2014; Lerebourg &
83 Coquart, 2020; Schulz & Curnow, 1988; Tanaka & Seals, 2008). Therefore, regarding the peak
84 performance ages, being able to identify and predict the potential of talented young runners in
85 the aim to achieve in middle-distance and in longer-term performances in long-distance, can be
86 of great interest to coaches and athletes. A recent study of Santos et al. (Santos et al., 2020),
87 calculating an improvement rate, proposed a new method to evaluate whether a training strategy
88 induced an optimal performance progression. For example, to break the athletic world record
89 (WR) in the 400-m, improvements of 0.97 to 6.83% and 0.41 to 5.85% per year (including
90 performances achieved at least 5 years before the WR) were indicated for women and men,
91 respectively (Santos et al., 2020). However, only sprint events (i.e., 100-m and 400-m) in world-
92 class elite athletes have been studied by Santos et al. (Santos et al., 2020), and it does not appear
93 to be any evaluation method in the literature for middle- and long-distance track events (i.e.,
94 800-m, 1500-m and 5000-m), which would provide coaches with an alternative tool for the
95 follow-up of their athletes.

96 In the first place, the main objective of this study was to highlight the understanding of personal
97 best achievement through an analysis of the multiple performances per season of the athletes
98 composing the national-top 10 in middle- and long-distance track events (i.e., 800-m, 1500-m,
99 and 5000-m). Secondly, this study aimed to propose, as a practical application, an improvement
100 index, in order to evaluate the progression of athletes, predict potential, and/or possibly detect
101 high potential athletes.

102

103 **Methods**

104 *1. Experimental approach*

105 Data were extracted from publicly available database of the French Athletics Federation
106 (*Fédération Française d’Athlétisme*: FFA). The characteristics and performances of the first
107 ten women and men athletes in the official rankings all-time (*i.e.*, of all years) of 800-m, 1500-
108 m and 5000-m events were collected. For each athlete, the date of birth, nationality, best race
109 time of the career: BRT (*i.e.*, personal best, the one that got the athlete in their top 10 National
110 ranking position), BRT in the season immediate before the BRT (BRT-1), and the BRT in the
111 last seasons before the BRT (BRTs) were recorded (Figure 1). The outliers, identified by box
112 plot diagrams and performances that were not achieved in the continuity of the seasons
113 preceding the BRT were excluded.

114 *Figure 1*

115
116 This study was approved by the National Ethics Committee for Research in Sports Sciences
117 (CERSTAPS 2019-22-02-31).

118 2. Participants

119 The performances of 60 athletes (*i.e.*, top 10 National of women and men performing over the
120 three selected distances) were considered (Table 1). The ranges of years of BRT (in line with
121 the all-time national top-10) retained by sex and distance were 1986 to 2020, 1987 to 2010 and
122 1994 to 2014 on 800-m, 1500-m and 5000-m respectively for women; and 1979 to 2020, 1995
123 to 2019 and 1995 to 2019 on 800-m, 1500-m and 5000-m, respectively, for men (Table 1). For
124 women, these performances in the 800-m, 1500-m and 5000-m represented, respectively, $95 \pm$
125 1% , $94 \pm 1\%$ and $93 \pm 1\%$ of the current outdoor WR (*i.e.*, 1 min 53 s by Kratochvilova
126 Jarmilain in 1983; 3 min 50 s by Dibaba Genzebe in 2015; and 14 min 6 s by Gidey Letesenbet
127 in 2020). For men, these performances represented, respectively, $97 \pm 1\%$, $97 \pm 1\%$ and $95 \pm$
128 1% of the current outdoor world records (*i.e.*, 1 min 40 s by Lekuta Rudisha David in 2012; 3
129 min 26 s by El Guerrouj Hicham in 1998; and 12 min 35 s by Cheptegei Joshuain in 2020).

130 *Table 1 here*

131 2.3. Design and Procedures

132
133 In accordance with the Santos’ study (Santos et al., 2020), the difference between the
134 BRT and the BRT-1 corresponded to the 1-season improvement rate (1-SIR), whereas the
135 multi-season improvement rate (M-SIR) represented to the average of improvement rate in the
136 season immediately before and seasons before as shown in Figure 1. These variables (1-SIR

137 and M-SIR) were calculated individually. Then, these measures were reported as a relative
138 value (%). Descriptive analysis (mean values, standard deviation (SD) and range) of all data
139 was performed and is displayed in Table 2 and Figure 2.

140

141 *Table 2 here*

142

143 *Figure 2*

144

145 2.4. Statistical Analysis

146

147 Standard statistical methods were used to calculate the means and SD. Normal Gaussian
148 distribution was verified by the Shapiro-Wilk test and homogeneity of variance by the Levene
149 test. A two-way analysis of variance (ANOVA) for independent measures was used for
150 normally distributed data to compare the dependent variable (*i.e.*, age, number of performances,
151 BRT, BRTs, 1-SIR or M-SIR) according to sex (men vs women) and distance (800-m vs 1500-
152 m vs 5000-m). When these data did not pass the test for normality, a Kruskal-Wallis test was
153 used. Then, the post-hoc analyses or Mann-Whitney tests were used to identify the significant
154 differences for parametric data and nonparametric data, respectively. Statistical significance
155 was set at $p < 0.05$. If post-hoc were conducted, a Bonferroni correction was applied by
156 considering the number of comparisons. For parametric data, the magnitude of the differences
157 was assessed by effect size (ES) and classified using the Cohen scale (Cohen, 1992). It was
158 considered trivial ($ES < 0.2$), small ($0.2 \leq ES < 0.5$), moderate ($0.5 \leq ES < 0.79$), and large (ES
159 ≥ 0.8). For nonparametric data, the magnitude of difference was assessed by the rank biserial-
160 correlation (r_B), that can be considered as an ES and is viewed as an equivalent to the correlation
161 coefficient (Goss-Sampson, 2019). It was considered trivial ($r_B < 0.1$), small ($0.1 \leq r_B < 0.3$),
162 moderate ($0.3 \leq r_B < 0.5$), and large ($r_B \geq 0.5$). Furthermore, in agreement with the Santos' study
163 (Santos et al., 2020), using the descriptive data of M-SIR (in min) a list of one thousand random
164 entries was generated in each group (800-m, 1500-m and 5000-m; women and men) to compose
165 a Gauss curve for the index. Although the generated numbers are random, the mean (*i.e.*, central
166 value) of each curve and the distance (*i.e.*, value of SD) of M-SIR are real, which will serve for
167 the purpose of this manuscript. In particular, the value of SD will indicate whether a certain
168 athlete is improving more or less than the best athletes of the event. All analyses were performed
169 with Jeffrey's Amazing Statistics Program (JASP Team, version 0.14.1, computer software).

170

171 **Results**

172 Mean values and SDs of variables (*i.e.*, age, number of performances, BRT, BRTs, 1-SIR, and
173 M-SIR) are presented in Table 2 and Figure 2.

174
175 *Age*

176 There was significant main effect for distance for the entire sample of runners (men and women,
177 without distinction between sex) only ($p = 0.001$). Post-hoc analysis showed that 800-m runners
178 were younger than 1500-m and 5000-m runners respectively ($p < 0.01$) with a large magnitude
179 of difference ($ES = 0.967$ and $ES = 1.221$, respectively). There was no significant effect between
180 sexes ($p = 0.433$) and no significant interaction between the effects of sex and distance ($p =$
181 0.796).

182 *Number of performances registered*

183 There was no significant effect for sex ($p = 0.130$) but significant difference for distance event
184 ($p = 0.017$). Post-hoc analysis showed that number of performances registered in women was
185 significantly higher in 800-m compared to 5000-m event ($p = 0.012$) with a large magnitude of
186 difference ($r_B = 0.670$).

187 *BRT*

188 There were significant main effects for both sexes ($p = 0.026$) and distance ($p < 0.01$). Post-hoc
189 analysis showed that BRT were significantly lower in men compared to women, regardless of
190 the distance event ($p < 0.001$) with a large magnitude of difference ($r_B = 1.000$) and revealed
191 that BRT were significantly different between each distance, for each sex ($p < 0.01$) with a large
192 magnitude of difference ($r_B = 1.000$).

193 *BRTs*

194 There were significant main effects for both sexes ($p = 0.027$) and distance ($p < 0.01$). Post-hoc
195 analysis showed that BRTs were significantly lower in men compared to women regardless of
196 the distance event ($p < 0.01$) with a large magnitude of these differences ($r_B = 1.000$), and
197 revealed that BRTs were significantly different between each distance, for each sex ($p < 0.01$)
198 with also a large magnitude of these differences ($r_B = 1.000$).

199

200 *1-SIR (min)*

201 There was no significant effect for sex ($p = 0.159$) but significant effect for distance ($p < 0.01$).
202 Post-hoc analysis showed that 1-SIR (min) was significantly higher in 5000-m compared to
203 800-m and 1500-m events, for each sex ($p < 0.01$) with a large magnitude of difference between
204 distances ($r_{B800-5000} = 1.000$ and $r_{B1500-5000} = 1.000$ for women; $r_{B800-5000} = 0.989$ and $r_{B1500-5000}$
205 $= 0.944$ for men) .

206

207 *1-SIR (%)*

208 There was significant effect for sex ($p = 0.036$). Post-hoc analysis revealed that 1-SIR (%) was
209 higher in women compared to men in 5000-m event only ($p = 0.021$). However, after a
210 Bonferroni correction, considering the number of comparisons, this result appeared not
211 significant, although the magnitude of the difference between the men's 5000-m event and the
212 women's 5000m race was large ($r_B = 0.741$). There was no significant effect for distance ($p =$
213 0.361).

