

UNIVERSITE TOULOUSE III - PAUL SABATIER

FACULTE DE SANTE DES SCIENCES PHARMACEUTIQUES

Année : 2024

2024/TOU3/2010

THESE

POUR LE DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE

Présentée et soutenue publiquement par

Camille JOLY

Le 5 Mars 2024

Le machine learning, outil d'intérêt dans la promotion du vieillissement en bonne santé : de la compréhension des mécanismes aux développements thérapeutiques

Directrice de thèse : Pr. Victorine Douin

Co-directrice de thèse : Dr. Isabelle Ader

JURY

Président : Pr. Daniel Cussac
1er assesseur : Dr. Laure Rouch
2ème assesseur : Dr. Marc Dubourdeau

PERSONNEL ENSEIGNANT

du Département des Sciences Pharmaceutiques

de la Faculté de santé

au 08 mars 2023

Professeurs Emérites

Mme BARRE A.	Biologie Cellulaire
M. BENOIST H.	Immunologie
Mme NEPVEU F.	Chimie analytique
Mme ROQUES C.	Bactériologie - Virologie
M. ROUGE P.	Biologie Cellulaire
M. SALLES B.	Toxicologie

Professeurs des Universités

Hospitalo-Universitaires

Mme AYYOUB M.	Immunologie
M. CESTAC P.	Pharmacie Clinique
M. CHATELUT E.	Pharmacologie
Mme DE MAS MANSAT V.	Hématologie
M. FAVRE G.	Biochimie
Mme GANDIA P.	Pharmacologie
M. PARINI A.	Physiologie
M. PASQUIER C.	Bactériologie - Virologie
Mme ROUSSIN A.	Pharmacologie
Mme SALLERIN B. (Directrice-adjointe)	Pharmacie Clinique
M. VALENTIN A.	Parasitologie

Universitaires

Mme BERNARDES-GENISSON V.	Chimie thérapeutique
Mme BOUTET E.	Toxicologie - Sémiologie
Mme COSTE A.	Parasitologie
Mme COUDERC B.	Biochimie
M. CUSSAC D. (Doyen-directeur)	Physiologie
Mme DERA EVE C.	Chimie Thérapeutique
M. FABRE N.	Pharmacognosie
Mme GIROD-FULLANA S.	Pharmacie Galénique
M. GUIARD B.	Pharmacologie
M. LETISSE F.	Chimie pharmaceutique
Mme MULLER-STAU MONT C.	Toxicologie - Sémiologie
Mme REYBIER-VUATTOUX K.	Chimie analytique
M. SEGUI B.	Biologie Cellulaire
Mme SIXOU S.	Biochimie
M. SOUCHARD J-P.	Chimie analytique
Mme TABOULET F.	Droit Pharmaceutique
Mme WHITE-KONING M.	Mathématiques

Maîtres de Conférences des Universités

Hospitalo-Universitaires

M. DELCOURT N.	Biochimie
Mme JUILLARD-CONDAT B.	Droit Pharmaceutique
Mme KELLER L.	Biochimie
M. PUISSET F.	Pharmacie Clinique
Mme ROUCH L.	Pharmacie Clinique
Mme ROUZAUD-LABORDE C	Pharmacie Clinique
Mme SALABERT A.S.	Biophysique
Mme SERONIE-VIVIEN S (*)	Biochimie
Mme THOMAS F. (*)	Pharmacologie

Universitaires

Mme ARELLANO C. (*)	Chimie Thérapeutique
Mme AUTHIER H.	Parasitologie
M. BERGE M. (*)	Bactériologie - Virologie
Mme BON C. (*)	Biophysique
M. BOUJILA J. (*)	Chimie Analytique
M. BROUILLET F.	Pharmacie Galénique
Mme CABOU C.	Physiologie
Mme CAZALBOU S. (*)	Pharmacie Galénique
Mme CHAPUY-REGAUD S. (*)	Bactériologie - Virologie
Mme COLACIOS C. (*)	Immunologie
Mme ECHINARD-DOUIN V. (*)	Physiologie
Mme EL GARAH F.	Chimie Pharmaceutique
Mme EL HAGE S.	Chimie Pharmaceutique
Mme FALLONE F.	Toxicologie
Mme FERNANDEZ-VIDAL A.	Toxicologie
Mme GADEA A.	Pharmacognosie
Mme HALOVA-LAJOIE B.	Chimie Pharmaceutique
Mme JOUANJUS E.	Pharmacologie
Mme LAJOIE-MAZENC I.	Biochimie
Mme LEFEVRE L.	Physiologie
Mme LE LAMER A-C. (*)	Pharmacognosie
M. LE NAOUR A.	Toxicologie
M. LEMARIE A.	Biochimie
M. MARTI G.	Pharmacognosie
Mme MONFERRAN S	Biochimie
M. PILLOUX L.	Microbiologie
M. SAINTE-MARIE Y.	Physiologie
M. STIGLIANI J-L.	Chimie Pharmaceutique
M. SUDOR J. (*)	Chimie Analytique
Mme TERRISSE A-D.	Hématologie
Mme TOURRETTE-DIALLO A. (*)	Pharmacie Galénique
Mme VANSTEELANDT M.	Pharmacognosie

(*) Titulaire de l'habilitation à diriger des recherches (HDR)

Enseignants non titulaires

Assistants Hospitalo-Universitaires

M. AL SAATI A	Biochimie
Mme BAKLOUTI S.	Pharmacologie
Mme CLARAZ P.	Pharmacie Clinique
Mme CHAGNEAU C.	Microbiologie
Mme DINTILHAC A.	Droit Pharmaceutique
M. LE LOUEDEC F.	Pharmacologie
Mme RIGOLOT L.	Biologie Cellulaire, Immunologie
Mme STRUMIA M.	Pharmacie Clinique

Attaché Temporaire d'Enseignement et de Recherche (ATER)

Mme HAMZA Eya	Biochimie
Mme MALLI Sophia	Pharmacie Galénique
M. TABTI Redouane	Chimie Thérapeutique

Table des matières

Listes des abréviations 7

1 Introduction 8

1.1 *Le vieillissement : un problème de santé publique* 8

1.2 *Une nouvelle discipline : La gérosceince* 9

1.3 *La compréhension multidimensionnelle du vieillissement : Une approche complexe et intégrative* 11

1.4 *Stratégies thérapeutiques en développement pour lutter contre les effets délétères du vieillissement* 12

1.5 *Le machine learning : un nouvel outil d'intérêt* 13

2 Vieillir en bonne santé : état actuel des connaissances 14

2.1 *Le vieillissement* 14

2.1.1 Clinique : de la robustesse à la fragilité 14

2.1.2 Systémique et cellulaire : Homéostasie et allostasie 17

2.1.3 Mécanismes cellulaires et moléculaires du vieillissement 19

2.1.3.1 Lésions génétiques 20

2.1.3.2 Télomérase 21

2.1.3.3 La sénescence 22

2.1.3.4 Épигénétique 23

2.1.3.5 Protéostase 23

2.1.3.6 Dysfonction mitochondriale 24

2.1.3.7 Épuisement des cellules souches 25

2.1.3.8 La dysbiose 26

2.1.3.9 Inflammation 26

2.1.4 Modèles d'études du vieillissement 28

2.1.4.1 Différentes échelles d'âges 28

2.1.4.1.1 Age chronologique 28

2.1.4.1.2 Age physiologique ou biologique 29

2.1.4.2	Modèles d'études	30
2.1.4.2.1	Modèles animaux	30
2.1.4.2.2	In vitro	33
2.1.4.2.3	Les pathologies humaines de vieillissement accéléré	34
2.2	Les stratégies thérapeutiques	35
2.2.1	Stratégies actuelles : non médicamenteuse	35
2.2.1.1	La nutrition	35
2.2.1.2	L'activité physique	38
2.2.2	Stratégies en développement : médicamenteuses	39
2.2.2.1	La senescence	39
2.2.2.2	Les senothérapeutiques	40
3	Le machine learning	44
3.1	Généralités	44
3.1.1	Une nouvelle ère : la recherche translationnelle	44
3.1.2	Intérêt par rapport aux statistiques	45
3.1.3	Découverte et intérêt de l'IA en santé et en sciences	46
3.2	Définition et principes de base du machine learning	49
3.2.1	Différents types d'apprentissage	50
3.2.1.1	L'apprentissage supervisé	50
3.2.1.2	L'apprentissage non supervisé	51
3.2.1.3	L'apprentissage par renforcement	52
3.2.2	La création des modèles : Différents algorithmes utilisés	53
3.2.2.1	Algorithmes pour l'apprentissage supervisé	55
3.2.2.1.1	La régression linéaire simple ou multiple	55
3.2.2.1.2	Le support vector machines (SVM)	57
3.2.2.1.3	KNN ou K Nearest Neighbor	58
3.2.2.1.4	Les forêts aléatoires (random forest et XG-boost)	59

3.2.2.2	Algorithmes pour l'apprentissage non supervisé	62
3.2.2.2.1	K-means	62
3.2.2.2.2	Analyse en composantes principales (ACP)	63
3.2.2.2.3	t-distributed Stochastic Neighbor Embedding (t-SNE)	64
3.2.3	La particularité du deep learning ou apprentissage profond	66
3.2.4	Explicabilité des modèles ou la compréhension des facteurs qui influencent le modèle	68
3.3	<i>Applications actuelles</i>	70
3.3.1	L'analyse des « big data »	70
3.3.2	Prédictions	73
3.3.2.1	La recherche clinique : L'âge biologique et les trajectoires de vieillissement	73
3.3.2.2	La recherche fondamentale : Gènes, voies dérégulées	77
3.3.3	Avancées thérapeutiques	78
3.3.3.1	Nouveau paradigme dans l'approche de découverte thérapeutiques	78
3.3.3.2	Intérêt du machine learning dans la recherche de thérapeutiques	80
3.3.3.3	La médecine de précision	84
4	Perspectives	85
5	Bibliographie.	89

Listes des abréviations

ACP : Analyse en Composantes Principales.
ADL : Activités de la vie quotidienne (Activities of Daily Living).
ADMmt : ADN Mitochondrial.
ADN : Acide Désoxyribonucléique.
AMM : Autorisation de Mise sur le Marché.
AMPK : AMP-activated protein kinase.
ARN : Acide Ribonucléique.
ATP : Adénosine Triphosphate.
C.elegans : Caenorhabditis elegans.
CI : Capacité intrinsèque.
CRP : C-réactive protéine (protéine réactive C).
DL : Deep Learning.
DRESS : Direction de la recherche, des études, de l'évaluation et des statistiques (France).
FDA : Food and Drug Administration (Agence du médicament au États-Unis).
IA : Intelligence Artificielle.
IADL : Activités Instrumentales de la Vie Quotidienne (Instrumental Activities of Daily Living).
IL6 : Interleukine 6.
INSEE : Institut national de la statistique et des études économiques (France).
IPSC : Cellules Souches Pluripotentes Induites (Induced Pluripotent Stem Cells).
IR : Infrarouge.
IRM : Imagerie par Résonance Magnétique.
K-means : K-moyennes.
ML : Machine Learning.
NAD : Nicotinamide Adénine Dinucléotide.
PBMC : Cellules mononucléaire périphérique
ROS : Reactive Oxygen Species (Espèces Réactives de l'Oxygène).
SASP : Senescence-Associated Secretory Phenotype.
SHAP : SHapley Additive exPlanations.
SIRT1 : Sirtuin 1.
SVM : Support Vector Machine.
t-SNE : t-distributed Stochastic Neighbor Embedding.
TORC1 : Target of Rapamycin Complex 1.
UV : Ultraviolets.
XG-boost : Extreme Gradient Boosting.

1 Introduction

1.1 Le vieillissement : un problème de santé publique

En France, l'espérance de vie des hommes et des femmes est en constante augmentation (Figure 1).

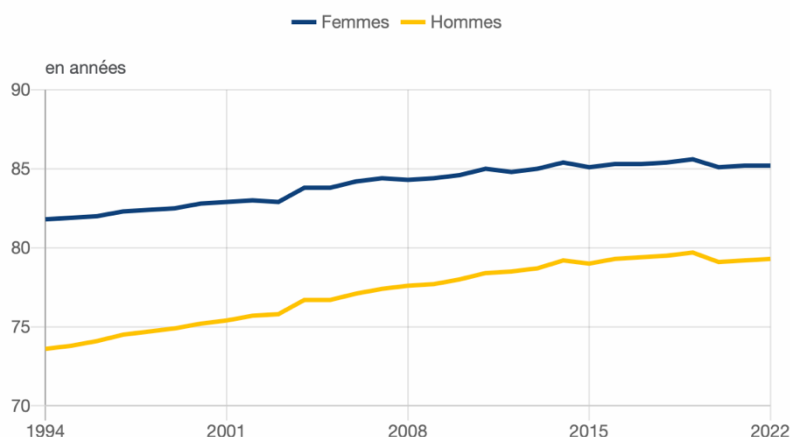


Figure 1 : Espérance de vie à la naissance des hommes et des femmes en France
INSEE, 2023

De plus, les dernières projections de l'INSEE pour 2070 font état d'une proportion d'individus de plus de 65 ans de 29% (INSEE, 2023). La population est donc **vieillissante**. Ces résultats peuvent notamment être attribués à l'avancée de la médecine avec une réduction de la mortalité précoce, une meilleure prévention et prise en charge thérapeutique des maladies infectieuses et cancers (antibiotique, vaccin, immunothérapies...) et à l'amélioration de l'accès aux soins (offre de soin, sécurité sociale et remboursement). Cependant, il est important de noter que l'espérance de vie ne correspond pas nécessairement à l'espérance de vie en bonne santé. Selon l'OMS, la santé est définie comme un état de complet bien-être physique, mental et social, et ne se limite pas simplement à l'absence de maladie ou d'infirmité. Cette définition met en lumière la notion que la santé englobe des aspects tels que le bien-être mental, émotionnel et social, allant au-delà de la simple absence de pathologies. A l'heure actuelle,

l'espérance de vie en bonne santé se situe aux alentours de 65 ans pour les hommes et les femmes et ne connaît pas la même progression que l'espérance de vie (DRESS, 2023). Il apparaît donc que **les individus vivront plus longtemps mais en incapacité ou mauvaise santé**. En effet l'augmentation de la durée de vie s'accompagne à l'échelle populationnelle d'une augmentation des pathologies chroniques, telles que les maladies cardiovasculaires (Hypertension, accident coronaire et cerebro-vasculaire), les maladies neurodégénératives (maladie de Parkinson, maladie d'Alzheimer), les maladies métaboliques (diabète, goutte, obésité) ou encore l'apparition de cancers (Niccoli et Partridge, 2012). Le vieillissement pose donc dans un premier temps un réel problème **sociétal**. En effet, les personnes âgées sont soumises à une poly médication complexe qui peut favoriser le risque de dépendance (état durable de la personne entraînant des incapacités et requérant des aides pour réaliser des actes de la vie quotidienne) et d'institutionnalisation ce qui impacte leur qualité de vie.

Le second enjeu est **économique** car une telle démographie exerce des pressions sur le système de santé. Selon le Ministère de la Santé et des Solidarités, le déficit de la branche maladie et le déficit de la branche retraite du régime général se creusent car l'augmentation de l'âge est directement corrélée à la consommation de soins médicaux et aux besoins sociaux (INSEE, 2023). En ce sens, l'objectif actuel est de **promouvoir le vieillissement en bonne santé** (Campisi et al, 2019).

1.2 Une nouvelle discipline : La gérosceience

L'émergence de la gérosceience en 2011, initiée par le « Gerosceience Interest Group », atteste de l'importance actuelle de comprendre les **mécanismes sous-jacents au vieillissement**. Cette discipline novatrice prône une approche interdisciplinaire, plaçant le vieillissement au centre de la réflexion (Sierra F, 2016). La gérosceience se distingue par son ambition d'explorer les

liens entre les processus biologiques liés au vieillissement et le développement des maladies associées à l'âge. L'hypothèse sous-jacente suggère que diverses affections partagent des **mécanismes moléculaires, métaboliques et physiologiques communs**, malgré des conséquences cliniques distinctes. Plutôt que de traiter chaque maladie liée à l'âge de manière isolée, la gérosceince cherche à identifier les processus fondamentaux qui sous-tendent le vieillissement et à développer des approches thérapeutiques visant à ralentir ou à atténuer ces processus (Kemoun et al, 2021). En effet, de nombreux mécanismes communs observés au cours du vieillissement sont associés à diverses maladies chroniques comme l'ostéoporose, la fragilité, le diabète, l'arthrose, la démence, ou les cancers (Guo et al, 2022 ; Figure 2).

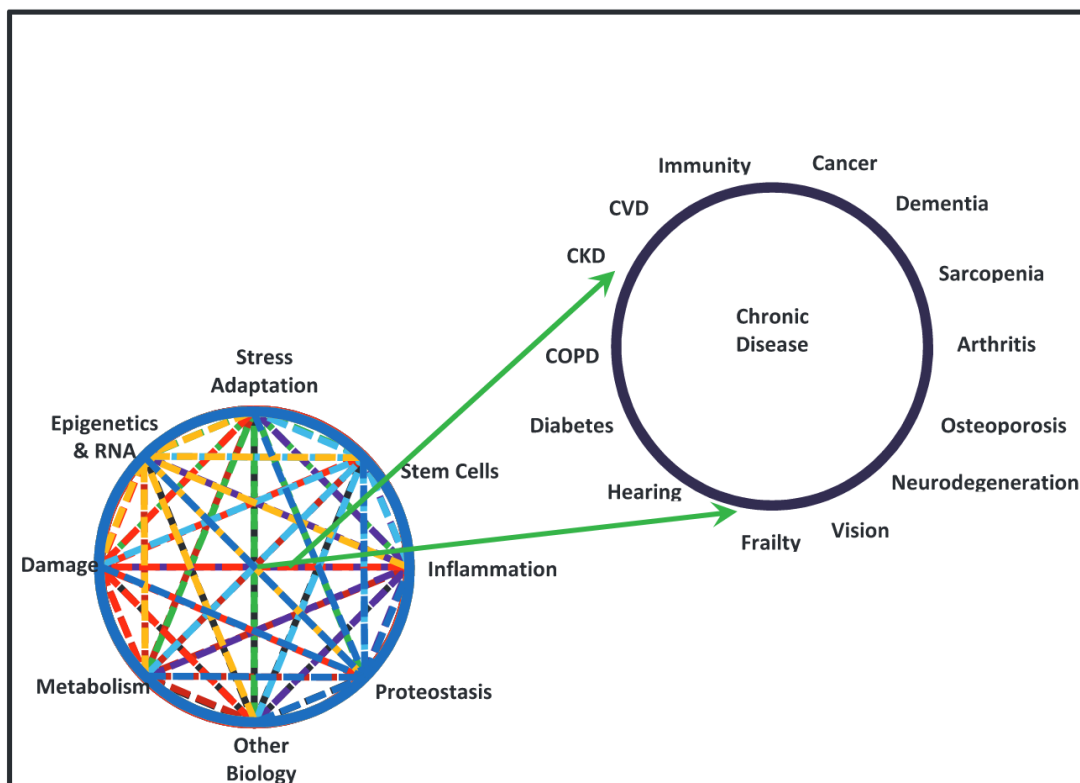


Figure 2 : Mécanismes transverses à l'origine des maladies chroniques liées à l'âge
Burch et al, 2014

Parmi les applications de la gérosceince, il y a la volonté **d'anticiper** les signes cliniques ou biologiques révélateurs de phénotypes cliniques de vieillissement en mauvaise santé, tels que la diminution de la **capacité intrinsèque** (CI) et la **fragilité** (Palliyaguru DL et al, 2019). La

capacité intrinsèque (CI) représente une approche novatrice axée sur la fonction, introduite par l'Organisation mondiale de la santé (OMS) dans le but de promouvoir une longévité en bonne santé en atténuant les conséquences du handicap et de la dépendance liés au vieillissement (Beard et al, 2016). La CI est définie comme la conjonction des capacités physiques et mentales, lesquelles, en association avec les facteurs environnementaux, déterminent la capacité fonctionnelle d'un individu à faire ce qu'il souhaite et à être intégré dans la société (Cesari et al, 2018). La fragilité au sens large, est un état clinique progressif lié à l'âge, caractérisé par une détérioration des capacités physiologiques conduisant à une vulnérabilité accrue de l'individu (Clegg A et al, 2013) qui augmente le risque de vieillir en mauvaise santé en favorisant une entrée plus rapide dans la dépendance. De ce fait, un objectif actuel est d'identifier des **marqueurs biologiques prédictifs** de la fragilité et de la CI afin de suivre la trajectoire de vieillissement individuel et d'accompagner au mieux les individus.

1.3 La compréhension multidimensionnelle du vieillissement : Une approche complexe et intégrative

Le vieillissement est un processus physiologique obligatoire **multi-échelle et multi-factoriel** avec des **différences intra- et inter-individuelles** (López-Otín et al, 2013), ce qui en fait un processus complexe à étudier. Dans l'objectif de promouvoir le vieillissement en bonne santé à l'échelle **populationnelle**, il faut d'abord comprendre la physiologie du vieillissement à l'échelle de **l'individu** et ce à toutes les échelles de l'organisme (clinique, systémique, cellulaire) en développant différents **modèles d'études** complémentaires. Pour étudier le vieillissement d'un individu, il faut prendre en compte des facteurs génétiques, environnementaux mais également sociaux (accès aux soins, alimentation, activités

physiques, relations sociales). Pour un individu, il y a donc beaucoup d'informations ou données qu'il faut connecter entre elles ce qui constitue une complexité supplémentaire à la compréhension des mécanismes du vieillissement.

1.4 Stratégies thérapeutiques en développement pour lutter contre les effets délétères du vieillissement

Concernant les thérapeutiques, la recherche se concentre aujourd'hui sur une approche qui n'est plus pathologie centrée (un médicament pour une maladie) mais **fonction-centrée**. Ces fonctions que nous définirons par la suite regroupent des aspects essentiels de la santé liés au vieillissement comme la capacité de maintenir une mobilité adéquate, de garantir une cognition optimale ou encore de conserver une bonne santé psychologique. L'objectif est de limiter et ralentir les manifestations fonctionnelles du vieillissement plutôt que de traiter indépendamment chaque maladie (Felipe Sierra, 2016). Les thérapeutiques **actuelles sont non médicamenteuses**, basées sur la nutrition ou l'activité physique. Le vieillissement n'étant pas une pathologie, l'OMS et la FDA (Agence du médicament aux États-Unis) restent réticentes à autoriser la commercialisation de médicaments ayant une indication contre le vieillissement. Néanmoins, certains médicaments, initialement commercialisés pour d'autres indications et utilisés dans le cadre du repositionnement des autorisations de mise sur le marché (AMM), sont actuellement en essais cliniques. Ces essais visent à démontrer l'efficacité de ces médicaments en tant que thérapies préventives et curatives du vieillissement. Ces agents, regroupés sous le terme de "**géroprotecteurs**", sont conçus pour offrir une protection contre les processus liés au vieillissement (Moqri et al, 2023). Un autre besoin actuel est donc de faciliter la recherche, le développement et la validation de nouvelles thérapeutiques (Campisi et al, 2019).

1.5 Le machine learning : un nouvel outil d'intérêt

Au 17^{ème} siècle Isaac Newton concevait l'individu comme un ensemble de cellules qu'il fallait étudier de façon linéaire. Au 20^{ème} siècle, la science a changé de paradigme et l'individu est devenu un système **complexe** ou le « tout » est un ensemble **hyper connecté**. Au niveau biologique, le plus important n'est plus la cellule elle-même mais les interactions entre ces cellules. En transposant cette théorie à l'étude du vieillissement par exemple, il devient essentiel **d'incorporer une diversité de niveaux de compréhension**, englobant les découvertes au niveau des individus, des organes ou des cellules, au sein d'un réseau interconnecté (Hood L, 2004). Cette approche s'avère indispensable pour appréhender de manière complète les **mécanismes complexes** associés au processus de vieillissement. Dans le but d'assister l'Homme dans ses actions et ses raisonnements, une branche de l'intelligence artificielle que l'on appelle le **machine learning** (apprentissage automatique) a su faire ses preuves dans la recherche en santé en raison de sa capacité à analyser de grandes quantités de données, à repérer des modèles complexes et à fournir des informations précieuses pour la prévention, le diagnostic et le traitement des maladies. La recherche sur le vieillissement bénéficie grandement des applications du machine learning, que ce soit dans l'élaboration de modèles prédictifs et informatifs sur la physiologie du vieillissement ou dans la découverte et le développement de nouvelles thérapies.

Premièrement, nous résumerons les données disponibles concernant la recherche sur le **vieillessement en bonne santé** ainsi que les données concernant la recherche et le développement de nouvelles thérapeutiques. *Dans un second temps*, nous décrirons les **méthodes de machine learning** les plus courantes et utiles dans notre domaine ainsi que les

applications actuelles qui cherchent à créer des modèles prédictifs de la trajectoire du vieillissement et à simplifier la découverte de nouvelles thérapeutiques.

2 Vieillir en bonne santé : état actuel des connaissances

Le vieillissement en bonne santé est un processus continu d'optimisation des capacités de l'organisme afin de maintenir la santé physique et psychologique, l'indépendance et la qualité de vie des individus (OMS).

2.1 Le vieillissement

Le vieillissement physiologique (en dehors de toutes maladies) représente l'ensemble des modifications se produisant au cours de l'avancée en âge, qui interviennent à tous les niveaux de l'organisme (gènes, molécules, cellules, tissus). Bien qu'inhérents à l'espèce humaine, ces mécanismes sont mal connus et il n'est pas possible de savoir pourquoi seulement certains individus vieilliront en bonne santé. De plus, le vieillissement est le facteur de risque majeur du déclin des capacités intrinsèques associées aux maladies chroniques et à la fragilité (Niccoli et Partridge, 2012). Il est donc nécessaire de mieux appréhender les mécanismes du vieillissement afin de pouvoir limiter et/ou prévenir son impact sur les capacités intrinsèques des personnes âgées.

2.1.1 Clinique : de la robustesse à la fragilité

L'OMS a identifié **cinq domaines de fonctions** ou capacités intrinsèques (Figure 3) afin de suivre cliniquement l'évolution de la santé des personnes âgées, à savoir : la mobilité, la

cognition, la vitalité, l'aspect psychologique et la sensorialité (vision et audition) (World report on ageing and health, 2015).



Figure 3 : Cinq domaines de fonctions ou capacités intrinsèques définis par l'Organisation Mondiale de la Santé, afin de suivre cliniquement l'évolution de la santé des personnes âgées
Kemoun et al, 2022

La **capacité Intrinsèque (CI)**, comme mentionnée précédemment se définit comme la combinaison des capacités physiques et mentales d'un individu (Figure 4). Elle est le reflet des interactions avec l'environnement et des facteurs intrinsèques aux individus (Cesari et al, 2018).

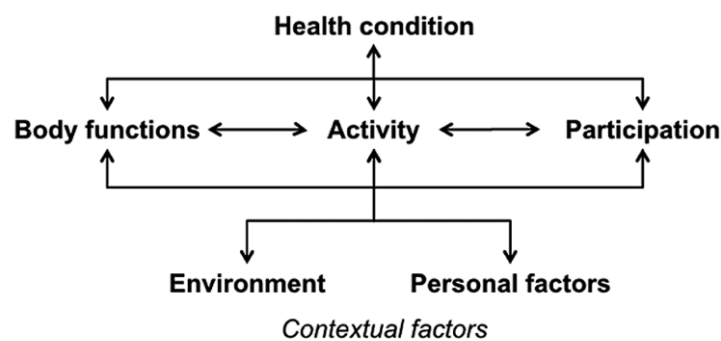


Figure 4 : Classification internationale du fonctionnement, du handicap et de la santé qui a permis le développement du concept de capacité intrinsèque.
Cesari et al, 2018

Un maintien optimal de ces capacités permet un vieillissement en bonne santé (Beard et al, 2016). Quand au moins l'une de ces fonctions décline on parle d'individus pré-fragiles tandis que le déclin des 5 fonctions fait basculer l'individu dans la fragilité. La **fragilité est un état clinique progressif lié à l'âge**, caractérisé par une détérioration des capacités physiologiques conduisant à une vulnérabilité accrue de l'individu (Clegg et al, 2013) qui augmente le risque de vieillir en mauvaise santé en favorisant une entrée plus rapide dans la dépendance. L'individu **robuste** possède quant à lui des capacités intrinsèques optimales que l'on retrouve généralement chez les individus jeunes et en bonne santé.