214

215 *M-SIR (min)*

216 There was no significant effect for sex ($p = 0.297$) but significant effect for distance ($p < 0.01$).
217 Post-hoc analysis showed that M-SIR (min) were significantly higher in 5000-m compared to
218 800-m and 1500-m events for each sex ($p < 0.01$) with a large magnitude of difference between
219 distances ($r_{B800-5000} = 1.000$ and $r_{B1500-5000} = 0.926$ for women; $r_{B800-5000} = 0.944$ and $r_{B1500-5000}$
220 $= 0.764$ for men) and significantly higher in 1500-m compared to 800-m for women ($p = 0.012$)
221 with a large magnitude of difference ($r_B = 0.704$).

222 *M-SIR (%)*

223 There was significant effect for sex ($p = 0.045$). Post-hoc analysis showed that M-SIR (%) were
224 significantly higher in 5000-m for women compared to men ($p = 0.008$) and the magnitude of
225 the difference between the men's 5000-m event and the women's 5000m race was large ($r_B =$

226 0.815). There was no significant effect for distance ($p = 0.712$). The index for optimal
227 performance improvement rate assessment was generated from the actual data of the top 10
228 National best athletes of each category (women, men) in the events of 800-m, 1500-m and
229 5000-m and are shown in Figure 3. The results of the current study suggest that for a woman to
230 reach the 800-m national record, she needs to improve her performance annually from 1.02 to
231 3.06 seconds. In contrast, a man usually needs to improve his performance from 1.02 to 2.58
232 seconds. In 1500-m, women generally must improve annually from 1.32 to 9.36 seconds and
233 men from 1.62 to 4.5 seconds. In 5000-m, women generally must improve annually from 7.14
234 to 48.9 seconds and men from 3.72 to 11.52 seconds. A practical approach of these results is
235 provided in more details in the section on practical applications.

236 *Figure 3 here*

237 **Discussion**

238 The main objective of this study was to highlight the understanding of personal best
239 achievement through an analysis of the multiple performances per season of the athletes
240 composing the national-top 10 in middle- and long-distance track events (*i.e.*, 800-m, 1500-m,
241 and 5000-m). Then, the second objective was to propose, as a practical application, an
242 improvement index, to evaluate the progression of athletes, predict potential, and possibly
243 detect high potential athletes.

244 While the overall results of this current study showed a higher rate of progression for women
245 compared to men, the sex comparison was significant for the M-SIR (%) in the 5000-m event
246 only. Indeed, the results showed that women improved performance annually before reaching
247 the all-time national-top 10 and their BRT by an average of 3.10% from the last season's record
248 while men had an average of 0.97% improvement.

249 Naturally and in accordance with previous studies (S. N. Cheuvront et al., 2005; Coast et al.,
250 2004; Coquart et al., 2015; Haugen et al., 2018), the results of this current study showed that
251 for the same running distances (*i.e.*, 800-m, 1500-m and 5000-m), women have significantly
252 lower performances (*i.e.*, BRT and BRTs) than men, probably because of their lower aerobic
253 capacity (*e.g.*, $\dot{V}O_2\text{max}$) and differences in biological determinants (*e.g.*, lower muscular
254 strength and higher body fat) (S. Cheuvront et al., 2002; Sparling & Cureton, 1983). However,
255 the comparison of the rate of improvement between men and women for all distances showed,
256 overall, a higher rate of improvement for women, particularly on the 5000-m. Beyond

257 physiological differences between women and men, historical backgrounds, societal influences
258 (sports participation and training) (Berthelot et al., 2015; S. N. Cheuvront et al., 2005), cultural
259 and sociopolitical factors (Berthelot et al., 2015; Capranica et al., 2013), behavioral
260 characteristics (pacing profile and tactics) (Emerson & Hill, 2014; Filipas et al., 2018; Hanley
261 & Williams, 2020) could play a greater role in potential for progression differences. We can
262 only speculate for these possible explanations to the sex progress differences, but it is
263 reasonable to assume that, for example, changes in the societal acceptance of women's events
264 running are a likely cause and have resulted in more opportunities for training and taking part
265 to competition to allow them to improve their record (Berthelot et al., 2015; Capranica et al.,
266 2013; Charlet, 2004; S. N. Cheuvront et al., 2005; Thibault et al., 2010). Although there was
267 no significant difference in this current study, the number of average performances per athletes
268 for the 5000-m event was lower for women than for men (2.3 and 4.1 performances,
269 respectively) and could be a possible explanation for this difference in performance
270 improvement rate between sex (Coast et al., 2004). Women may be more reluctant to run this
271 distance than men, but when they trained to run this distance in competition, they created more
272 of a gap with their last BRT. The current findings related to sex differences in rate of
273 progression are mainly in accordance with those reported by previous studies in different
274 athletic distances (Berthelot et al., 2015; S. N. Cheuvront et al., 2005; Coast et al., 2004; Haugen
275 et al., 2018; Hollings et al., 2014; Santos et al., 2020). For example, Haugen et al. (Haugen et al.
276 et al., 2018) observed greater improvement in women than men corresponding in middle- and
277 long-distance track running at $0.4, \pm 0.2\%$ between 2002 to 2016. However, these previous
278 studies (Berthelot et al., 2015; S. N. Cheuvront et al., 2005; Coast et al., 2004; Haugen et al.,
279 2018; Hollings et al., 2014; Santos et al., 2020) examined a large number of performances in
280 world-ranked top (*i.e.*, world-class elite athletes; *e.g.*, top 12-16 athletes for Hollings et al.
281 (Haugen et al., 2018) or top 10 or top 11-100 for Haugen et al. (Haugen et al., 2018)), while the
282 present analyses included season best results in all-time national-ranked top 10 of French
283 athletes, which could explain differences in values of performances progression. In addition, if
284 the data may be confounded by the lower number of women performances, particularly in the
285 longer distance as 5000-m, the differences of improvement rates could be explained also by the
286 years of performance studied. In this study, no time limit per athlete was set (except that
287 performances that were not achieved in the continuity of the seasons preceding the BRT were
288 not considered) in contrast to the studies of Haugen et al. (Haugen et al., 2018) who analyzed
289 the individual season bests for world-ranked athletes from 2002 to 2016 or Cheuvront et al. (S.
290 N. Cheuvront et al., 2005) who analyzed WR running performances in 1500-m and 5000-m

291 from the 1910s to 2004, for examples. In fact, it seems important to not exclude the
292 professionalization, the research, but also the technological advancements leading to an
293 increased effectiveness of some androgenic drugs on the physiology of women athletes in the
294 improvement rate of performances. In fact, the technological improvements in the last decades
295 (*e.g.*, running shoes, spike shoes) are probably an important factor that should be considered.
296 Indeed, the evolution of shoes (*i.e.*, running or spikes) promotes the absorption of shocks with
297 super-elastic foam to decrease the risk of the medical events (less traumatic) and promotes the
298 restitution of energy with carbon plates to optimize the performance (Bermon et al., 2021;
299 Hoogkamer et al., 2017). However, even though over the past decade, some studies showed
300 that women's performance tends to improve more than men's and catch up some of the gender
301 gap (Haugen et al., 2018), some previous work suggests that gender differences have plateaued
302 and that speculation that women are approaching men's times is not supported (Ostarello &
303 Suess, 2001; Sparling et al., 1998). Indeed, it has previously been suggested that differences in
304 running time performance between sexes reduced with increasing distances (Bam et al., 1997;
305 Whipp & Ward, 1992). Although, the sex physiological difference is at the advantage to men
306 in running performance (*e.g.*, higher $\dot{V}O_2\text{max}$...) (S. Cheuvront et al., 2002; S. N. Cheuvront et
307 al., 2005; Sparling & Cureton, 1983), some biological factors in favor of women may also come
308 into play like a better tolerance of hot and humid racing conditions (due to their smaller body
309 size) (Dennis & Noakes, 1999; Marino et al., 2000), and better substrate utilization (higher of
310 fat and lower carbohydrate consumptions for training athletes) during endurance exercise (Tate
311 & Holtz, 1998). However, more recent studies (S. N. Cheuvront et al., 2005; Coast et al., 2004),
312 in contrast to the suggestions, showed that the gap in performance increase with distance rather
313 than decrease at distances beyond the marathon. However, given the key physiological
314 parameters) (Billat & Koralsztein, 1996; Di Prampero et al., 1986; Hill et al., 1924; Kenney &
315 Hodgson, 1985) dependent on the sex (Trappe et al., 1996; Vickers & Vertosick, 2016) and
316 determining middle-distance or long-distance races (*e.g.*, 800-m, 1500-m or 5000-m), this
317 explanation seems relatively limited and more suited to long-distance or ultra-distance races
318 (*e.g.*, longer than marathon).

319 Last, the results of this current study showed significantly higher rate of improvement (*i.e.*, 1-
320 SIR and M-SIR in minutes) for the longest distance mainly (*i.e.*, 5000-m) than the shortest
321 distances (*i.e.*, 800-m and 1500-m). These results showed similar trends to the study by
322 Lerebourg et al. (Lerebourg & Coquart, 2020) which showed a significant increase in the level
323 of performance of women for the longer distances (*e.g.*, 5000-m). In middle- and long distance

324 track-running, the main contribution comes from the oxidative phosphorylation pathway (*i.e.*,
325 aerobic metabolism), but the glycolytic pathway (*i.e.*, anaerobic metabolism) is also fundamental
326 (Duffield et al., 2005a, 2005b; Kenney & Hodgson, 1985). However, longer distances such as
327 the 5000-m require a higher 'oxidative phosphorylation pathway to maintain the endurance
328 while shorter distances such as the 800-m and 1500-m are characterized by higher
329 preponderance of the "glycolytic pathway" to maintain the speed (Chamari & Padulo, 2015;
330 Duffield et al., 2005a, 2005b; Kenney & Hodgson, 1985; Spencer et al., 1996). In addition, the
331 training volumes and number of years of running experience would permit to progress over
332 long distances, notably through the management of the race pace or the choice of the race
333 strategy and would allow to reach peak performance at older ages in the longest distances.
334 Moreover, and as mentioned earlier, the technological improvements for running shoes in the
335 last decades (*i.e.*, carbon fiber plate and super-elastic foam) can then allow longer track distance
336 runners (*e.g.*, 5000-m) to preserve themselves more in training, linked to a higher training
337 volume and higher ground impact (more distance covered) than short middle-distance runners
338 (*e.g.*, 800-m, 1500-m), while optimizing their performance (Bermon et al., 2021; Hoogkamer
339 et al., 2017).