La fragilité peut être scorée de plusieurs façons (Yaksic et al, 2019). Ici, nous mentionnerons seulement les deux principales échelles utilisées en clinique actuellement. L'échelle de fragilité clinique de **Rockwood** (Rockwood et al, 2005) est un index de fragilité globale et multi-domaines. Il s'agit d'un questionnaire qui évalue les dimensions physiques et cognitivo-sociales au travers d'un index des déficits cumulés (paramètres médicaux, dépendance ADL/IADL, dépression, comportement, âge, isolement social). Il y a 7 stades, un sujet est dit « fragile » à partir du stade 4. C'est un score complexe qui nécessite la réalisation de plus de 20 tests cliniques. Le **score de FRIED** (Fried et al, 2001) se mesure quant à lui par la réalisation de 5 paramètres cliniques (perte de poids involontaire, asthénie, fatigabilité musculaire, activité physique réduite, marche ralentie) qui classe les individus en 3 catégories. Un sujet « robuste » ne présente aucun critère de FRIED tandis qu'un sujet « fragile » en présente au moins trois. Les sujets n'entrant dans aucune de ces deux catégories sont qualifiés de « pré-fragiles ». Cependant ce score n'est validé que pour des individus de plus de 60 ans et ne traite que des fonctions de locomotion et de vitalité. Globalement, ces critères sont **non exhaustifs** donc **peu optimaux**. Récemment une équipe Toulousaine a élaboré un nouveau score de

capacité intrinsèque. L'originalité de ce nouveau score réside dans sa capacité à identifier les individus à risque, notamment ceux dont les scores CI se situent dans le 10e centile le plus bas. Ce groupe présente une prédisposition accrue à des problèmes de santé variés, tels que des maladies chroniques, des difficultés dans les activités quotidiennes ou des chutes. Cette étude révèle aussi qu'une réduction des réserves fonctionnelles pourrait ne pas être perceptible avant un âge avancé (Lu et al, 2023). L'objectif actuel est donc de développer des scores qui révèlent des phénotypes à surveiller **en amont des déclin fonctionnels**, afin de **prévenir la fragilité**.

2.1.2 Systémique et cellulaire : Homéostasie et allostasie

Premièrement décrit par **Claude Bernard** en 1865 en tant que milieu intérieur dans son *Introduction à l'étude de la médecine expérimentale* (Bernard, 1865), l'état **d'homéostasie** correspond à la **stabilité du milieu intérieur** et au maintien des constantes vitales (Cannon, 1929). Cet état représente un flux stationnaire maintenu constant grâce à des paramètres de référence du milieu intérieur tels que la glycémie, la température, la pression artérielle ou le pH. Des systèmes de capteurs, d'effecteurs et de rétro-contrôles contrôlent les **systèmes de réponse de l'organisme** (à toutes les échelles de l'organismes : du système endocrinien aux systèmes enzymatiques, antioxydant ou aux mécanismes de déméthylation sur un plan très moléculaire etc) qui maintiennent les valeurs de références dans leur plage en situation physiologique. Au cours de la vie, l'individu est soumis à des **stress**, qu'ils soient biologiques (infections, défaillance aiguë d'organes), ou bien psychologiques (chocs émotionnels ou traumatiques) qui vont nécessiter une **adaptation** de la part de l'organisme. Les mécanismes d'adaptation qui vont se mettre en place vont pouvoir modifier les valeurs de référence afin

qu'elles soient optimales pour le maintien de la vie sans lésions. Cette capacité d'adaptation s'appelle la **réserve physiologique** (Clegg et al, 2013) (Figure 5).

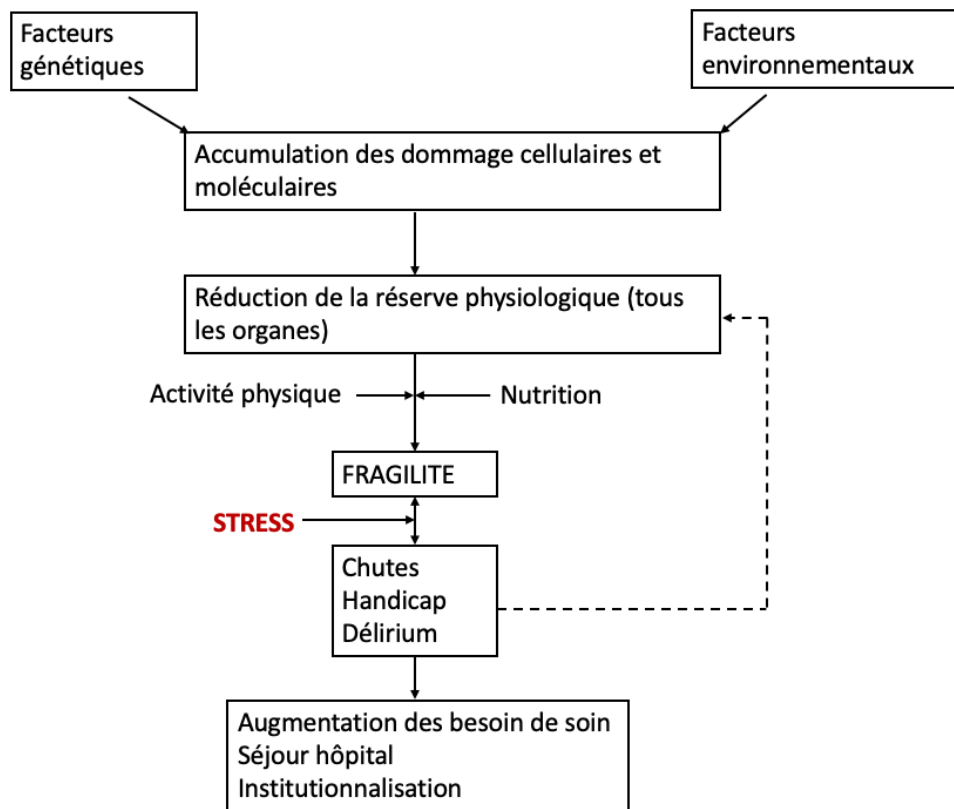


Figure 5 : De la perte de réserve physiologique à la fragilité
D'après Clegg et al, 2013

Cependant, l'accumulation de ces stress associée aux facteurs génétiques et environnementaux (Kirkwood et al, 2016) provoque, avec le temps, un **épuiement de la réserve physiologique**. De ce fait, il n'y a plus d'adaptation de la valeur de référence. La capacité de l'organisme à revenir dans un état initial est perturbée ce qui provoque la perte de l'équilibre du milieu intérieur qu'on appelle aussi l'**état d'allostasie** (Figure 6). L'**allostasie** représente une dérégulation biologique systémique au sein de l'organisme qui devient finalement visible au niveau clinique de par la perte de fonction (menant à la fragilité) et le recours accru à des soins médicaux (menant à la dépendance). L'individu jeune ou robuste aura une **capacité adaptative** ou **résilience** optimale afin de maintenir la stabilité de son

environnement intérieur tandis que l'individu âgé ou fragile conservera des **séquelles** cliniques ou biologiques à la suite de ces stress. Par exemple, l'insulino-résistance (souvent observée chez les sujets âgés) exposera l'organisme à un stress hyperglycémique prolongé pouvant conduire à un **diabète de type 2**. Lorsqu'une personne est exposée à un stress prolongé, le niveau de **cortisol** (hormone produite en pics le matin et qui diminue tout au long de la journée) peut augmenter de manière significative. À court terme, cela peut être une réponse normale et adaptative, mais lorsqu'il devient chronique, cela peut conduire à une résistance au cortisol. Cette dernière peut être à l'origine de **troubles immunitaires** (augmentation de la susceptibilité aux infections), **métaboliques** (insulino-résistance) ou encore de troubles de **sommeil** qui auront des conséquences cliniques (Adam et al, 2017).

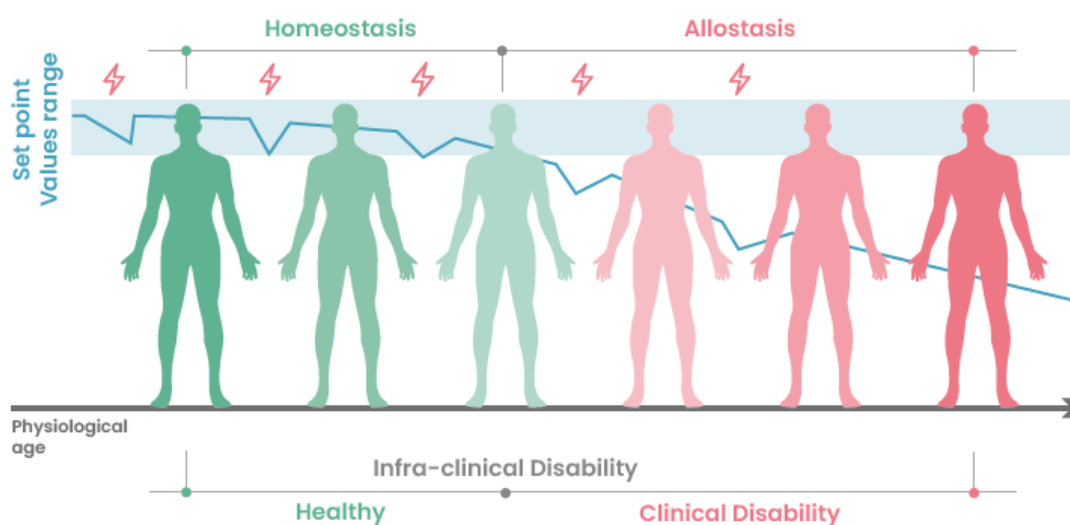


Figure 6 : Bascule de l'état d'homéostasie vers l'état d'allostasie associé à des conséquences cliniques délétères

Ader et al, 2021 ; Kemoun et al 2022

2.1.3 Mécanismes cellulaires et moléculaires du vieillissement

Les mécanismes cellulaires et moléculaires ayant un impact sur les processus de vieillissement, ont été récapitulés sous le terme de piliers du vieillissement ou « **Hallmarks of aging** ». Ces piliers peuvent être catégorisés en hallmarks primaires, qui conduisent à des dommages cellulaires, en hallmarks antagonistes, qui représentent les réponses et

mécanismes compensatoires face aux dommages, et en hallmarks intégratifs, une combinaison des deux premiers responsables de la perte de fonction et du phénotype de vieillissement observé. Les hallmarks évoluent constamment ; initialement regroupés en 7 (Kennedy et al, 2014) ou 9 piliers (Lopez-Otin et al, 2013), ils sont aujourd'hui bien plus nombreux (Schmauck-Medina et al, 2022 ; Lopez-Otin et al, Cell 2023) (figure 7) et représentent la **physiologie globale** de l'individu.

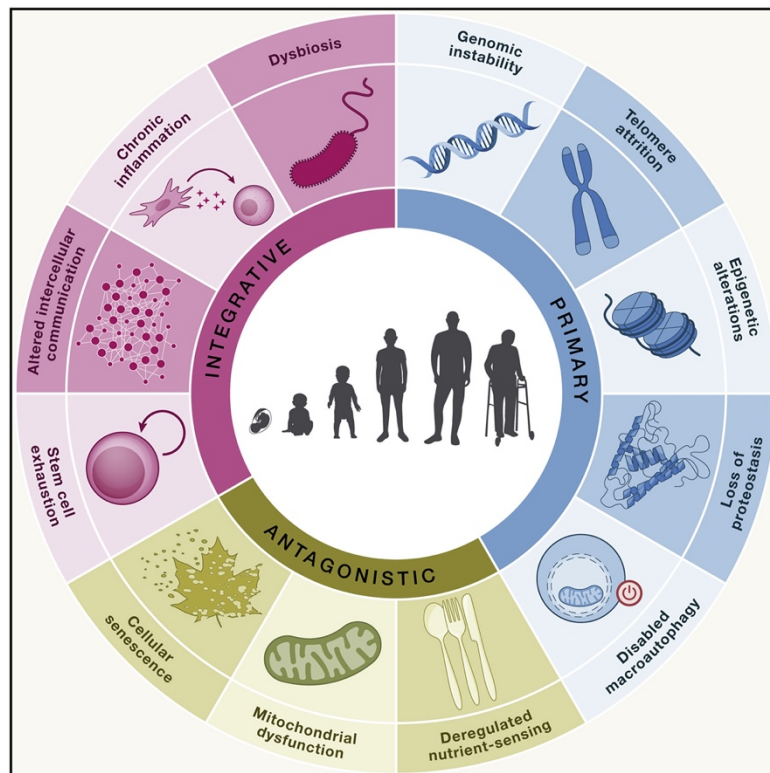


Figure 7 : Les piliers du vieillissement
Lopez-Otin et al, 2023

Nous allons explorer certains de ces piliers pour mieux appréhender leur importance tant dans la compréhension des mécanismes que dans le développement de thérapies visant à contrer le processus de vieillissement.

2.1.3.1 Lésions génétiques

Au cours du vieillissement, on observe l'accumulation de lésions génétiques liées à divers dommages extrinsèques (chimique, radiation UV, IR) et intrinsèques (espèces réactives de

l'oxygène ROS, erreurs de réplication) qui **compromettent la stabilité et l'intégrité de l'ADN** (Schumacher et al, 2021).

En réponse à ces lésions, l'organisme en bonne santé dans un état d'homéostasie met en place des systèmes de réparation de l'ADN. Parmi ces **mécanismes de réparation** on retrouve, la **ligature d'extrémités non homologues** et la **recombinaison homologue** pour les doubles brins ou encore la **réparation par excision de base** en cas de dommages à un seul nucléotide sur un simple brin. De plus, des enzymes telles que la superoxyde dismutase, la catalase ou la glutathion peroxydase, qui permettent de limiter le taux de radicaux libre (ROS) sont dérégulées au cours du vieillissement contribuant à l'augmentation des lésions génétiques.

2.1.3.2 Télomérase

Il existe une différence dans l'activité de la télomérase entre les cellules souches et somatiques qui réside dans leur capacité à influencer la longueur des télomères, des séquences répétitives situées à l'extrémité des chromosomes. Dans les cellules souches, la télomérase est active et maintient la longueur des télomères. Cela signifie que, même lors de divisions cellulaires successives, les télomères ne subissent pas de raccourcissement significatif, préservant ainsi l'intégrité génétique et la capacité de ces cellules à se renouveler de manière régénérative. Cette caractéristique distingue les cellules souches, qu'elles soient embryonnaires ou adultes, en leur permettant de conserver leur potentiel de différenciation et leur capacité de division à long terme. En revanche, dans les **cellules somatiques**, la télomérase devient généralement inactive ou présente une activité limitée au fil du temps. Cela entraîne un raccourcissement progressif ou **attrition des télomères** à chaque division cellulaire. Lorsque les télomères deviennent trop courts, la cellule entre en sénescence. Cette réduction de la longueur des télomères dans les cellules somatiques est associée au processus

de vieillissement et à diverses conditions pathologiques. De plus, il a été montré chez la souris qu'une activation de la télomérase pouvait ralentir le vieillissement physiologique sans augmenter l'incidence des cancers ce qui en fait un levier d'action potentiel (Way JW et al, 2016).

2.1.3.3 La sénescence

La sénescence, initialement décrite en 1961 par Hayflick and Moorhead, correspond à l'arrêt irréversible de la prolifération cellulaire induit par le raccourcissement des télomères (Campisi et al, 2018) en réponse aux stress et aux dommages cellulaires. Elle s'accompagne de la sécrétion de facteurs pro-inflammatoires regroupés dans le phénotype sécrétoire associé à la sénescence **ou SASP** (Li et al, 2023). La sénescence est **globale et intégrative** car elle est la résultante de nombreux mécanismes présent au cours du vieillissement (voir les autres piliers du vieillissement), ce qui en fait une cible de choix dans le développement de thérapeutique (Zhang L et al, 2022) que nous détaillerons dans le paragraphe consacré aux sénotheraques. Il existe une sénescence **physiologique** plurifactorielle, de bas grade tout au long de la vie (Lozano-Torres et al, 2019). Les cellules sénescents sont présentes chez tous les individus. En effet, la sénescence est un mécanisme obligatoire, qui n'est pas pathologique s'il reste **transitoire**. La senescence permet notamment l'arrêt de prolifération des cellules endommagées (rôle dans la suppression des tumeurs), favorise la régénération cellulaire après une blessure ou encore permet la vascularisation du placenta (rôle dans l'accouchement) (Muñoz-Espín et al, 2014). Il existe également une sénescence **induite** par le stress, prématurée, aiguë et pathologique (Lozano-Torres et al, 2019) qui contribue au déclin du potentiel fonctionnel et régénératif des tissus, à l'inflammation mais aussi à l'oncogenèse dans les organismes vieillissants (Hernandez-Segura et al, 2018). De plus, chez les personnes

âgées, les mécanismes d'élimination des cellules sénescents sont défaillant ce qui conduit à **l'accumulation** dans l'organisme de cellules endommagées et non fonctionnelles qui contribuent au développement de **maladies chroniques**, métaboliques, neurodégénératives etc.

2.1.3.4 Épigénétique

L'épigénétique englobe les modifications de l'expression génique qui ne découlent pas de changements dans la séquence d'ADN elle-même. Ces ajustements épigénétiques ont le pouvoir d'influencer l'activation ou la répression des gènes, jouant ainsi un rôle crucial dans la régulation de divers processus biologiques. Elle concerne des **facteurs** qui ne sont pas déterminés par la séquence d'ADN. Des modifications épigénétiques, telles que la **méthylation** de l'ADN (Unnikrishnan A et al, 2018), **l'acétylation** des histones, ont des impacts durables sur les cellules tout au long de la vie. Ces facteurs épigénétiques peuvent être sensibles à des **influences environnementales** telles que la nutrition, le stress, les hormones, les médicaments et les expositions à des polluants au fil du temps. Dans un organisme sain, des **systèmes enzymatiques** régulent ces altérations. Cependant, au cours du vieillissement, des changements dans la méthylation de l'ADN, par exemple, peuvent contribuer à l'activation de gènes oncogènes ou à l'inhibition de gènes suppresseurs de tumeurs, favorisant ainsi le développement du cancer (Pal et al, 2016).

2.1.3.5 Protéostase

La **protéostase, ou homéostasie des protéines**, est essentielle pour le bon fonctionnement des cellules, garantissant l'intégrité structurale et fonctionnelle des protéines par des mécanismes complexes de repliement, de stabilisation, et de dégradation. Il est connu que

l'activité des deux principaux systèmes protéolytiques impliqués dans le contrôle de la qualité des protéines, à savoir le système autophagie-lysosomique et le système ubiquitine-protéasome, diminuent avec le vieillissement (Klaips et al, 2018). Lorsque ces systèmes de régulation ne parviennent plus à maintenir l'équilibre, des perturbations importantes surviennent, pouvant contribuer au développement de maladies neurodégénératives liées à l'âge telles que la maladie d'Alzheimer (Cozachenco et al, 2023) ou de Parkinson. Ces dernières se caractérisent par l'accumulation d'agrégats protéiques, perturbant le fonctionnement normal du système nerveux central et entraînant des symptômes de démences contribuant à la perte de fonction cognitive. Ces défauts de protéostase sont aussi retrouvés dans des pathologies comme la sarcopénie qui contribue quant à elle à la perte de mobilité chez le sujet âgé (Paez et al, 2023).

2.1.3.6 Dysfonction mitochondriale

La mitochondrie est un organite que l'on retrouve dans la plupart des cellules (hors érythrocytes matures). Elle assure la **bioénergétique** de l'organisme car elle est le lieu de formation principal **d'ATP** (Adénosine Triphosphate), le carburant énergétique des cellules ainsi que du flux **REDOX** (flux d'électron). La synthèse d'ATP qui se fait par phosphorylation oxydative est couplée à la respiration mitochondriale. La respiration se fait par génération d'un potentiel transmembranaire mitochondrial à l'aide d'une pompe à proton via les complexes mitochondriaux I, III et IV de la chaîne de transport des électrons. En raison de la disponibilité d'oxygène dans les mitochondries, celles-ci génèrent des espèces réactives de l'oxygène (ROS) (Angelova et al, 2018). Avec l'âge, on observe une défaillance de la chaîne respiratoire mitochondriale associée à une diminution d'énergie pour les cellules. Cette défaillance pourrait être à l'origine d'une production accrue de ROS ce qui endommagerait la

chaîne respiratoire elle-même (Guo et al, 2023). D'autres mécanismes contribuent aux dysfonctionnements bioénergétiques comme l'accumulation de mutations dans l'ADN mitochondrial, l'oxydation des protéines mitochondriales, les altérations de la composition lipidique des membranes mitochondriales, les perturbations de la dynamique mitochondriale, et les défauts dans la mitophagie, un processus de dégradation spécifique aux mitochondries défectueuses. De plus, l'entraînement d'endurance et le jeûne régulier semblent atténuer la dégénérescence mitochondriale, potentiellement en induisant l'autophagie (Memme et al, 2021), ce qui souligne l'importance des mécanismes cellulaires dans le vieillissement et la possibilité d'interventions liées au mode de vie pour retarder ces processus dégénératifs.

2.1.3.7 Épuisement des cellules souches

Les cellules souches sont des cellules indifférenciées, présentes chez tous les êtres vivants multicellulaires, qui peuvent se différencier en cellule spécialisées ou se maintenir à l'état souche via une division symétrique ou asymétrique. Elles jouent un rôle essentiel dans le **développement et le maintien de l'intégrité** de l'organisme. Le **pool** de cellules souches présent à la naissance diminue avec l'âge dû notamment à l'attrition des télomères et à l'augmentation des dommages à l'ADN qui favorisent la sénescence des cellules. Progressivement il s'installe donc une diminution des capacités régénératives au sein des tissus. Cet épuisement des cellules souches, à la fois de leur nombre mais aussi de leur capacité régénératrice, est l'une des causes majeures du vieillissement de l'organisme (Sameri et al, 2020).

2.1.3.8 La dysbiose

La dysbiose correspond au déséquilibre de la composition du **microbiote**. Le microbiote est l'ensemble des **micro-organismes** qui colonisent la muqueuse intestinale. Il est indispensable pour des fonctions de catabolisme, de biosynthèse des vitamines et acides aminés mais aussi dans la détoxification de xénobiotique. Il a également un rôle dans la modulation du système immunitaire et dans la résistance aux pathogènes. Au cours du vieillissement, la composition des bactéries de notre corps est **modifiée** (Biagi et al, 2017). On observe une **accumulation** de la charge bactérienne lors du vieillissement. La flore intestinale et les métabolites associés affectent directement ou indirectement les activités biologiques de l'hôte (Maynard et al, 2018). Il a été mis en évidence des liens entre la composition du microbiote et la longévité en bonne santé, avec des rôles dans l'immunité innée, la sarcopénie, la fragilité, et les dysfonction cognitive (O'Toole et al, 2015).

2.1.3.9 Inflammation

L'inflammation est un processus complexe impliquant des réponses immunitaires et des réactions cellulaires qui ont pour **but de protéger** l'organisme contre des agressions extérieures par des micro-organismes tels que les bactéries, les virus ou des toxines. Au cours du vieillissement, il existe une augmentation générale de l'inflammation appelée « **inflamm-aging** » (Franceschi et al 2007) qui est liée à plusieurs facteurs. Les cellules endommagées, des produits de dégradation de constituants cellulaires libérés dans certaines conditions de mort cellulaire s'accumulent avec l'âge (Franceschi et al, 2008), **augmentant la charge en antigènes** et déclenchant une réponse immunitaire. Cette réponse, caractérisée par une production accrue de cytokines pro-inflammatoires (TNF-alpha, IL-6 etc), contribue à la dysfonctionnalité des tissus et à l'apparition de maladie liées à l'âge. Cependant, les récepteurs de

reconnaissance de l'inflammation (TLR, IL-6R, CLR, NLR etc), les molécules de signalisation, les cytokines ou encore les cellules immunitaires impliqués dans l'immunité innée ou adaptative sont régulées de manière moins efficace avec l'âge. De ce fait, il n'y a pas de résolution complète de l'inflammation en réponse aux pathogènes ce qui crée une **inflammation constante de bas bruit** (Franceschi et al, 2018). Les dysfonctionnements immunitaires et les changements dans les voies inflammatoires contribuent au vieillissement, créant un **effet d'entraînement** car l'inflammation favorise elle-même les autres marqueurs du vieillissement. En effet l'inflammation altère la fonction mitochondriale, favorise la sénescence ou encore dérégule les voies de l'autophagie. L'inflammation chronique est identifiée comme une cible potentielle pour traiter le vieillissement avec par exemple la stimulation du système immunitaire pour éliminer les cellules sénescents. Bien que l'application clinique reste à prouver, les approches de ciblage immunitaire similaire aux thérapies anticancéreuses ont un fort potentiel (Baechle et al, 2023).

Dans le contexte d'un vieillissement pathologique, les mécanismes de réparation et de contrôle de ces piliers du vieillissement fonctionnent de manière moins efficace, entravant ainsi l'adaptation de l'organisme. Cibler de manière individuelle ou combinée les piliers du vieillissement représente une nouvelle orientation thérapeutique visant à adopter une approche intégrative face aux processus sous-jacents aux maladies plutôt que de se concentrer sur une pathologie spécifique. Actuellement, ces marqueurs reflètent la physiologie globale de l'organisme, ce qui rend difficile leur hiérarchisation ou la focalisation sur un seul aspect du vieillissement. Ces piliers acquièrent une plus grande pertinence lorsqu'ils sont envisagés comme un réseau interconnecté plutôt que comme des processus isolés. Les interactions complexes entre eux au cours des processus physiopathologiques

rendent leur description et leur étude plus compliquées, nécessitant ainsi le développement de nouveaux outils pour une compréhension plus approfondie.

2.1.4 Modèles d'études du vieillissement

2.1.4.1 Différentes échelles d'âges

2.1.4.1.1 Age chronologique

L'âge chronologique est un indicateur **simple, direct et non discutable** qui correspond au nombre d'années écoulées depuis la naissance. Il est utilisé dans de nombreux contextes, tels que la médecine, la recherche, l'assurance, la planification de la retraite, etc. Il apporte une information sur la durée de vie d'un individu et ces applications diverses sont intéressantes à l'échelle populationnelle. Par exemple, il trouve un intérêt dans les études épidémiologiques ou pour développer des campagnes de prévention et des stratégies globales de traitement liées à l'âge. En recherche, il permet d'établir des profils de santé toujours à l'échelle populationnelle car il nous informe sur le déclin global supposé des différents systèmes et organes avec l'âge (Figure 8). Bien qu'utile dans de nombreux domaines, il présente donc des limites dans la compréhension complète du processus de vieillissement car il ne capture pas pleinement la diversité des trajectoires de vieillissement entre les individus. En effet, l'âge chronologique ne prend **pas en compte les différences inter-individuelles** en matière de santé, de style de vie, de génétique, de stress, de facteurs environnementaux, qui peuvent influencer significativement le vieillissement d'une personne. (Ferrucci et al, 2020).