340 Thus, to assess the optimal performance improvement rate and increase the chances of a BRT
341 that could lead to success in national or major sporting events, understanding and seeking to
342 improve performance variables in middle- and long-distance track running seems essential for
343 coaches and athletes. This measure of performance improvement rate proposed to know the
344 annual performance gain of top-national athletes, could help the coaches or athletes in the
345 intervention of training to make modification or optimizing the progress (Santos et al., 2020).
346 Indeed, for athletes who fall within these limits, it is reasonable to assume that the training
347 strategy is effective (Santos et al., 2020). However, athletes do not always manage to improve
348 their performance every year, and for those who improve below this range, a change in training
349 strategy should be considered (Santos et al., 2020). Full details about performance improvement
350 limits are presented in the "Practical Applications" section.

351 The main limitation of this study was that the running performances were not executed in the
352 same time scale (*i.e.*, no time limit per athlete, actual best all-time top 10 runners) and not
353 experiencing similar technological (*e.g.*, running shoes, recovery mode...) or environmental
354 conditions (*e.g.*, temperature, barometric pressure, precipitations...), with identical physical
355 and psychological (*e.g.*, motivation and stress) conditions, and using the same opposition field

356 and/or race strategy. Another limitation of the study could be the limited number of
357 performances recorded for some athletes, for many reasons (*e.g.*, injuries, personal reasons,
358 change of athletic distance, doping...) to calculate progression rates (*i.e.*, 1-SIR or M-SIR).

359 **Practical Applications**

360 The present study seeks to assist coaches working with men and women athletes in middle- and
361 long-distance track running (*i.e.*, 800-m, 1500-m and 5000-m events). Therefore, in agreement
362 with Santos' study (Santos et al., 2020), graphs have been produced that can help to verify the
363 improvement rate of the athletes, so that coaches can be assisted by it and know if their training
364 methods are producing a rate of performance increase like National Record holders (Figure 3).

365 To apply the results from the index, an athlete should consider his/her best race time from
366 present season minus his/her race time from the last immediate season (or the average of the
367 last seasons). This number is his/her rate of improvement and could be compared to the actual
368 best all-time top ten national runners (*i.e.*, of all years) within his/her event (Figure 3 and Table
369 3).

370 For example, a man races the 5000-m-dash with her best race time in five past seasons being
371 13.38 min and 14.87 min (*i.e.*, 13 min 23 s and 14 min 52 s, respectively). In 2020, this man
372 performed 13.25 min (*i.e.*, 13 min 15 s) (Table 3).

373 Considering the Table 3 and plot from Figure 3 for men within the 5000-m and average rate of
374 0.32 min (*i.e.*, 19.2 s), this man falls within the very dark grey range. Meaning that the training
375 strategy seems to be very effective, and a projection considering her age and target race event
376 could be planned.

377 **Conclusion**

378 Women improved better annually than men before reaching the 5000-m national top 10 or their
379 personal record. Whatever the sex, the improvement rate was higher for the longer distance
380 than the shorter track distances. In addition, this study provide data that can help athletes and
381 coaches guide performance improvement in percentage and absolute values for those seeking
382 the top 10 National or national record of 800-m, 1500-m and 5000-m or can help whether some
383 adjustments are needed in the training process. Future research could focus on other distances
384 in athletics such as road races (*i.e.*, 10 km, half marathon, marathon).

385 **Acknowledgments**

386 The authors are grateful to the French Athletics Federation (Fédération Française d'Athlétisme)
387 for data diffusion, the Institut de Recherche Interdisciplinaire Homme et Société (IRIHS) and
388 the Orthodynamica Center at Mathilde Hospital 2 for fundings.

389 **Disclosure statement**

390 No potential conflict of interest was reported by the authors.

391 **References**

- 392 Bam, J., Noakes, T. D., Juritz, J., & Dennis, S. C. (1997). Could women outrun men in
393 ultramarathon races? *Medicine and Science in Sports and Exercise*, 29(2), 244-247.
394 <https://doi.org/10.1097/00005768-199702000-00013>
- 395 Bermon, S., Garrandes, F., Szabo, A., Berkovics, I., & Adami, P. E. (2021). Effect of
396 Advanced Shoe Technology on the Evolution of Road Race Times in Male and
397 Female Elite Runners. *Frontiers in Sports and Active Living*, 3, 653173.
398 <https://doi.org/10.3389/fspor.2021.653173>
- 399 Berthelot, G., Sedeaud, A., Marck, A., Antero-Jacquemin, J., Schipman, J., Sauliere, G.,
400 Marc, A., Desgorces, F.-D., & Toussaint, J.-F. (2015). Has athletic performance
401 reached its peak? *Sports Medicine*, 45(9), 1263-1271.
- 402 Billat, L. V., & Koralsztein, J. P. (1996). Significance of the velocity at VO₂max and time to
403 exhaustion at this velocity. *Sports Medicine (Auckland, N.Z.)*, 22(2), 90-108.
404 <https://doi.org/10.2165/00007256-199622020-00004>
- 405 Blythe, D. A. J., & Király, F. J. (2016). Prediction and quantification of individual athletic
406 performance of runners. *PLoS One*, 11(6), e0157257.
407 <https://doi.org/10.1371/journal.pone.0157257>
- 408 Capranica, L., Piacentini, M. F., Halson, S., Myburgh, K. H., Ogasawara, E., & Millard-
409 Stafford, M. (2013). The gender gap in sport performance : Equity influences equality.
410 *International Journal of Sports Physiology and Performance*, 8(1), 99-103.
411 <https://doi.org/10.1123/ijsp.8.1.99>
- 412 Chamari, K., & Padulo, J. (2015). « Aerobic » and « Anaerobic » terms used in exercise
413 physiology : A critical terminology reflection. *Sports Medicine - Open*, 1(1), 9.
414 <https://doi.org/10.1186/s40798-015-0012-1>
- 415 Charlet, S. (2004). *L'Athlétisme à travers les siècles*. Editions Alan Sutton.

- 416 Cheuvront, S., Moffatt, R., & DeRuisseau, K. (2002). *Body Composition and Gender*
417 *Differences in Performance* (p. 177-200). <https://doi.org/10.1201/9781420041323.ch8>
- 418 Cheuvront, S. N., Carter, R., Deruisseau, K. C., & Moffatt, R. J. (2005). Running
419 performance differences between men and women:an update. *Sports Medicine*
420 *(Auckland, N.Z.)*, 35(12), 1017-1024. [https://doi.org/10.2165/00007256-200535120-](https://doi.org/10.2165/00007256-200535120-00002)
421 00002
- 422 Coast, J. R., Blevins, J. S., & Wilson, B. A. (2004). Do gender differences in running
423 performance disappear with distance? *Canadian Journal of Applied Physiology =*
424 *Revue Canadienne De Physiologie Appliquee*, 29(2), 139-145.
425 <https://doi.org/10.1139/h04-010>
- 426 Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, 1(3),
427 98-101.
- 428 Coquart, J. B., Mercier, D., Tabben, M., & Bosquet, L. (2015). Influence of sex and specialty
429 on the prediction of middle-distance running performances using the Mercier et al.'s
430 nomogram. *Journal of Sports Sciences*, 33(11), 1124-1131.
431 <https://doi.org/10.1080/02640414.2014.986499>
- 432 Dennis, S. C., & Noakes, T. D. (1999). Advantages of a smaller bodymass in humans when
433 distance-running in warm, humid conditions. *European Journal of Applied Physiology*
434 *and Occupational Physiology*, 79(3), 280-284. <https://doi.org/10.1007/s004210050507>
- 435 Di Prampero, P. E., Atchou, G., Brückner, J. C., & Moia, C. (1986). The energetics of
436 endurance running. *European Journal of Applied Physiology and Occupational*
437 *Physiology*, 55(3), 259-266. <https://doi.org/10.1007/bf02343797>
- 438 Duffield, R., Dawson, B., & Goodman, C. (2005a). Energy system contribution to 400-metre
439 and 800-metre track running. *Journal of Sports Sciences*, 23(3), 299-307.
440 <https://doi.org/10.1080/02640410410001730043>

- 441 Duffield, R., Dawson, B., & Goodman, C. (2005b). Energy system contribution to 1500- and
442 3000-metre track running. *Journal of Sports Sciences*, 23(10), 993-1002.
443 <https://doi.org/10.1080/02640410400021963>
- 444 Emerson, J., & Hill, B. (2014). Gender Differences in Competition : Running Performance in
445 1,500 Meter Tournaments. *Eastern Economic Journal*, 40(4), 499-517.
- 446 Filipas, L., Nerli Ballati, E., Bonato, M., La Torre, A., & Piacentini, M. F. (2018). Elite Male
447 and Female 800-m Runners' Display of Different Pacing Strategies During Season-
448 Best Performances. *International Journal of Sports Physiology and Performance*, 1-5.
449 <https://doi.org/10.1123/ijsp.2018-0137>
- 450 Gembris, D., Taylor, J. G., & Suter, D. (2007). Evolution of athletic records : Statistical
451 effects versus real improvements. *Journal of Applied Statistics*, 34(5), 529-545.
- 452 Goss-Sampson, M. (2019). *Statistical analysis in JASP: A guide for students*. JASP.
- 453 Hanley, B., & Williams, E. L. (2020). Successful Pacing Profiles of Olympic Men and
454 Women 3,000 m Steeplechasers. *Frontiers in Sports and Active Living*, 2, 21.
- 455 Haugen, T. A., Solberg, P. A., Foster, C., Morán-Navarro, R., Breitschädel, F., & Hopkins,
456 W. G. (2018). Peak Age and Performance Progression in World-Class Track-and-
457 Field Athletes. *International Journal of Sports Physiology and Performance*, 13(9),
458 1122-1129. <https://doi.org/10.1123/ijsp.2017-0682>
- 459 Hill, A. V., Long, C. N. H., & Lupton, H. (1924). Muscular Exercise, Lactic Acid, and the
460 Supply and Utilisation of Oxygen. *Proceedings of the Royal Society of London. Series*
461 *B, Containing Papers of a Biological Character*, 96(679), 438-475. JSTOR.
- 462 Hollings, S. C., Hopkins, W. G., & Hume, P. A. (2014). Age at Peak Performance of
463 Successful Track & Field Athletes. *International Journal of Sports Science &*
464 *Coaching*, 9(4), 651-661. <https://doi.org/10.1260/1747-9541.9.4.651>