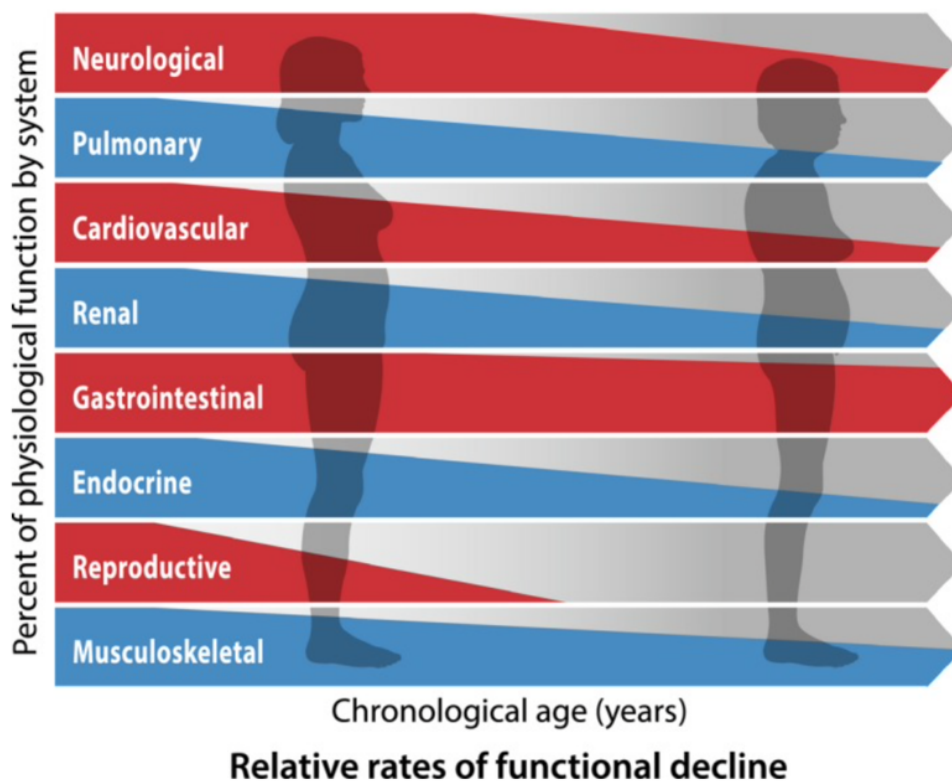


Figure 8 : Déclin des fonctions physiologiques selon l'âge chronologique
Khan et al, 2017

2.1.4.1.2 Age physiologique ou biologique

La notion d'âge biologique est apparue en 1988 (Baker et Spratt, 1988) et a remis en question l'utilisation de l'âge chronologique comme prédicteur de la physiologie de l'individu. Aujourd'hui on sait que l'âge chronologique n'a que peu de valeur en comparaison avec l'âge physiologique qui permet de comprendre l'hétérogénéité **inter-individuelle** du **vieillessement**. En effet pour des individus du même âge chronologique, on peut observer un âge biologique ou physiologique différent. L'âge physiologique est un concept qui intègre les dégradations moléculaires et cellulaires de l'organisme (Moqri et al, 2023). Cet « âge » fictif reflète la capacité d'adaptation et donc l'**usure physiologique** de l'individu ce qui permet d'établir une potentielle durée de santé fonctionnelle des individus. L'âge physiologique représente un facteur de risque dans le développement de maladies chroniques et métaboliques qui

augmentent le risque de fragilité et de dépendance (Ferrucci et al, 2020). Il pourrait être utilisé pour développer des programmes de prévention et de traitement ciblés dans le cadre d'une médecine personnalisée (Moqri et al, 2023). Ces dernières années, la prédiction (que nous détaillerons ensuite) de l'âge biologique a été réalisée en utilisant une gamme variée de biomarqueurs, tels que la longueur des télomères, la méthylation de l'ADN, la protéomique, la métabolomique, la glycomique, les données de capteurs portables, et les biomarqueurs cliniques sanguins. Certaines études tentent d'établir des **scores** pour définir un âge biologique à partir de l'étude de certains piliers du vieillissement (Hartmann et al, 2023) ou de biomarqueurs sanguins (Bots et al, 2023). **Aucun consensus** n'a encore été établi car les cohortes d'études doivent être larges et prendre en compte de nombreux paramètres et variables qu'il est difficile d'analyser.

2.1.4.2 Modèles d'études

D'un point de vue **éthique**, il n'est pas possible d'étudier directement le vieillissement chez les sujets âgés fragiles. Il est donc nécessaire de développer des modèles permettant de récapituler au mieux les phénotypes et mécanismes intrinsèques du vieillissement.

2.1.4.2.1 Modèles animaux

Le modèle d'étude du vieillissement dépend de la question posée. De façon générale les modèles animaux sont de plus en plus controversés de par leur éloignement de la physiologie humaine et de par les nouvelles directives en matière d'éthique (Règles des 3R : Réduction, Raffinement, Remplacement). Pour autant, à l'heure actuelle, ces modèles restent **nécessaires** dans les études de longévité (Taormina et al, 2019) et de compréhension des mécanismes et voies impliqués dans le vieillissement. Nous détaillerons donc dans les premier

temps les modèles usuels utilisés pour les études sur les mécanismes du vieillissement. Les **modèles classiques unicellulaires** comme les levures, ou **d'invertébrés** comme les nématodes ou la drosophile ont joué un rôle déterminant dans l'identification des gènes conservés et des voies qui retardent le vieillissement et prolongent la durée de vie (Guarente et al, 2000). Par exemple le **Caenorhabditis elegans**, un ver nématode, qui est exploité pour ses cycles de vie courts et sa génétique bien connue, est particulièrement utile pour les études sur les voies dérégulées impliquées dans la longévité (Vellai et al, 2003 ; Shen et al, 2018). Cependant, les modèles invertébrés manquent d'organes et de systèmes essentiels au vieillissement humain comme le sang, un système de circulation fermé, des cellules souches somatiques dans plusieurs organes et un système immunitaire adaptatif (Brunet et al, 2020). Les modèles de **vertébrés traditionnels** comme les souris ou les poissons possèdent des organes et des systèmes similaires à ceux des humains. Les **souris**, avec leur cycle de vie plus court que l'Homme mais qui reste relativement long à appréhender, ont une génétique bien caractérisée, ce qui rend leur utilisation très fréquente pour les questions de l'âge biologique (Santin et al, 2020). Il est possible de les manipuler génétiquement, comme dans le cas des souris Progeria (Mutation du gène de la lamine A qui induit un phénotype de vieillissement accéléré) (Koblan et al, 2021) ou porteuses de diverses mutations (maladies mitochondriales induites par cassure de l'ADNmt par exemple), ce qui permet d'étudier les mécanismes du vieillissement. Les **rats**, ayant un cycle de vie plus long, sont préférés pour étudier des phénomènes de vieillissement à long terme, en particulier pour les maladies cardiovasculaires et les troubles métaboliques. Certaines études sont aussi menées sur les **zebrafish africains** (Pfefferli et al, 2015) qui sont des poissons dotés de capacité de régénération intéressante, et qui sont utiles dans les études sur la sarcopénie par exemple (faiblesse musculaire que l'on retrouve avec l'âge) (Daya et al, 2020). Bien que ces **modèles usuels** aient fourni des

informations importantes, leur éloignement de la physiologie humaine limite leur extrapolation chez l'Homme (par exemple les télomères des souris sont très longs comparés aux télomères chez l'Homme). De plus, la durée de vie d'environ 3 ans et les coûts d'élevage élevés pour les rongeurs sont un frein aux études sur le vieillissement. Les poissons zèbres n'ont pas été largement utilisés pour la recherche sur le vieillissement car ces poissons vivent relativement longtemps (de 3 ans à 5 ans) (Brunet et al, 2020).

De **nouveaux modèles** ont émergé comme par exemple le killifish turquoise africain **Nothobranchius furzeri** (Cellerino et al, 2016). Ce poisson présente une croissance, une maturation sexuelle et une production de descendance rapides (et donc un vieillissement rapide). De plus, il répond aux interventions qui peuvent retarder le vieillissement, comme le jeûne intermittent. Mais comme les modèles vertébrés traditionnels tels que la souris et le poisson zèbre, il présente des différences avec les humains en termes de longévité globale, de susceptibilité aux maladies, de comportement, et de réponses aux médicaments. Les **oiseaux** quant à eux, en raison de leur longévité et des similitudes avec les humains au niveau des systèmes nerveux, immunitaire, et des organes, offrent un autre modèle intéressant notamment dans les études sur la longévité et sur les mécanismes neuro-endocriniens (Travin et al, 2016). Les primates (singes rhésus ou ouistitis) (Mitchell et al, 2015) ou les chiens (Kaeberlein et al, 2016) permettent quant à eux d'étudier la variabilité individuelle et les comportements complexes liés au vieillissement, et de tester des interventions thérapeutiques contre le vieillissement. Mais chez ces modèles les études sont plus coûteuses et plus longues, ce qui les rend utiles pour la validation plutôt que pour la découverte de thérapies anti-vieillessement. Ces derniers soulèvent des questions **éthiques** car ces animaux conscients sont maintenus en captivité toute leur vie. De plus ces modèles, bien

qu'indispensables, ne sont pas suffisants pour une compréhension totale des mécanismes du vieillissement.

2.1.4.2.2 In vitro

Les modèles *in vitro* sont des systèmes de culture de cellules ou de tissus qui permettent de **simuler des aspects de la biologie** de l'organisme vivant dans un environnement contrôlé en laboratoire. Ces modèles présentent plusieurs avantages notamment en matière d'éthique, de contrôle (variables expérimentales), de **rapidité et de coût**. Ils permettent également de réaliser des études à grande échelle en cultivant un nombre important de cellules humaines avec une précision importante en fonction du choix de tissus ou d'organes.

Les modèles *in vitro* d'étude du vieillissement intègrent une **variété de types cellulaires** pour mieux représenter la complexité des processus biologiques liés au vieillissement. Les cellules primaires dérivées de tissus humains, telles que les fibroblastes, sont couramment employées dans l'exploration des altérations liées à l'âge. En effet ces cellules, lorsqu'elles sont maintenues en culture sur de longues périodes, acquièrent la plupart des caractéristiques établies du vieillissement (raccourcissement des télomères, dommages à l'ADN, altérations épigénétique, perte de la protéostase, dommages mitochondriaux, senescence cellulaire etc). Par exemple, les fibroblastes dermiques cultivés *in vitro* ou isolés *ex vivo* présentent des modifications de leurs fonctions (plasticité cellulaire, fonction du cytosquelette et gestion des pré-ARNm) qui ne sont actuellement pas considérées comme des caractéristiques du vieillissement mais qui sont d'un potentiel d'intérêt pour la découverte de biomarqueurs cellulaires (Tigges et al, 2014). À titre d'exemple, des recherches réalisées *in vitro* avec des fibroblastes humains ont mis en évidence une corrélation entre les propriétés mécaniques des cellules et l'âge chronologique des individus (Philips et al, 2017). D'autres études ont mis en

évidence des modifications dans la régulation de la sénescence cellulaire, un processus clé du vieillissement (Hartmann et al, 2023). Une corrélation a également été observée entre le profil bioénergétique des fibroblastes de peau, en particulier la chaîne respiratoire mitochondriale, et le déclin fonctionnel clinique chez les patients atteints de la maladie de Huntington, une affection neurodégénérative caractérisée par un déclin des capacités intrinsèques. (Gardiner et al, 2018).

Les **cellules souches pluripotentes induites** (iPSC) (Takahashi et al, 2006), sont également un outil intéressant car la technique permet la reprogrammation de cellules différenciées, comme des fibroblastes de peau, en cellules souches pluripotentes, capables de se différencier en divers types cellulaires du corps humain ce qui facilite ainsi l'étude de divers tissus (Stadtfeld et al, 2010). De plus, des **lignées cellulaires spécifiques**, telles que les cellules endothéliales, neuronales ou musculaires, sont employées pour modéliser les changements spécifiques associés au vieillissement dans ces types cellulaires particuliers. Ces modèles *in vitro*, en combinant différentes lignées cellulaires, offrent une approche **complémentaire aux modèles animaux** pour comprendre les mécanismes sous-jacents du vieillissement et contribuent à l'identification de cibles thérapeutiques potentielles. Tous ces modèles d'études permettent de générer de grandes quantités de données qu'il faut ensuite intégrer ensemble.

2.1.4.2.3 Les pathologies humaines de vieillissement accéléré

Il existe plusieurs types de vieillissement, un vieillissement **physiologique** qui apparaît avec l'avancée en âge et un vieillissement **prématuré** en lien avec certaines pathologies. Le **syndrome de Hutchinson-Gilford** (Progeria), conséquence d'une mutation à l'origine de la production de la protéine lamine A défectueuse, perturbant l'architecture nucléaire et entraînant un vieillissement accéléré, est fréquemment utilisé dans des études, que ce soit in

vitro (Philips et al, 2017) ou in vivo (Koblan, et 2021). Cette pathologie, caractérisée par une espérance de vie réduite à 12 ans, permet une comparaison directe avec les altérations observées dans les modèles de vieillissement physiologique, validant les caractéristiques de vieillissement identifiées. La **sclérose latérale amyotrophique** (SLA), une maladie neurodégénérative ayant des facteurs génétiques, représente également un modèle de vieillissement en raison de ses impacts durables et irréversibles sur les capacités intrinsèques (Gerou et al, 2021). D'autres maladies génétiques comme les **maladies mitochondriales**, peuvent causer des symptômes variés tels que des troubles musculaires, des troubles neurodégénératifs, des troubles de la croissance, du fait de l'importance de la mitochondrie dans l'homéostasie énergétique et le stress oxydant (Grillo et al, 2021). Cette liste, non exhaustive, de modèles dérivés de pathologies humaines démontre qu'ils sont considérés comme des modèles de vieillissement accélérés en raison de leurs répercussions sur les fonctions cellulaires majeures ou piliers du vieillissement qui impactent directement les capacités intrinsèques.