- 465 Hoogkamer, W., Kram, R., & Arellano, C. J. (2017). How Biomechanical Improvements in
466 Running Economy Could Break the 2-hour Marathon Barrier. *Sports Medicine*
467 (*Auckland, N.Z.*), 47(9), 1739-1750. <https://doi.org/10.1007/s40279-017-0708-0>
- 468 Kenney, W. L., & Hodgson, J. L. (1985). Variables predictive of performance in elite middle-
469 distance runners. *British Journal of Sports Medicine*, 19(4), 207.
470 <https://doi.org/10.1136/bjism.19.4.207>
- 471 Lerebourg, L., & Coquart, J. B. (2020). Changes in performances/characteristics of French
472 female runners over the last 12 years. *Research in Sports Medicine*, 1-11.
- 473 Marino, F. E., Mbambo, Z., Kortekaas, E., Wilson, G., Lambert, M. I., Noakes, T. D., &
474 Dennis, S. C. (2000). Advantages of smaller body mass during distance running in
475 warm, humid environments. *Pflugers Archiv: European Journal of Physiology*,
476 441(2-3), 359-367. <https://doi.org/10.1007/s004240000432>
- 477 Ostarello, J., & Suess, E. (2001). The gender gap : 100m to 42km. *ISBS-Conference*
478 *Proceedings Archive*.
- 479 Santos, P. A., Sousa, C. V., Barbosa, L. P., Aguiar, S. S., Sales, M. M., Simões, H. G.,
480 Nikolaidis, P. T., & Knechtle, B. (2020). Breaking the athletics world record in the
481 100 and 400 meters : An alternative method for assessment. *The Journal of Sports*
482 *Medicine and Physical Fitness*, 60(10), 1317-1321. [https://doi.org/10.23736/S0022-](https://doi.org/10.23736/S0022-4707.20.10753-9)
483 [4707.20.10753-9](https://doi.org/10.23736/S0022-4707.20.10753-9)
- 484 Schulz, R., & Curnow, C. (1988). Peak performance and age among superathletes : Track and
485 field, swimming, baseball, tennis, and golf. *Journal of gerontology*, 43(5), NaN-NaN.
486 <https://doi.org/10.1093/geronj/43.5.P113>
- 487 Sparling, P. B., & Cureton, K. J. (1983). Biological determinants of the sex difference in 12-
488 min run performance. *Medicine and Science in Sports and Exercise*, 15(3), 218-223.

- 489 Sparling, P. B., O'Donnell, E. M., & Snow, T. K. (1998). The gender difference in distance
490 running performance has plateaued : An analysis of world rankings from 1980 to
491 1996. *Medicine and Science in Sports and Exercise*, 30(12), 1725-1729.
492 <https://doi.org/10.1097/00005768-199812000-00011>
- 493 Spencer, M. R., Gastin, P. B., & Payne, W. (1996). Energy system contribution during 400 to
494 1500 metres running. *New Studies in Athletics*, 11(4), 59-66.
- 495 Tabben, M., Bosquet, L., & Coquart, J. B. (2016). Effect of performance level on the
496 prediction of middle-distance-running performances using a nomogram. *International
497 journal of sports physiology and performance*, 11(5), 623-626.
- 498 Tanaka, H., & Seals, D. R. (2008). Endurance exercise performance in Masters athletes : Age-
499 associated changes and underlying physiological mechanisms. *The Journal of
500 Physiology*, 586(1), 55-63. <https://doi.org/10.1113/jphysiol.2007.141879>
- 501 Tate, C. A., & Holtz, R. W. (1998). Gender and fat metabolism during exercise : A review.
502 *Canadian Journal of Applied Physiology = Revue Canadienne De Physiologie
503 Appliquee*, 23(6), 570-582. <https://doi.org/10.1139/h98-032>
- 504 Thibault, V., Guillaume, M., Berthelot, G., Helou, N. E., Schaal, K., Quinquis, L., Nassif, H.,
505 Tafflet, M., Escolano, S., Hermine, O., & Toussaint, J.-François. (2010). Women and
506 Men in Sport Performance : The Gender Gap has not Evolved since 1983. *Journal of
507 Sports Science & Medicine*, 9(2), 214-223.
- 508 Trappe, S. W., Costill, D. L., Vukovich, M. D., Jones, J., & Melham, T. (1996). Aging among
509 elite distance runners : A 22-yr longitudinal study. *Journal of Applied Physiology
510 (Bethesda, Md.: 1985)*, 80(1), 285-290. <https://doi.org/10.1152/jappl.1996.80.1.285>
- 511 Vickers, A. J., & Vertosick, E. A. (2016). An empirical study of race times in recreational
512 endurance runners. *BMC Sports Science, Medicine & Rehabilitation*, 8(1), 26.
513 <https://doi.org/10.1186/s13102-016-0052-y>

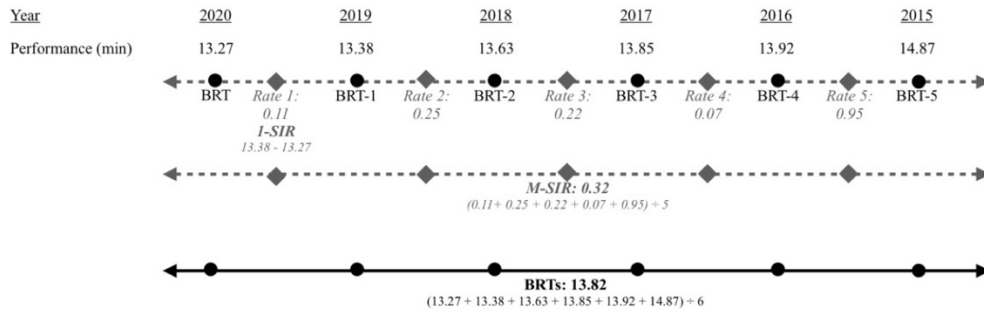
514 Whipp, B. J., & Ward, S. A. (1992). Will women soon outrun men? *Nature*, 355(6355), 25.

515 <https://doi.org/10.1038/355025a0>

516

517

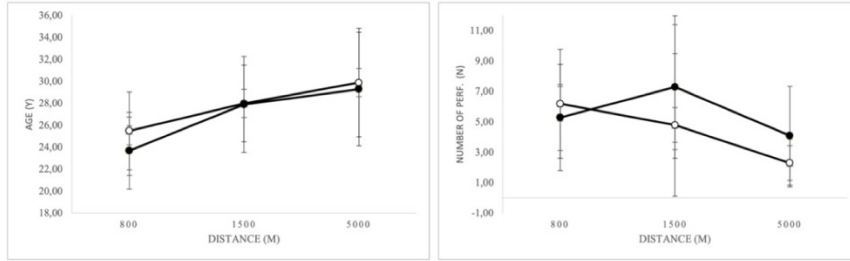
518 **Figure 1.** Definitions of the BRT, BRT-1, BRTs, 1-SIR and M-SIR with an example of man athlete in
 519 the 5000-m for 6 years.
 520



521
 522
 523
 524
 525
 526
 527

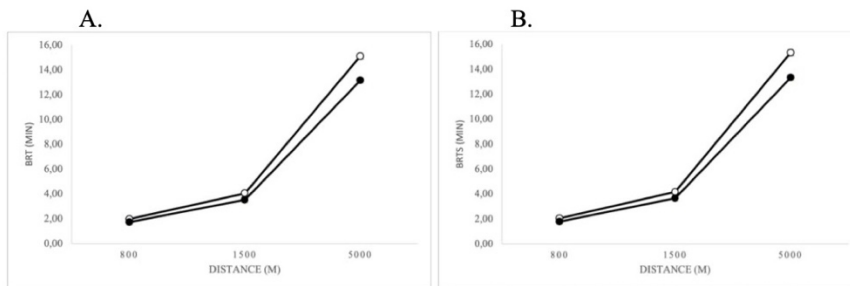
Note: BRT: best race time of the career (*i.e.*, personal best); BRT-1: BRT in the season immediate before the BRT;
 BRTs: average of BRT in the last seasons before the BRT; 1-SIR: difference between the BRT and the BRT-1;
 M-SIR: average of improvement rate in the season immediate before and seasons before.

528 **Figure 2.** Descriptive plots of each variable according to sex and distance of best all-time
 529 National top 10 athletes in middle- and long-distance track running (Mean and SD).