2.2 Les stratégies thérapeutiques

2.2.1 Stratégies actuelles : non médicamenteuse

2.2.1.1 *La nutrition*

La **restriction calorique**, caractérisée par une réduction calorique de 20 % à 40 % chez les mammifères supérieurs a démontré, dans plusieurs études, la capacité de diminuer l'expression des gènes associés au stress oxydatif, tout en améliorant les dommages oxydatifs dans divers tissus. Cette intervention a également été associée à une réduction des risques cardiovasculaire, une amélioration de la sensibilité à l'insuline, une optimisation de la fonction mitochondriale, ainsi qu'une diminution des dommages oxydatifs à l'ADN (Golbidi et al, 2017).

Par conséquent, dans différents modèles animaux, du ver aux primates non humains la restriction calorique allonge la durée de vie en bonne santé en atténuant de nombreuses conséquences pathologiques du vieillissement (Colman et al, 2009). Un mécanisme pouvant être impliqué est la voie de **l'autophagie**, processus cellulaire qui implique la dégradation et le recyclage des composants cellulaires. L'autophagie joue un rôle crucial dans le maintien de l'homéostasie cellulaire en éliminant les composants dysfonctionnels et en recyclant leurs éléments constitutifs pour soutenir les fonctions cellulaires (Bagherniya et al, 2018). Il s'agit d'un mécanisme fondamental de contrôle de la qualité cellulaire, d'adaptation au stress et d'équilibre énergétique. Une dérégulation de l'autophagie a été impliquée dans diverses maladies, notamment les troubles neurodégénératifs, le cancer et les affections métaboliques (Escobar et al 2019). La restriction calorique semble agir via l'activation de la voie AMPK et Sirtuine associé à l'inhibition de la voie mTOR (Figure 9). L'AMPK, activée en cas de faible disponibilité d'énergie, régule le métabolisme cellulaire en stimulant des processus énergétiques tout en inhibant des voies coûteuses en énergie. Les sirtuines, famille d'enzymes dépendantes du NAD⁺, régulent divers processus cellulaires, de la modulation de l'activité protéique au maintien de l'intégrité de l'ADN, et sont associées à la longévité, reflétant leur rôle central dans le bon fonctionnement cellulaire. Enfin, la voie mTOR, sensible aux signaux environnementaux tels que les nutriments et les facteurs de croissance, orchestre la croissance cellulaire, la prolifération et le métabolisme et est associé aux cancers.

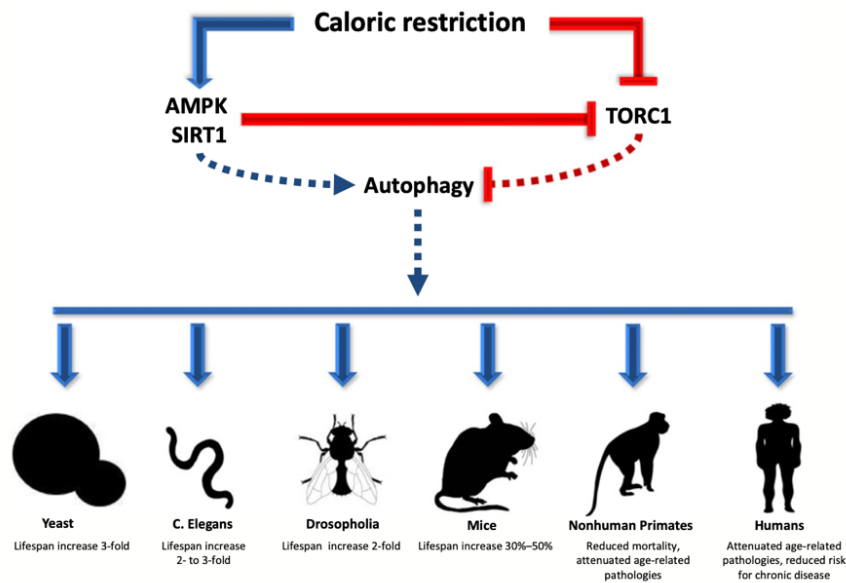


Figure 9 : Impact de la restriction calorique chez plusieurs espèces

Escobar et al 2019

Le **jeûne intermittent** a également été associé à l'activation de voies similaires à celles de la restriction calorique (Anton et al, 2013). Dans le cadre de l'**étude NU-AGE**, axée sur l'évaluation des effets d'un régime méditerranéen sur la santé des personnes âgées, le groupe suivant ce régime a présenté une réduction significative de la fragilité, une amélioration de la fonction physique et une meilleure cognition par rapport au groupe témoin (Ghosh et al, 2020).

De plus, une **alimentation équilibrée** riche en nutriments (fruits, légumes), en protéines pauvre en graisse ou en aliments à base d'acide gras insaturés trans industriels, a été associée à un risque plus faible de maladies chroniques et à une meilleure santé globale chez les personnes âgées.

En conclusion les études récentes montrent que le facteur alimentation est essentiel à prendre en compte puisque certains facteurs se sont révélés bénéfiques pour un vieillissement en bonne santé comme l'adhésion à un régime de type méditerranéen, au

régime d'Okinawa, au régime DASH (Dietary Approaches to Stop Hypertension) ou encore à la restriction calorique (Bojang et al, 2023).

2.2.1.2 L'activité physique

L'OMS recommande au moins **150 minutes d'activité physique modérée par semaine**, soulignant les bienfaits évidents de l'exercice sur la santé (OMS). L'exercice d'endurance chronique favorise notamment un vieillissement sain en atténuant l'apparition des maladies liées à l'âge (Arem et al, 2015). Des données longitudinales indiquent une réduction d'environ 30 % du risque de décès chez les individus physiquement actifs (Schnohr et al, 2015). Les mécanismes mis en jeu semblent être médiés par une protéine, l'AMPK, qui possède une fonction d'enzyme. Cette protéine est sensible au « stock » d'énergie (ATP) dans la cellule. Plus le taux d'ATP est bas, plus l'AMPK augmente. L'activation de cette protéine permet entre autres d'augmenter l'absorption du glucose dans les muscles, la synthèse de mitochondries, et l'autophagie (Escobar et al, 2018) dans différents organes (Figure 10).

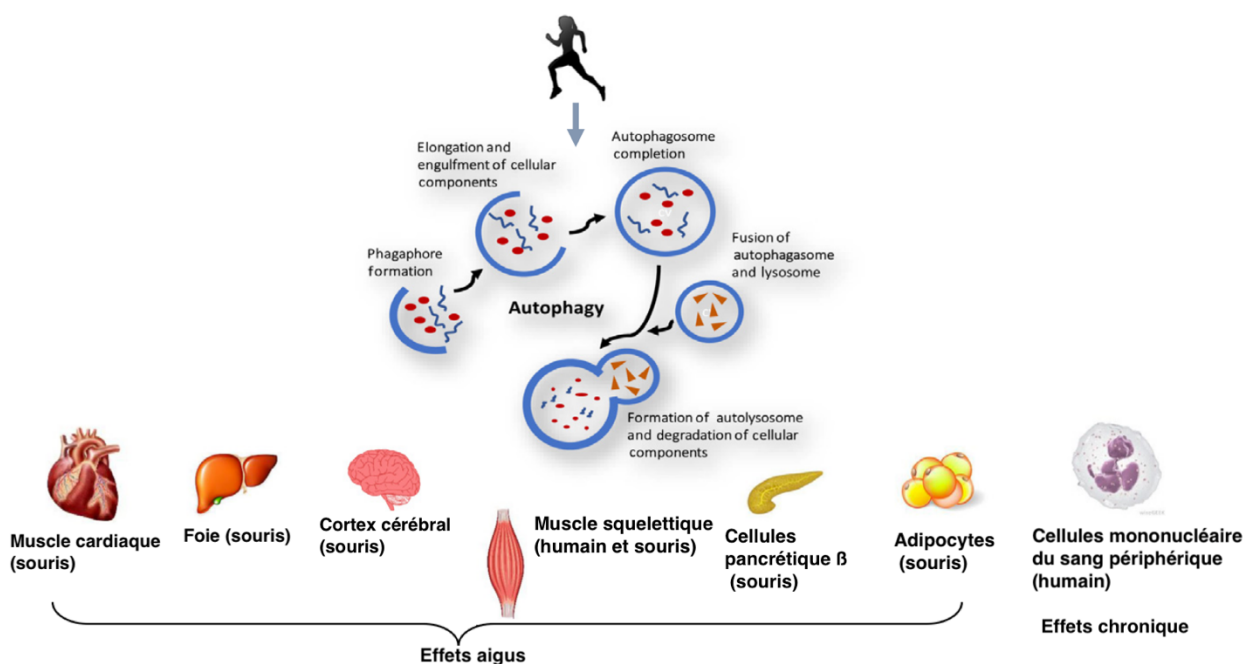


Figure 10 : Influence de l'exercice physique sur l'autophagie et impact multi-organes
Escobar et al, 2018

Le mode de vie des individus joue donc un rôle essentiel et non négligeable dans la promotion du vieillissement en bonne santé. À l'avenir, il pourrait être envisagé d'associer ces approches avec des interventions médicamenteuses afin d'agir de manière synergique avec celles-ci.

2.2.2 Stratégies en développement : médicamenteuses

2.2.2.1 *La sénescence*

Pour rappel, la sénescence (vu dans le paragraphe : Senescence, des Piliers du vieillissement) qui correspond à un arrêt stable du cycle cellulaire est associée à la sécrétion de facteurs pro-inflammatoires, de facteurs de croissance et de protéases regroupés dans le phénotype sécrétoire associé à la sénescence **ou SASP** (Li et al, 2023). Chez le sujet âgé, on retrouve la présence chronique de cellules sénescents dans divers tissus tel que le tissu cartilagineux, osseux, musculaire ou encore au niveau du tissu adipeux, de l'intima artérielle, du rein ou du poumon. Il a été montré qu'elles étaient associées aux développements de pathologies comme l'arthrose, l'ostéoporose, la sarcopénie, la lipodystrophie du sujet âgé, l'athérosclérose, ou des insuffisances rénales et respiratoires (Schmeer et al, 2019). De plus, des études ont montrés que la **transplantation** de cellules sénescents pouvait contribuer (chez la souris) au développement de pathologies comme l'arthrose (Xu et al, 2017) ou pouvait limiter la régénération cardiaque post-infarctus du myocarde (Lewis-McDougall et al 2019), notamment par la production de facteurs pro-inflammatoires. De ce fait, les cellules sénescents semblent être des **cibles thérapeutiques d'intérêt**, pour lesquels des fenêtres thérapeutiques contrôlées doivent être définies afin de maintenir leurs effets bénéfiques tout en **évitant leur accumulation**.

2.2.2.2 Les senothérapeutiques

Les nouvelles thérapies ciblant les cellules sénescentes sont regroupées sous le terme de **Senothérapeutiques**. Ces dernières ont différents objectifs : les **sénolytiques** qui tuent sélectivement les cellules sénescentes, les **sénomorphiques** qui modulent leurs fonctions et leur morphologie et enfin les **sénostatiques** qui régulent leur accumulation. Cependant, le classement est controversé en raison de la complexité des voies d'action des molécules, compliquant leur spécificité d'action ainsi que de la variabilité de phénotype des cellules sénescentes. De plus, le fait que le vieillissement ne soit pas considéré comme une maladie, complique le développement de nouveaux médicaments, c'est-à-dire avec une 1^{ière} AMM. Les essais cliniques actuels se concentrent donc sur le repositionnement d'AMM de médicaments déjà autorisés. Afin de traiter des individus fragiles (pour inverser leur phénotype) ou en bonne santé (en prévention), il est crucial de cibler une fenêtre thérapeutique étroite en raison de potentiels effets indésirables peu acceptables. Nous verrons ici un exemple non exhaustif de molécules candidates intéressantes. Plusieurs molécules candidates, telles que la metformine, ont initialement une AMM **contre le diabète de type 2** dans lequel on observe une dérégulation du métabolisme qui fait partie des piliers du vieillissement. Des études sur des invertébrés et des souris ont révélé une prolongation de la durée de vie et de la vie en bonne santé sous l'effet de la metformine (Glossmann et al, 2019) notamment par l'induction d'un état similaire à celui de la restriction calorique (Onken et al, 2010). Les mécanismes d'action de la metformine ne sont pas entièrement compris, mais des indications suggèrent qu'elle provoquerait une inhibition du complexe 1 de la chaîne mitochondriale (Wheaton et al, 2014), réduisant la charge énergétique cellulaire (augmentation du rapport AMP/ATP). De ce fait, l'AMPK serait activé ce qui entrainerait notamment une diminution du taux de glucose (Foretz et al, 2010 ; Hunter et al, 2018) (Figure 11). Il est intéressant de constater que les voies

mises en jeu avec la metformine sont similaires à celles induites par la restriction calorique ou l'exercice physique (Cf partie exercice physique et restriction calorique).

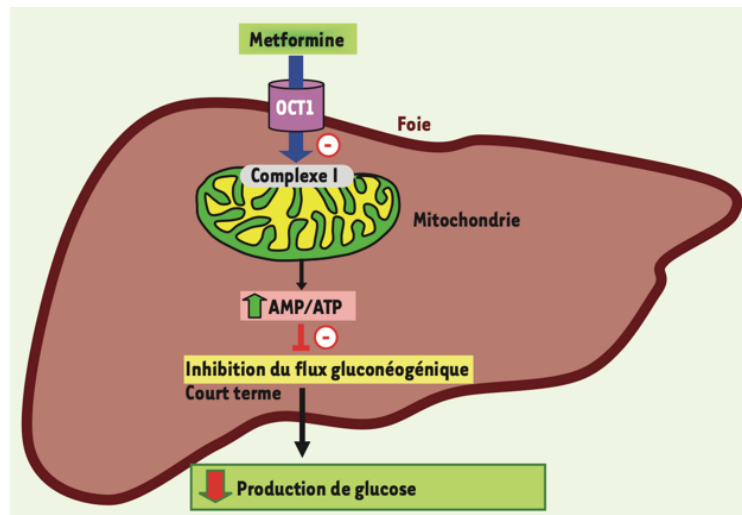


Figure 11 : Proposition de mécanisme de la metformine au niveau du foie chez l'homme
Foretz et al, 2010

Des données pré-cliniques et cliniques suggèrent que la metformine, via la suppression des voies pro-inflammatoires, la protection des mitochondries et de la fonction vasculaire et des actions directes sur les cellules souches neuronales, peut protéger contre les maladies neurodégénératives (Triggle et al, 2022). **L'étude TAME** (Targeting Aging with METformin) (Barzilai et al, 2017), est le premier essai clinique approuvé par la FDA, chez l'homme avec pour indication une molécule contre les effets du vieillissement. Cette étude américaine regroupe une série d'essais cliniques multicentriques en double aveugle. Elle se déroulera sur 6 ans dans 14 instituts de recherche qui ont recrutés plus de 3 000 personnes âgées de 65 à 79 ans. L'étude qui demande des fonds importants devrait démarrer dans les prochaines années. L'étude **MILES** (Metformine dans l'étude de la longévité) quant à elle beaucoup plus petite, a réuni 16 participants de plus de 60 ans. Les résultats de cette étude ne sont pas encore disponibles. Cette molécule est donc une candidate de repositionnement d'AMM très

intéressante, de par son faible coût mais aussi sa sécurité bien établie avec des effets indésirables rares (acidose lactique) ou fréquents (problèmes gastro-intestinaux) mais peu contraignants (Chen et al, 2022).