530

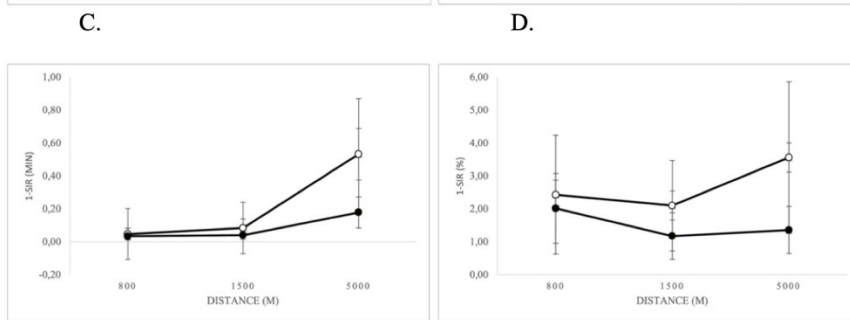
531



532

533

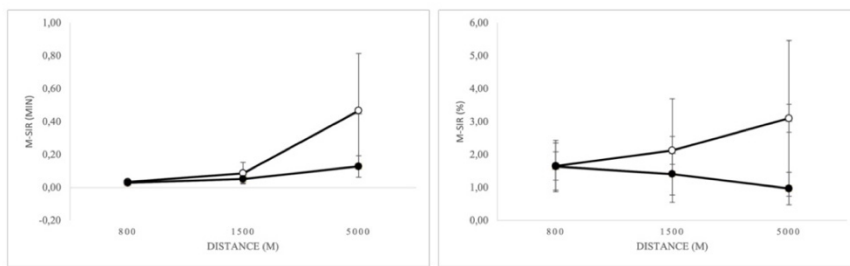
534



535

536

537



538

539

540

541

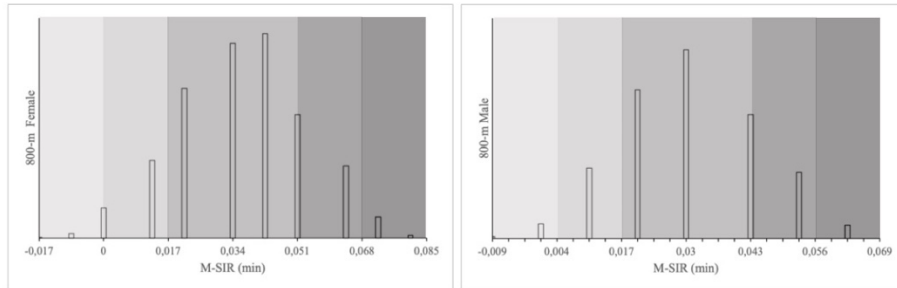
542

543

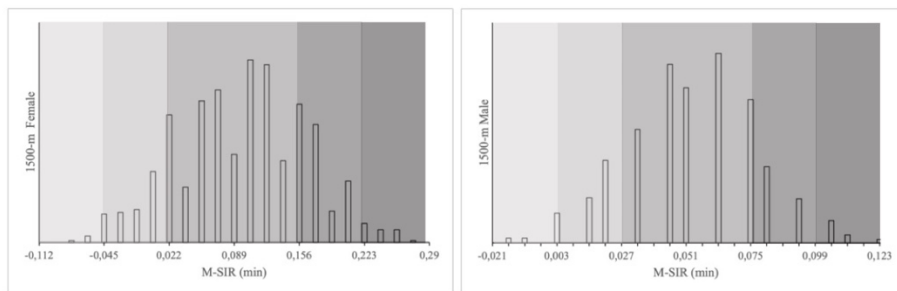
Legend: The white and black dots represent women and men respectively of each variable: age (A), number of performances (B), BRT (min) (C), BRTs (min) (D), 1-SIR (min) (E), 1-SIR (%) (F), M-SIR (min) (G), and M-SIR (%) (H).

544 **Figure 3.** Improvement rate index for the 800-m, 1500-m and 5000-m races for men and women
 545 (min).

546

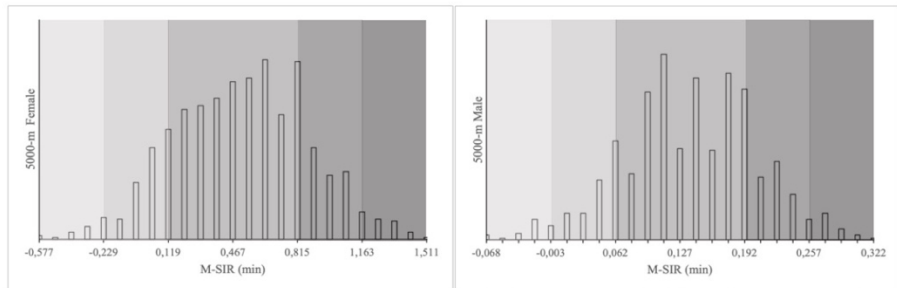


547



548

549



550

551

552

553

554

555

556

557

558

559

560

561

562

563

Legend: M-SIR: multi-season improvement rate average (min). Grey range (± 1 SD): improvement within the average of the top 10 runners; Light grey range ($- 2$ SD): improvement below the average of the top 10, a change in training strategy should be considered; Very light grey range ($- 3$ SD): improvement far below the average of the top 10, a major change in training strategy should be considered; Dark grey range ($+ 2$ SD): improvement above the top 10 average, training strategy is effective and possible causes should be investigated; Very dark grey range ($+ 3$ SD): improvement far above the top 10 average, training strategy is very effective and possible causes should be investigated; doping agencies should also consider additional attention.

564 **Table 1.** Extracted data of best all-time National top 10 records retained by sex and distance
 565 (Name, BRT, date and age of BRT).

566

Distance	Rank	Men			Women		
		BRT (min and s)	Date of BRT	Age of BRT(y)	BRT (min and s)	Date of BRT	Age of BRT (y)
800-m	1	1 min 42 s	18/07/2014	22	1 min 56 s	09/09/1995	24
	2	1 min 43 s	08/09/2002	24	1 min 58 s	05/06/2016	23
	3	1 min 43 s	12/09/1979	25	1 min 58 s	15/09/2020	24
	4	1 min 43 s	13/07/1987	27	1 min 58 s	09/09/2010	31
	5	1 min 44 s	27/08/2016	25	1 min 58 s	18/07/2008	23
	6	1 min 44 s	02/07/2004	23	1 min 59 s	27/06/1992	25
	7	1 min 44 s	14/08/2020	22	1 min 59 s	15/07/1986	20
	8	1 min 44 s	29/08/2001	21	1 min 59 s	31/07/2004	27
	9	1 min 45 s	16/08/1995	27	1 min 59 s	21/07/2004	31
	10	1 min 45 s	06/08/2004	21	1 min 59 s	19/07/2002	27
1500-m	1	3 min 28 s	05/09/2003	25	3 min 59 s	16/07/2010	31
	2	3 min 30 s	15/08/2003	25	4 min 00 s	11/07/2006	33
	3	3 min 31 s	19/07/2002	33	4 min 01 s	16/07/2010	30
	4	3 min 32 s	21/07/1999	26	4 min 01 s	09/09/2005	26
	5	3 min 32 s	12/07/1995	29	4 min 02 s	23/08/1996	25
	6	3 min 32 s	19/07/2013	35	4 min 04 s	04/07/2003	29
	7	3 min 32 s	02/07/2004	21	4 min 04 s	18/07/1994	29
	8	3 min 33 s	10/08/1996	25	4 min 05 s	04/08/1999	30
	9	3 min 33 s	24/08/2019	32	4 min 05 s	19/08/1987	21
	10	3 min 33 s	02/07/2013	28	4 min 07 s	06/07/2001	26
5000-m	1	12 min 58 s	28/07/2000	28	14 min 43 s	03/09/2004	30
	2	13 min 02 s	28/08/1998	28	14 min 47 s	01/08/2000	26
	3	13 min 10 s	06/07/2012	32	14 min 58 s	07/09/1999	30
	4	13 min 10 s	25/08/2000	26	15 min 07 s	23/05/2014	24
	5	13 min 11 s	20/06/2006	29	15 min 11 s	14/07/1998	24
	6	13 min 12 s	18/05/2014	36	15 min 12 s	14/07/2006	27
	7	13 min 12 s	07/09/1996	21	15 min 13 s	20/06/1999	34
	8	13 min 13 s	17/06/1995	35	15 min 14 s	06/07/2013	32
	9	13 min 13 s	01/08/1998	35	15 min 16 s	02/08/2014	40
	10	13 min 13 s	05/07/2019	23	15 min 16 s	21/08/1994	32

567 *Legend:* BRT: best race time

568

569 **Table 2.** Descriptive data of best all-time National top 10 records in middle- and long-
 570 distance track running in athletics (Mean values, SD and range).
 571