Les **inhibiteurs de la dipeptidyl peptidase-4 (DPP-4)**, également connus sous le nom de Gliptines, constituent une autre classe de médicaments couramment utilisés pour traiter le diabète de type 2. Des essais en cours suggèrent que cette classe de molécules pourrait présenter un intérêt dans le contexte de la sarcopénie (Witham et al, 2023) dont le facteur de risque principal est l'avancée en âge. La sarcopénie se caractérise par une perte de force et de masse musculaires ce qui provoque des difficultés pour la fonction de locomotion liée au phénotype de fragilité (Cruz-Jentoft et al, 2019).

Un autre exemple est l'**aspirine**, ou acide acétylsalicylique, un médicament largement prescrit aux multiples indications telles qu'antalgique, antipyrétique, anti-inflammatoire à forte dose, et antiagrégant plaquettaire. Son mécanisme d'action, bien compris, repose sur l'inhibition des prostaglandines par son effet sur les cyclo-oxygénases (COX) 1 de manière irréversible et COX 2 de manière réversible. Plusieurs études ont montré que l'utilisation de cette molécule était corrélée à la diminution de l'incidence de plusieurs cancers (Chan et al, 2007 ; Rothwell et al, 2011) ce qui a motivé les chercheurs à la considérer candidat de repositionnement d'AMM contre le vieillissement. Cependant, une étude chez des personnes âgées en bonne santé a révélé que l'utilisation d'aspirine n'a pas prolongé la survie sans incapacité sur cinq ans, et a en plus entraîné un taux **d'hémorragie majeure** plus élevé que le placebo (McNeil et al, 2018). La balance bénéfice n'a pas permis de garder ce candidat médicament.

Une association de molécules, le **Dasatinib + la quercétine**, est également à l'étude (Hickson et al, 2019). Le Dasatinib, inhibiteur de tyrosines kinase utilisé en tant qu'anticancéreux, combiné à la quercétine, un flavonoïde aux propriétés antioxydantes antagonistes de la voie

des Phosphoinositides 3-Kinases (PI3K), montrent une **synergie** d'action induisant la mort des cellules endothéliales et des pré-adipocytes sénescents par inhibition des voies de survie. Les données d'un essai clinique de phase 1 sur des patients atteints de la maladie rénale du diabétique indiquent que l'association dasatinib + quercétine réduit la charge de cellules sénescents du tissu adipeux, diminue les macrophages activés par les cellules sénescence et diminue les facteurs du SASP circulant tel que IL-1, IL-6 et des métalloprotéase (Hickson et al, 2019).

D'autres molécules sont également candidates comme les inhibiteurs des molécules de la famille Bcl-2 lié à la résistance des cellules à l'apoptose ou encore la Fisétine, un flavonoïde retrouvé dans les fruits et les légumes qui a des propriétés anti-oxydante, anti-inflammatoires etc (Zang et al, 2023).

Globalement, de nombreuses molécules semblent avoir une influence sur les cellules sénescents et être des potentiels candidat-médicaments contre les effets délétères du vieillissement (Figure 12). L'acceptation de ces thérapeutiques repose sur le contrôle des effets indésirables car bien que minimes ils pourraient ne pas être acceptable dans le cadre d'un traitement préventif chez un individu jeune ou âgé à priori en bonne santé. Il est donc important d'utiliser de nouveaux outils pour développer et gérer au mieux l'administration de ces nouvelles thérapeutiques.

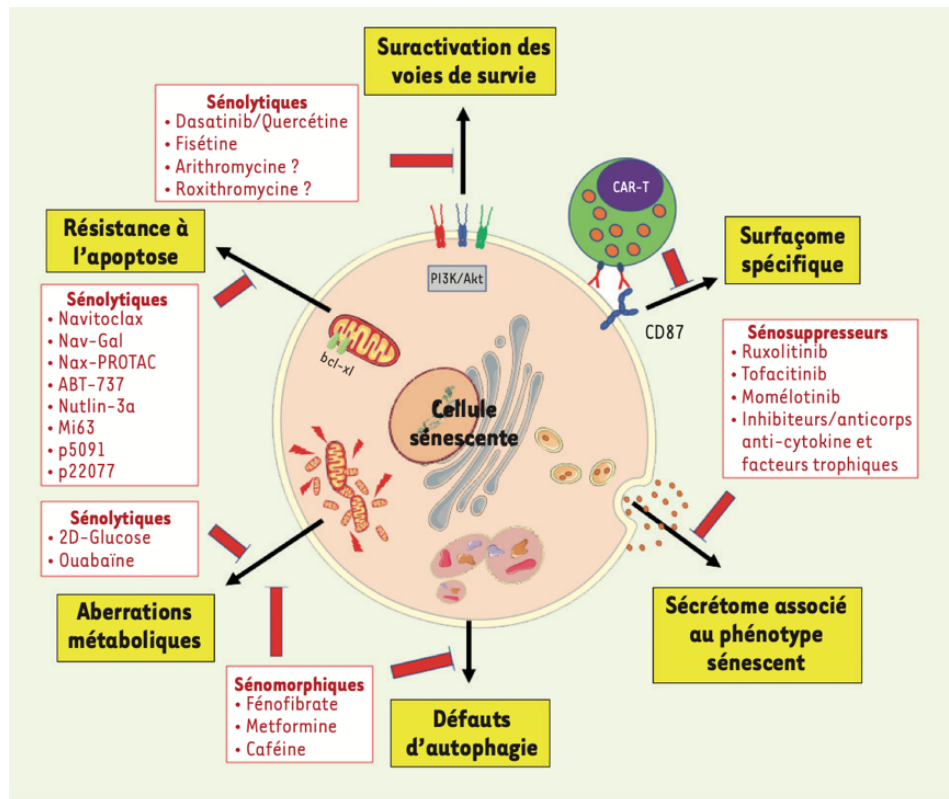


Figure 12 : Liste non exhaustive de Senothérapeutiques
Foretz et al, 2010

3 Le machine learning

3.1 Généralités

3.1.1 Une nouvelle ère : la recherche translationnelle

Nous vivons dans une ère de recherche translationnelle qui se caractérise par la **collaboration** étroite entre les chercheurs au laboratoire et les cliniciens au lit des patients. L'objectif est de faciliter la translation des découvertes scientifiques vers l'amélioration de la santé des individus. Les **progrès** technologiques notamment informatiques et l'**accès** contrôlé aux données médicales permettent une accélération de la compréhension des mécanismes des maladies, favorisant ainsi le développement de traitements plus personnalisés. Cette approche **multidisciplinaire** (clinicien, scientifique, bioinformaticien) qui nécessite l'expertise de chacun, favorise l'innovation et accélère le passage des découvertes en recherche

fondamentale vers des stratégies concrètes pour améliorer la prise en charge des patients (Figure 13). Dans l'objectif du vieillissement en bonne santé, grâce à la recherche fondamentale, les processus biologiques inhérents et les mécanismes physiopathologiques liées à l'âge sont mieux appréhendés grâce au développement de modèles *in vitro* et *in vivo*. Les **bioinformaticiens** gèrent les nombreuses bases de données disponibles (GenAge, GenDR, DrugAge, the longevity map, CellAge etc) ou nouvellement générées et permettent des liens entre les découvertes scientifiques et les retentissements cliniques par exemple. Enfin, les **cliniciens** identifient les patients qui bénéficieront d'interventions préventives et rentreront dans les essais cliniques (Kruempel et al, 2019 ; Lindsell et al, 2021 ; Moaddel et al, 2021).

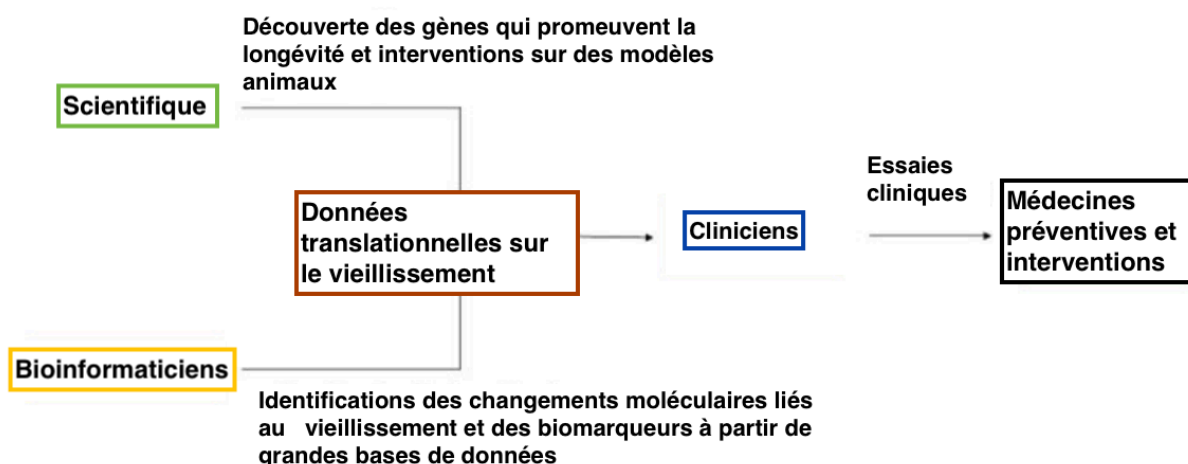


Figure 13 : Les différents acteurs de la recherche translationnelle
Kruempel et al, 2019

3.1.2 Intérêt par rapport aux statistiques

Les **statistiques** jouent un rôle essentiel dans la recherche en permettant l'analyse **quantitative** de données, et plus particulièrement dans le cadre de cette thèse, des données liées aux processus biologiques, aux facteurs environnementaux et aux caractéristiques génétiques associés au vieillissement. Elles fournissent des outils pour évaluer les **tendances**,

identifier les **corrélations** et établir des associations significatives entre divers paramètres. Cependant, les statistiques traditionnelles sont peu adaptées aux masses de données disponibles aujourd'hui. De plus les statistiques sont une méthode très stratifiée de par la multiplicité et la spécificité des tests qui augmente la probabilité d'obtenir une significativité uniquement par hasard. Un des biais concerne la sélection du seuil de p-value dont la valeur est arbitrairement définie par la communauté scientifique et qui conditionne la significativité et donc l'interprétation des résultats. Le **machine learning complète** ces approches statistiques traditionnelles en exploitant la capacité des algorithmes à détecter des motifs complexes et non linéaires dans de grandes quantités de données, offrant ainsi une perspective plus exhaustive sur la compréhension des mécanismes physiologiques du vieillissement que nous détaillerons dans ce chapitre. Cette **complémentarité** permet d'explorer les relations et les interactions multidimensionnelles entre les données, ce qui améliore la capacité à prédire les phénomènes liés au vieillissement et pourra à terme permettre de développer des interventions personnalisées.

3.1.3 Découverte et intérêt de l'IA en santé et en sciences

L'intelligence artificielle (IA), en pleine expansion depuis **les années 50**, est une discipline issue de la science des données et se définit par l'ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de **l'intelligence humaine** (raisonnement, apprentissage...) (Choi et al, 2020) (Figure 14).

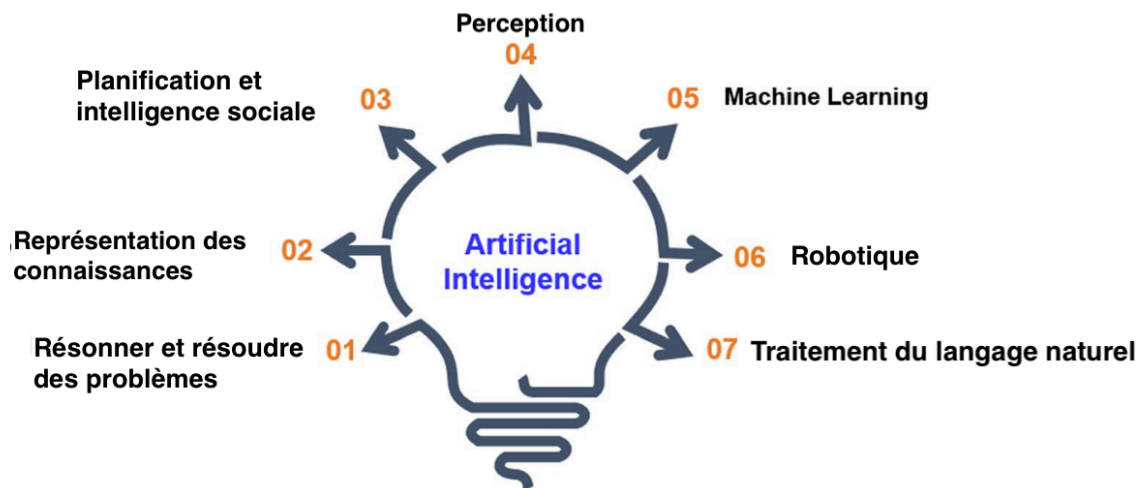


Figure 14 : Classification de l'IA
Gupta et al, 2021

L'intelligence artificielle a de nombreuses applications en santé et en sciences (Figure 15).