	Men		Women	
	Mean (SD)	Range (min-max)	Mean (SD)	Range (min-max)
Age (y)				
800-m	23.7 (2.3)	21 - 27	25.5 (3.5)	20 - 31
1500-m	27.9 (4.4)	21 - 35	28.0 (3.5)	21 - 33
5000-m	29.3 (5.2)	21 - 36	29.9 (4.9)	24 - 40
Number of Performances registered per athlete (n)				
800-m	5.3 (2.2)	2 - 8	6.2 (3.6) ^c	1 - 12
1500-m	7.3 (4.1)	1 - 13	4.8 (4.7)	1 - 16
5000-m	4.1 (3.3)	1 - 11	2.3 (1.6) ^a	1 - 6
BRT (min and s)				
800-m	1 min 44 s (1.2 s) ^{*b,c}	1 min 42 s - 1 min 45 s	1 min 58 s (1.2 s) ^{*b,c}	1 min 56 s - 1 min 59 s
1500-m	3 min 32 s (1.2 s) ^{*a,c}	3 min 28 s - 3 min 33 s	4 min 3 s (2.6 s) ^{*a,c}	3 min 59 s - 4 min 7 s
5000-m	13 min 9 s (5.4 s) ^{*a,b}	12 min 58 s - 13 min 13 s	15 min 6 s (12 s) ^{*a,b}	14 min 43 s - 15 min 16 s
BRTs (min and s)				
800-m	1 min 47 s (1.8 s) ^{*b,c}	1 min 44 - 1 min 49 s	2 min 3 s (2.4 s) ^{*b,c}	1 min 58 - 2 min 5 s
1500-m	3 min 39 s (5.4 s) ^{*a,c}	3 min 33 s - 3 min 46 s	4 min 10 s (6.6 s) ^{*a,c}	4 min 1 s - 4 min 21 s
5000-m	13 min 19 s (12 s) ^{*a,b}	13 min 18 s - 13 min 40 s	15 min 20 s (15 s) ^{*a,b}	14 min 55 s - 15 min 39 s
1-SIR (min and s)				
800-m	1.8 s (1.2 s) ^c	1.2 s - 4.2 s	3 s (2.4 s) ^c	0 s - 7.2 s
1500-m	2.4 s (1.8 s) ^c	0 s - 4.8 s	5.4 s (3 s) ^c	1.8 s - 12 s
5000-m	10.8 s (5.4 s) ^{a,b}	4.2 s - 21 s	31.8 s (20.4 s) ^{a,b}	13.8 s - 1 min 6 s
1-SIR (%)				
800-m	2.02 (1.06)	0.95 - 3.88	2.43 (1.80)	0 - 5.88
1500-m	1.18 (0.71)	0 - 2.35	2.11 (1.37)	0.82 - 4.98
5000-m	1.36 (0.72)	0.50 - 2.65	3.57 (2.30)	1.58 - 7.47
M-SIR (min and s)				
800-m	1.8 s (0.6 s) ^c	0.6 s - 3 s	1.8 s (1.2 s) ^{b,c}	1.2 s - 4.2 s
1500-m	3 s (1.2 s) ^c	1.2 s - 6 s	5.4 s (4.2 s) ^{a,c}	1.8 s - 13.2 s
5000-m	7.8 s (4.2 s) ^{a,b}	2.4 s - 13.8 s	28.2 s (21 s) ^{a,b}	9.6 s - 1 min 6 s
M-SIR (%)				
800-m	1.64 (0.72)	0.57 - 2.72	1.66 (0.78)	0.57 - 3.26
1500-m	1.41 (0.64)	0.62 - 2.68	2.13 (1.57)	0.81 - 5.27
5000-m	0.97 (0.49) [*]	0.32 - 1.77	3.10 (2.37) [*]	1.04 - 7.47

572 *Legend:* 1-SIR: 1-season improvement rate; M-SIR: multi-season improvement rate average; BRT: best race time
 573 (*i.e.*, PR); BRTs: best race times (*i.e.*, BRT in the last seasons before the PR); *Differences between sexes ($p <$
 574 0.05); ^aSignificant differences with 800-m; ^bSignificant differences with 1500-m; ^cSignificant differences with
 575 5000-m ($p <$ 0.05).

576
 577

578 **Table 3.** Example of a man runner in the 5000-m race for 6 years.
579

580

Year	Performance (min)	Rate of improvement (min)
2015	14.87	-
2016	13.92	0.95
2017	13.85	0.07
2018	13.63	0.22
2019	13.38	0.25
2020 (BRT)	13.27	0.11 (1 SIR)
Average rate of improvement		0.32 (M SIR)

Annexe 8. Projet : Conception et développement d'une gamme de Sandales Orthopédiques « sur-mesure » et personnalisables

L'entreprise Orthodynamica (OD) a pu faire le constat que certains de ses patients exprimaient des difficultés pour supporter leur traitement orthopédique en période estivale. En effet, les semelles orthopédiques (orthèses plantaires) ont pour inconvénient de ne pas pouvoir s'adapter aux chaussures ouvertes (sandales, tongs...). Face à une forte demande de la patientèle, OD a alors fait le choix d'opter pour la solution de sandales orthopédiques sur-mesure et personnalisables afin de régler le problème rencontré avec des orthèses plantaires lors de la saison estivale en y intégrant une semelle orthopédique directement dans une paire de chaussures ouvertes.

Le projet « Sandales » consistait alors, dans un premier temps, en la conception et le développement d'orthèses plantaires « sur-mesure » telles que des sandales orthopédiques personnalisables pour les patients ne supportant pas leur traitement orthopédique en période estivale (orthèses plantaires à placer dans leurs chaussures fermées) ou encore désireux de porter des chaussures ouvertes intégrant leur traitement podologique à leur domicile. Dans un second temps, le projet « Sandales » avait pour enjeu de concevoir des sandales orthopédiques sur-mesure de récupération destinées aux sportifs.

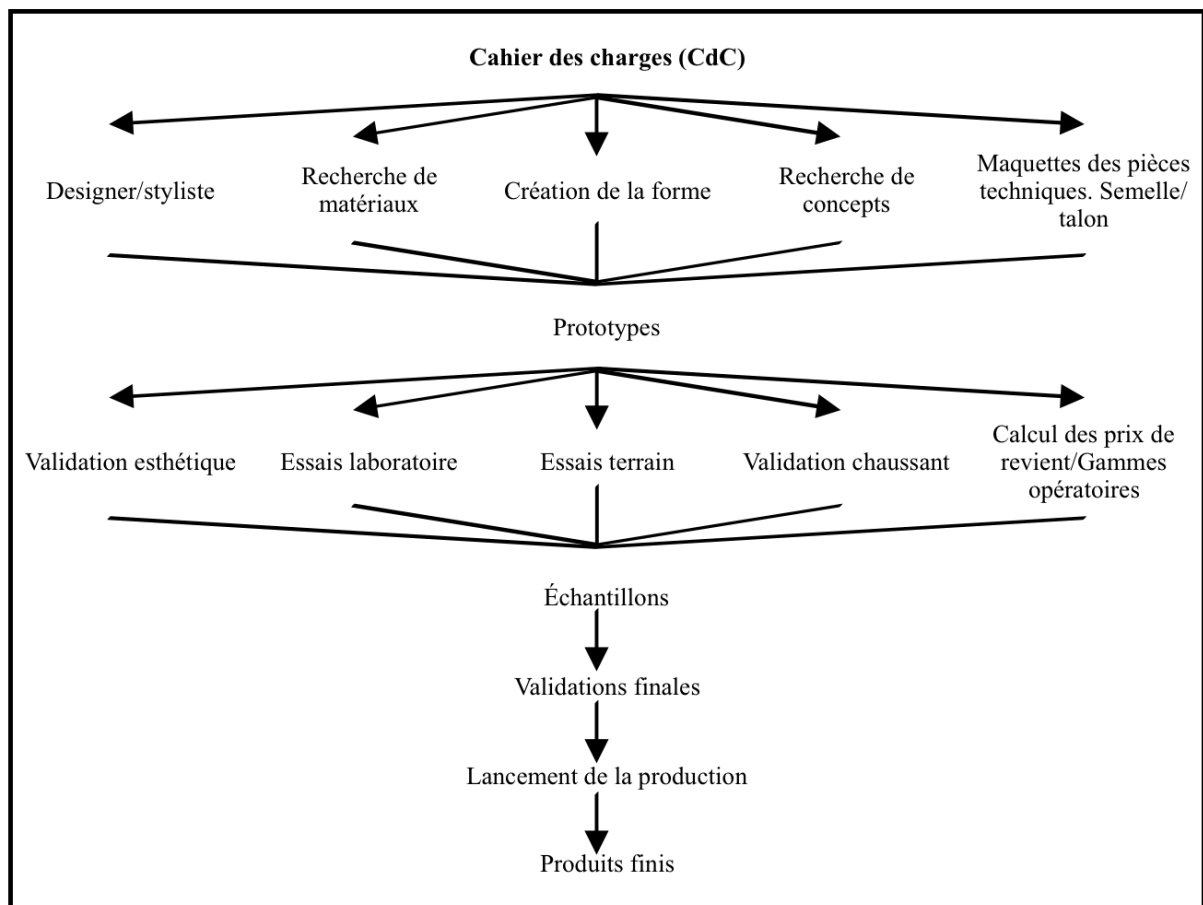
Cahier des charges (CdC) & Planning prévisionnel

Le cahier des charges reposait sur la conception et le développement d'une gamme de produits innovants, des sandales « sur-mesure », permettant de porter son traitement orthopédique lors des périodes estivales, des périodes chaudes ou encore à son domicile.

Au-delà du cahier des charges il était important de souligner que plusieurs opérations étaient nécessaires à la fabrication d'une chaussure, depuis la réflexion du styliste ou du designer, jusqu'à la mise en vente des produits finis.

Ce processus de fabrication nécessitait une multitude de savoir-faire, ainsi qu'une collaboration du fabricant avec de nombreux fournisseurs.

Le logigramme *ci-dessous* en donne l'illustration :



Processus de fabrication

Définition du cahier des charges de la gamme « Sandales sur-mesure »

Les charges ont été regroupées sous cinq grandes étapes :

Cahier des Charges Initial	
<p>PHASE I</p> <p>Stratégie, Définition & Identification du Concept</p> <p>Définition du Besoin</p>	<p>Analyse de l'Existant : Dresser un État des lieux et identifier les différents produits « leaders » ou modèles présents sur le marché.</p>
	<p>Étude du Besoin des utilisateurs (patients)</p>
	<p>Recherche de matériaux/ fournitures : Identifier les matériaux les plus adaptés aux produits (rapport qualité/prix). Rechercher des formes, des concepts, des modèles, des designs.</p>
	<p>Mesures Anthropométriques : Analyse des particularités anatomiques des pieds des utilisateurs dans l'objectif de proposer des produits/ taillants adaptés à leurs morphologies.</p>
<p>PHASE II</p> <p>Développement Technique & Production</p>	<p>Mise en place d'un cahier des charges mentionnant les spécificités techniques de chacun des produits (Diagramme FAST > Fonctions principales & Services > Solutions techniques)</p>
	<p>Définition Design Technique : Confection Technique des produits selon les besoins des utilisateurs ainsi que les contraintes biomécaniques & physiologiques liées à la marche et au port de sandales. Création de maquettes. Calcul des prix de revient/Gamme opératoire.</p>
	<p>Mise en place d'une étude de marché : Design Graphique (Design esthétique). Étudier les tendances de la mode.</p>
	<p>Prototypes & Ajustements : Production des premiers prototypes avec HL Industries. Réglages des semelles standards OD (semelles orthopédiques- orthèses plantaires).</p>
<p>PHASE III</p> <p>Tests, Retours & Validation Produits</p>	<p>Essais : laboratoire & terrain</p>
	<p>Retours Produits & Ajustements : Analyse & Discussion des résultats obtenus en interne. Validation ou Modification des Produits dans le but de proposer des produits adaptés aux utilisateurs.</p>
	<p>Validation des produits (chaussant, matériaux, esthétique) : Analyse & Discussion des résultats obtenus suite aux essais (terrain) et à l'étude de marché (esthétique, forme).</p>
	<p>Échantillons & Validation finale</p>
<p>PHASE IV</p> <p>Production</p>	<p>Lancement de la production : Commande des matériaux & Fabrication des premiers modèles (produits finis)</p>
<p>PHASE V</p> <p>Catalogue & Prix Produits</p>	<p>Création d'un catalogue « Présentation des modèles »</p>
	<p>Prix produits : Déterminer les gammes de prix en fonction des options.</p>

Cahier des charges (Cdc)

Le respect de l'ensemble de ces charges, listées précédemment, était essentiel pour mener à bien cette création de gamme de produits de « sandales sur-mesure ». Le cahier des charges a été tenu tout au long du projet et aucune modification n'a été apportée.

Description des cinq périodes du cahier des charges :

La première période comprenait l'analyse de l'existant (état de l'art) afin de dresser un état des lieux et identifier les différents produits « leaders » ou modèles présents sur le marché ; des recherches bibliographiques, physiologiques, anatomiques en lien avec la podologie et l'orthopédie sur les anomalies plantaires (hallux valgus, talalgie) ou une morphologie plantaire particulière (pieds creux, plat, pronateur) par exemple, dans le but de comprendre et d'identifier le besoin des utilisateurs (patients); des recherches de fournitures (cuir, formes de pied, bloc orthopédique, accessoires ou systèmes de serrage pour chaussures...) afin d'identifier les matériaux les plus adaptés aux produits (rapport qualité/prix). Cette première phase de « définition de concept » a permis également de prendre contact avec nos différents fournisseurs, notamment lors notre déplacement avec l'équipe HL Industries en Septembre 2018 à Milan (Italie) où se tenait le salon « Lineapelle ». Cette première phase était essentielle pour pouvoir enclencher la deuxième, soit la phase de « développement ».



Salon Lineapelle (Milano, Italie)

L'objectif, ensuite, était de travailler sur le développement « technique » des produits afin de lancer la production. Cette phase se résumait à la mise en place d'un cahier des charges (cf. Cdc) mentionnant les spécificités techniques de chacun des produits ; la définition du design technique (attribuer les fonctions à chaque produit) soit la confection technique des produits selon le besoin des utilisateurs (patients) ainsi que les contraintes biomécaniques et physiologiques liées à la marche et au port de sandales (réflexion accès sur l'interface « Homme-Matériel ») ; mise en place d'une étude de marché afin de définir les tendances actuelles de la mode et le design graphique soit l'esthétique des produits ; ainsi qu'à la production des premiers prototypes (essais des cuirs, des systèmes de serrage, du processus de fabrication...).



Étude de marché

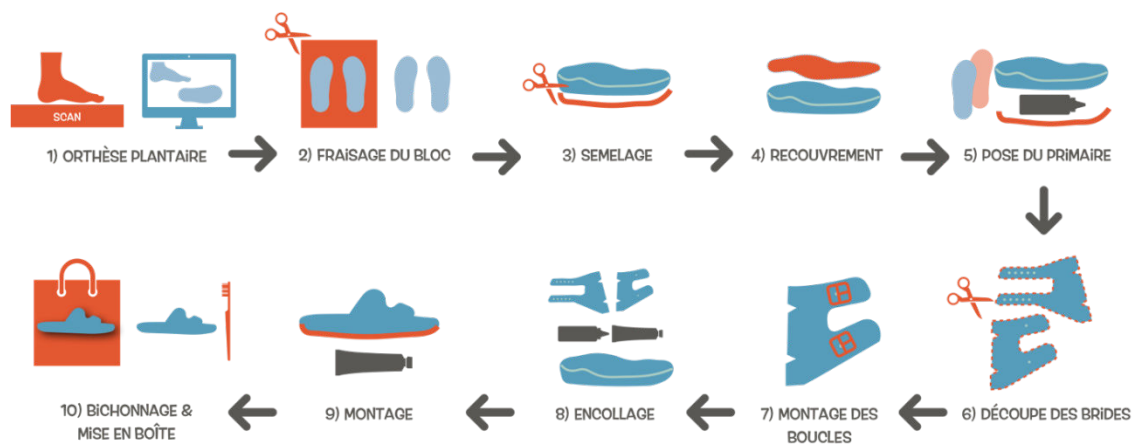
La troisième étape « test et validation », ayant pour objectif d'évaluer la qualité, le confort et la durabilité des sandales fabriquées, consistait en des essais en laboratoire (CTC de Lyon, réputé pour leur rigueur dans le processus de fabrication de chaussures) ainsi que des essais terrain (soit par ex., le port des sandales dans la vie de tous les jours) des prototypes ; des retours produits à l'aide de questionnaires afin de cerner les points forts et les limites des produits ; la validation des différents

produits compostant les sandales (forme du chaussant, matériaux des sandales, coloris des cuirs, des blocs ou finition des accessoires...); et la validation finale suite aux tests réalisés.

Description des analyses
Capacité d'absorption et de désorption d'eau des semelles premières
Frottement VESLIC à sec
Frottement VESLIC en humide
Résistance à la flexion des semelles - 100 000 cycles sans entaille
Résistance au déchirement des tiges, doublures et premières
Résistance au glissement sur sol carrelé avec solution de laurylsulfate de sodium
Résistance de la liaison
Vieillessement simulateur de marche

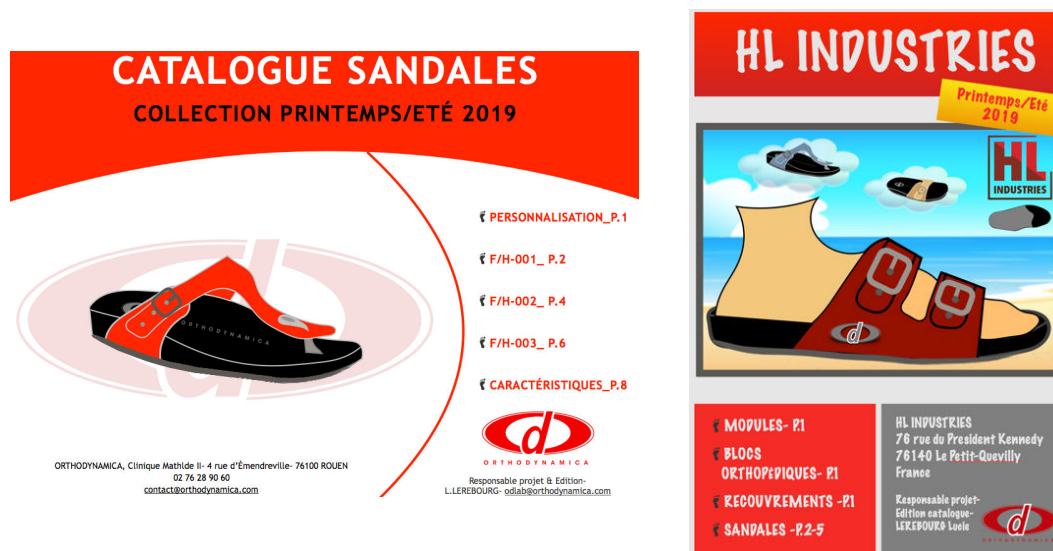
Exemples des essais en laboratoire réalisés au CTC

La quatrième étape se résumait au lancement de la production chez HL Industries. J'ai eu l'opportunité de réaliser un court-métrage pour les équipes HL Industries et Orthodynamica récapitulant l'ensemble des dix étapes du processus de fabrication (**cf. Processus de fabrication** : scan du pied du patient ; fraisage du bloc orthopédique ; le semelage ; le recouvrement ; la pose du primaire ; la découpe des brides ; le montage des boucles ; l'encollage ; le montage ; et le bichonnage et la mise en boîte) ainsi que les matériaux et outils nécessaires au montage d'une paire de sandales.



Processus de fabrication

La cinquième et dernière phase du projet « catalogue et prix » avait pour objectif de définir les prix de vente des sandales de HL Industries (fabricant et fournisseur) pour OD comprenant notamment les coûts des matériaux, des produits, de la main d'œuvre pour la fabrication et de définir également les prix de vente de l'entreprise OD pour les patients (TTC). Après avoir fixé les prix, l'enjeu était de réaliser deux catalogues, l'un destiné à l'entreprise et plus particulièrement aux praticiens d'OD mentionnant l'ensemble des prix d'achat auprès de HL Industries et l'un destiné à la clientèle mentionnant les prix de vente des sandales ainsi que les possibilités de choix concernant les modèles, les coloris et la finition, soit la partie personnalisable des sandales.



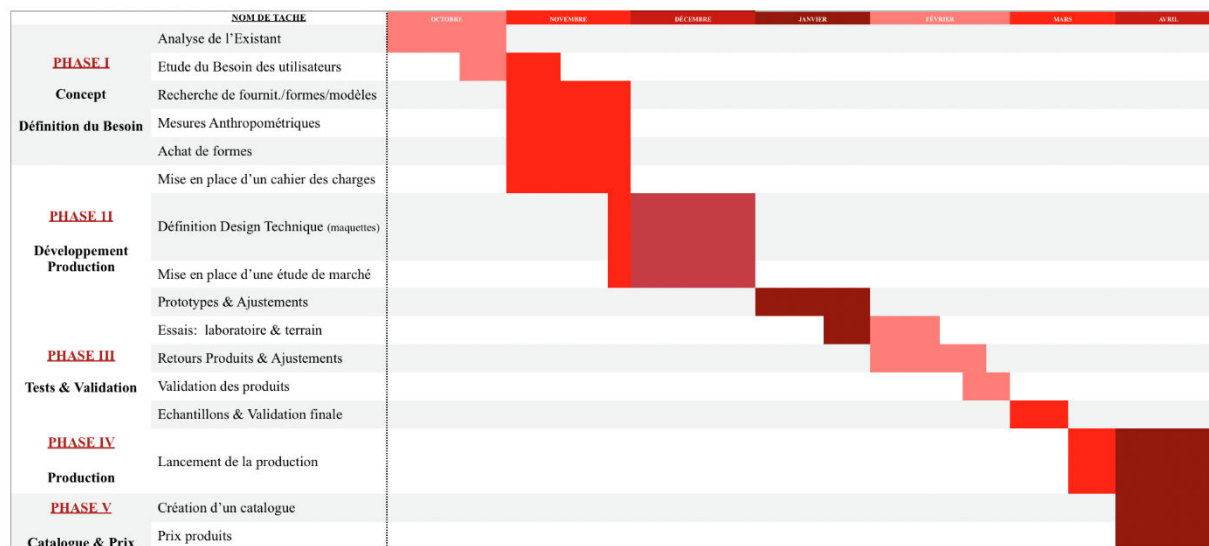
Catalogues Printemps-Été 2019

Planning prévisionnel de la gamme « Sandales sur-mesure »

Après rédaction du cahier des charges, l'objectif était de planifier l'ensemble des tâches nécessaires à la réalisation du projet.

Pour mener à bien la création de la gamme des « Sandales sur-mesure », avait été établi un planning prévisionnel afin de faciliter l'organisation et suivre l'avancement du projet. L'utilisation de cet outil a permis de définir des délais (en relation avec les

différentes unités de l'entreprise) et d'identifier les marges existantes sur certaines tâches. L'ensemble des tâches était présenté dans la colonne de gauche (**cf. *Planning prévisionnel***). Pour chaque tâche, la durée était matérialisée par une barre horizontale commençant à la date de début et finissant à la date de fin de la tâche. La stratégie adoptée était d'établir des délais assez larges pour les tâches les plus importantes (nécessitant plusieurs services de l'entreprise) afin de s'accorder de potentielles marges de manœuvre si les délais tenaient à être dépassés. Il était également possible de chevaucher plusieurs tâches sur l'échelle du temps.



Planning prévisionnel

Toutefois, ce planning prévisionnel a connu quelques modifications (**cf. *planning prévisionnel modifié***). En effet, nous avons abordé un manque de moyen financier à l'instant T avec l'entreprise OD, notamment lors de la phase III du projet, soit la partie « Tests et validation » (au CTC) qui avait pour objectif d'évaluer la qualité, le confort et la durabilité des sandales fabriquées. Ceci a donc considérablement altéré la planification du projet. La mise en route de la production prévue initialement au mois d'Avril 2019 a été repoussée au mois de Juin 2019. Par conséquent la saison 2019 a fortement été ralentie durant la période Février-Mai en

raison de ces tests en laboratoire qui ont retardé et repoussé l'enclenchement des phases suivantes (phases IV et V).

Cette mission en tant que chef de projet chez OD a réellement débuté en Septembre-Octobre 2018, soit quelques mois après le début de ma convention CIFRE (Juin 2018) et s'est concrétisée en Octobre 2019 où nous avons réalisé le premier bilan annuel de ce projet de création de gamme de produits de « sandales sur-mesure » avec les différents membres concernés. À l'heure actuelle, le projet est toujours d'actualité et poursuit son évolution. En effet, je veille toujours au suivi et au pilotage de ce projet dans le but d'apporter des modifications, des améliorations ou encore d'autres connaissances en sachant que d'autres projets comme la conception d'une gamme de « Sandales sur-mesure de récupération » pour le sportif, en lien avec mon travail de thèse ont émergé de ce dernier.

NOM DE TACHE		OCTOBRE	NOVEMBRE	DÉCEMBRE	JANVIER	FÉVRIER	MARS	AVRIL
PHASE I	Analyse de l'Existant							
	Etude du Besoin des utilisateurs							
	Concept							
	Définition du Besoin							
	Achat de formes							
PHASE II	Mise en place d'un cahier des charges							
	Définition Design Technique (maquettes)							
	Développement Production							
	Prototypes & Ajustements							
	Communication (moyens) & Stratégies							
PHASE III	Essais: laboratoire & terrain	EN ATTENTE ...						
	Retours Produits & Ajustements	RETARDÉ						
	Tests & Validation	RETARDÉ						
	Validation des produits	RETARDÉ						
PHASE IV	Echantillons & Validation finale							
	Production							
	Lancement de la production							
PHASE V	Catégorie & Prix & Communication							
	Création d'un catalogue							
	Prix produits							
PHASE VI	Lancement de la communication							

Planning prévisionnel modifié

Ainsi, nous pouvons observer quelques modifications entre le planning prévisionnel prévu initialement et le planning prévisionnel final avec notamment des ajustements

dans la durée et le contenu de certaines tâches (essais tests en laboratoire, par exemple). Toutefois, ces changements n'ont pas altéré la réalisation du projet « sandales sur-mesure » et montrent aujourd'hui, qu'il est important de faire preuve d'adaptation afin de mener à bien la réussite d'un projet. Les démarches réalisées et les procédés utilisés qui ont permis de mener à bien la création et le développement de la gamme « Sandales sur-mesure » constituent inévitablement la base de développement de la gamme, « Sandales sur-mesure de récupération », destinée aux sportifs.



UNIVERSITÉ DE ROUEN
UFR STAPS
LABORATOIRE CETAPS EA 3832
PROJET DE THÈSE CIFRE

Impact de sandales de récupération « sur-mesure », stimulant la semelle veineuse plantaire, sur la récupération, la fatigue et la performance chez le coureur à pied.

Présenté par :
Lucie LEREBOURG
 Sous la direction du Dr. Jérémie COQUART (MCF-HDR)
 &
 Sous le co-encadrement du Dr. Maxime L'HERMETTE (MCF)

Projet Scientifique (R&D)



Projet de Thèse CIFRE avec ORTHODYNAMICA et l'Université de Normandie, spécialité STAPS

Impact des sandales de récupération « sur-mesure », stimulant la semelle veineuse plantaire, sur la fatigue, la récupération et la performance chez le coureur à pied.

Étude n°1:
Revue systématique

Présenté par :
Lucie LEREBOURG
 Sous la direction du Dr. Jérémie COQUART (PU)
 &
 Sous le co-encadrement de Maxime L'HERMETTE (MCF)





Hypothèses générales : Nous pensons qu'au fur et à mesure de l'avancée dans la compétition simulée, le golfeur va accumuler du sang dans ses membres inférieurs, générant ainsi une fatigue et une chute de sa performance. De plus, nous supposons que le port de semelles de récupération peut stimuler davantage la semelle veineuse plantaire et améliorer le débit sanguin en diminuant l'accumulation du sang au niveau des membres inférieurs, réduisant ainsi la fatigue perçue. Nous pensons également que le port de semelles de récupération peut influencer sur la vitesse, la puissance, la stabilité du swing mais aussi la précision et par conséquent la performance golfique du joueur. Enfin, nous supposons aussi que la semelle peut apporter plus de confort et de bien-être au golfeur.

Pour permettre un retour veineux sanguin efficace, différents mécanismes de pompes sont mis en jeu.

1



Écrasement de la semelle veineuse plantaire
(Scurr et al., 1993)

2



Contraction musculaire du mollet
(Gardner et Fox, 1993 ; White et al., 1996 ; Uhl et al., 2009 ; Dagrada, 2009)

RÉSUMÉ

Cette thèse avait pour objectif de présenter les différents travaux réalisés sur la prédiction de la performance en course à pied afin d'aider les athlètes et les entraîneurs à optimiser leur processus d'entraînement. Ces études, en collaboration avec la Fédération Française d'Athlétisme (FFA), se sont appuyées sur le système d'information fédéral répertoriant notamment l'ensemble des résultats athlétiques, les bilans ou encore le nombre de licenciés. La première étude avait pour objectif d'exposer l'évolution des performances françaises des courses de demi-fond et de fond chez les femmes. Les études suivantes étaient principalement destinées à tester la validité, la justesse, et la précision de différentes méthodes de prédiction (*i.e.*, capacité à prédire les performances) sur des performances individuelles réelles d'athlètes de différents niveaux, hommes et/ou femmes. Les résultats se sont avérés valides et précis, quelle que soit la méthode de prédiction utilisée. Enfin, la dernière recherche était destinée à la prédiction du potentiel de performance. Cette étude a notamment mis en avant une analyse du taux d'amélioration des performances en demi-fond et en fond précédant la réalisation de records personnels chez les hommes et chez les femmes. Un index de progression à visée pratique, a également été proposé, afin d'évaluer l'évolution des performances et permettre une éventuelle détection et orientation des athlètes au fort potentiel.

Mots clés : Prédiction, estimation, validité, justesse, précision, record, course à pied, potentiel de performance

ABSTRACT

This thesis aimed to present the various studies conducted on the prediction of running performance to help athletes and coaches to optimize their training strategy. These studies, in collaboration with the *Fédération Française d'Athlétisme* (FFA), were based on the federal information system listing all athletic results, rankings and the number of members. The first study aimed to show the evolution of French performances in middle and long-distance races for women. The following studies were mainly designed to test the validity, accuracy, and precision of different prediction methods (*i.e.*, ability to predict performance) on actual individual performances of male and/or female athletes of different levels. The results were found to be valid and accurate regardless of the prediction method used. Finally, the last research was aimed at the prediction of performance potential. This study highlighted an analysis of the rate of improvement in middle and long-distance performance prior to personal bests in men and women. A practical progression index was also proposed to evaluate the evolution of performance and to allow for possible detection and orientation of high potential athletes.

Keywords: prediction, estimation, validity, accuracy, precision, record, running, performance potential