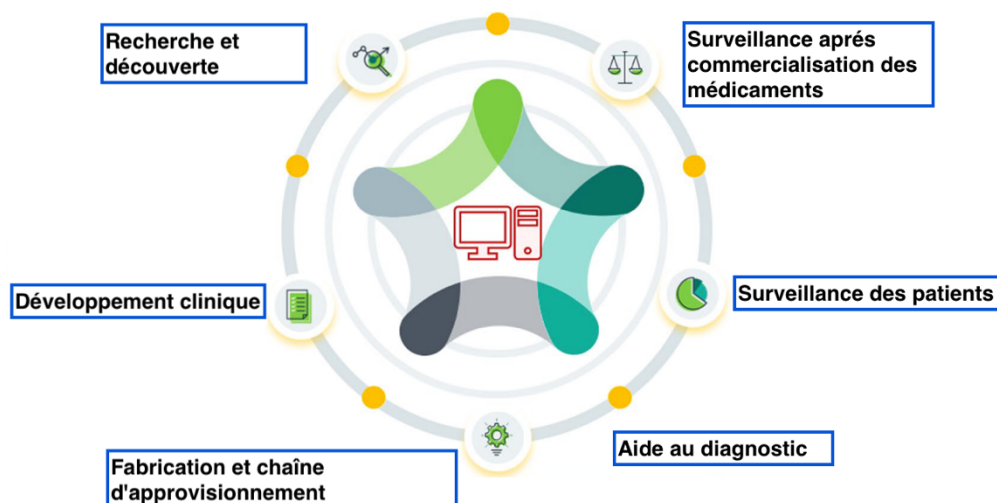


Figure 15 : IA dans la santé et l'industrie pharmaceutique
Gupta et al, 2021

L'objectif initial de l'IA a été de concevoir des machines capables de raisonner comme l'Homme, voire de le **remplacer** par ce qu'on appelait des « **systèmes experts** » comme par exemple le système mycin (Shortliffe et al, 1975) pour détecter les infections bactériennes ou le système sphinx (Fieschi, 1983) pour détecter les ictères afin de faciliter le diagnostic. Aujourd'hui, l'objectif dans le domaine de la santé est de concevoir des machines pouvant

assister l'expert dans des tâches complexes et redondantes comme par exemple aider aux dépistages des rétinopathies diabétiques via un fond d'œil « intelligent » (Yang et al, 2019) ou encore de **personnaliser les soins chez les personnes âgées** en évaluant les risques et en proposant des recommandations adaptées. En biologie, la modélisation permet de **traduire une question biologique** (observation) en un modèle mathématique théorique. La complexité des données (cliniques, omics, expérimentales) à manipuler a conduit à l'utilisation exponentielle de l'apprentissage automatique en laboratoire pour construire des modèles **informatifs et prédictifs** des processus biologiques connus ou inconnus (Mathis et al, 2018). Les sujets de prédictions sont multiples, recouvrant différents champs d'applications, comme l'évaluation de la fonction des protéines (Antczak et al, 2019) ou l'évaluation de la toxicité des médicaments (Mayr et al, 2016). Concernant les applications du machine learning dans la recherche sur le vieillissement, on retrouve par exemple la prédiction des gènes lié à l'âge (Li et al, 2022), l'estimation de l'âge biologique (Bortz et al, 2023 ; Naue et al, 2017) ou la prédiction de la trajectoire de santé (Rutledge et al, 2022 et al, 2022) que nous détaillerons par la suite. L'adoption du machine learning dans le domaine de la recherche sur le vieillissement a connu une croissance exponentielle. Entre 2000 et 2010, environ 20 000 articles ont été publiés sur le thème de l'intégration du machine learning dans la recherche sur le vieillissement. Cette tendance s'est encore accentuée entre 2010 et 2024, avec un nombre significativement plus élevé de publications, atteignant 54 300 articles sur la même période (données : google scholar).

3.2 Définition et principes de base du machine learning

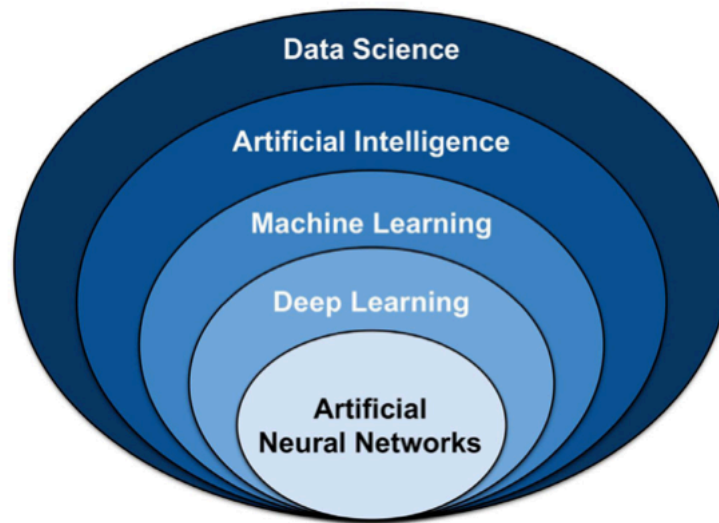


Figure 16 : Place du machine learning dans la science des données
Choi et al, 2020

Le machine learning est un **sous-ensemble** (Figure 16) de l'intelligence artificielle qui permet à un système informatique d'apprendre et de s'améliorer à partir de données, sans être explicitement programmé. Il repose sur le principe de base selon lequel les machines sont capables d'analyser des modèles et des structures dans les données pour prendre des **décisions ou effectuer des prédictions**. Le principe de base du machine learning est **d'entraîner** un modèle en lui fournissant des données d'entraînement. Le modèle tente ensuite de trouver des schémas ou des relations dans ces données afin de **généraliser** et de faire des prédictions sur de nouvelles données non vues auparavant. Il existe différents types d'apprentissage (que nous détaillerons en suivant) en fonction du type de données et du problème scientifique à résoudre. Une fois le modèle entraîné, il peut être utilisé pour faire des prédictions ou prendre des décisions sur de **nouvelles données**. Le modèle est évalué en comparant ses prédictions avec les résultats réels, et s'il ne parvient pas à obtenir des résultats satisfaisants, il est ajusté et entraîné à nouveau pour améliorer ses **performances** (Choi et al,

2020). Le choix de la métrique d'évaluation des performances du modèle dépend de la problématique. La **précision** (accuracy) mesure la proportion de prédictions correctes (vrai positif) parmi l'ensemble des prédictions. Le **rappel** (recall) évalue la proportion de vrais positifs parmi tous les exemples positifs réels, élément crucial pour détecter les faux négatifs. Par exemple, dans la prédiction de diagnostic de cancer il est très important de limiter les faux négatifs pour éviter un retard de diagnostic du patient. La **spécificité** (specificity) quantifie la proportion de vrais négatifs parmi tous les exemples négatifs réels, importante pour minimiser les faux positifs. Par exemple, dans la prédiction de découverte de nouveaux candidats médicament il faut éviter les faux positifs qui augmentent le nombre de molécule à étudier et augmentent le coût expérimental (Smer-Barretoo et al, 2022). Un outil souvent utilisé est le F1-score, qui se calcule à partir de la précision et du rappel, qui permet de limiter les faux positifs et les faux négatifs ce qui est essentiel avec des applications médicales. En général, un F1-score élevé indique que le modèle est capable de classier efficacement les données, avec un bon équilibre entre la précision et le rappel.

Le machine learning est lui-même composé de techniques spécifiques que nous allons maintenant détailler.

3.2.1 Différents types d'apprentissage

3.2.1.1 *L'apprentissage supervisé*

L'apprentissage supervisé nécessite des données **étiquetées** c'est-à-dire qu'on a une connaissance au préalable de ce qu'on attend comme résultats ou « sortie » pour nos échantillons (Baştanlar et al, 2014). Par exemple, à partir de données de patients (cliniques, biologiques) nous voulons prédire si l'individu est robuste, pré-fragile ou fragile. Les données d'entraînement sont donc étiquetées puis le modèle est entraîné avec ces données. Dans un

second temps, des données tests sont intégrées dans le modèle, et une prédiction pour chaque donnée est proposée (Figure 17).

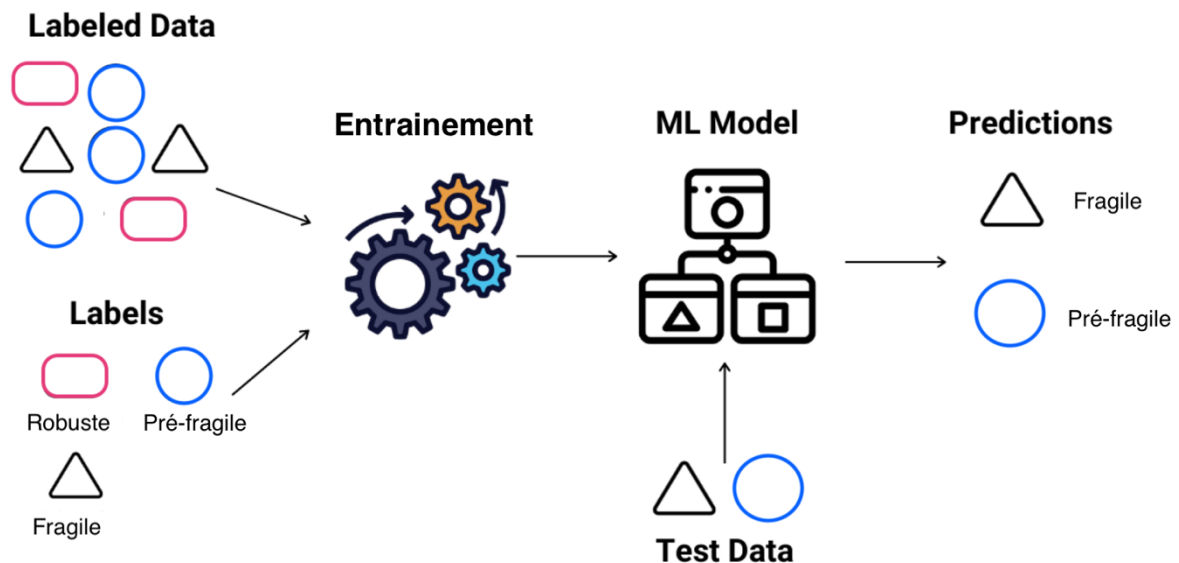


Figure 17 : Pipeline technique de machine learning supervisé

3.2.1.2 L'apprentissage non supervisé

L'apprentissage non supervisé recherche des structures ou des modèles dans les données **sans étiquettes, c'est ce qu'on appelle le « clustering »**. Les algorithmes vont détecter des caractéristiques communes entre les données et les classer sans à priori dans des groupes appelés « clusters » en fonction de leurs similitudes (Baştanlar et al, 2014). Par exemple, l'apprentissage non supervisé pourrait être employé afin d'identifier sans à priori des sous-groupes distincts de personnes âgées en fonction de leurs profils métaboliques. Par exemple, des données sur le métabolisme cellulaire, provenant de techniques omiques telles que la Spectrométrie de Masse ou la Résonance Magnétique Nucléaire, peuvent être recueillies sur un grand nombre de personnes d'âge différent. Ceci permet de disposer d'un nombre important de données métaboliques d'entrée à insérer dans le modèle. L'apprentissage non supervisé peut alors regrouper les individus en clusters distincts en fonction de similarités

métaboliques, sans être guidé par des étiquettes préexistantes (Figure 18). Ce processus pourrait révéler des profils métaboliques spécifiques associés à des sous-types de vieillissement.

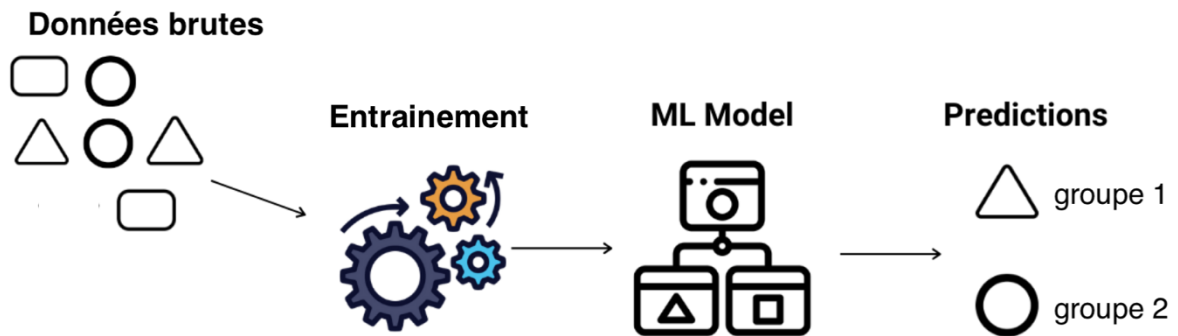


Figure 18 : Pipeline technique de machine learning non supervisé

3.2.1.3 L'apprentissage par renforcement

L'apprentissage par renforcement est une **approche itérative d'optimisation algorithmique** basée sur les actions entreprises par un agent et les réponses de l'environnement. Ce type d'apprentissage, développé au cours de années 50, est **partagé par tous les êtres vivants** car c'est un apprentissage par l'interaction avec l'environnement (Sutton et al, 2018). Cette méthode autorise les machines et les agents à ajuster automatiquement leur comportement dans un contexte donné afin de **maximiser** leurs performances vis-à-vis d'un objectif donné. Elle repose sur une rétroaction de récompense, appelée signal de renforcement, essentielle pour que l'agent apprenne quelles actions sont les plus bénéfiques. Par exemple, l'agent pourrait tester différents types d'action ou dans notre cas, de programmes d'exercices physiques, dans le but d'améliorer la fonction locomotrice des individus. L'environnement qui représente les états de la fonction locomotrice sera évalué avec un score attribué selon le

nouvel état de la fonction locomotrice après l'action. Ce score sera une récompense renvoyée à l'agent, le but étant pour ce dernier de maximiser sa récompense. (Figure 19).

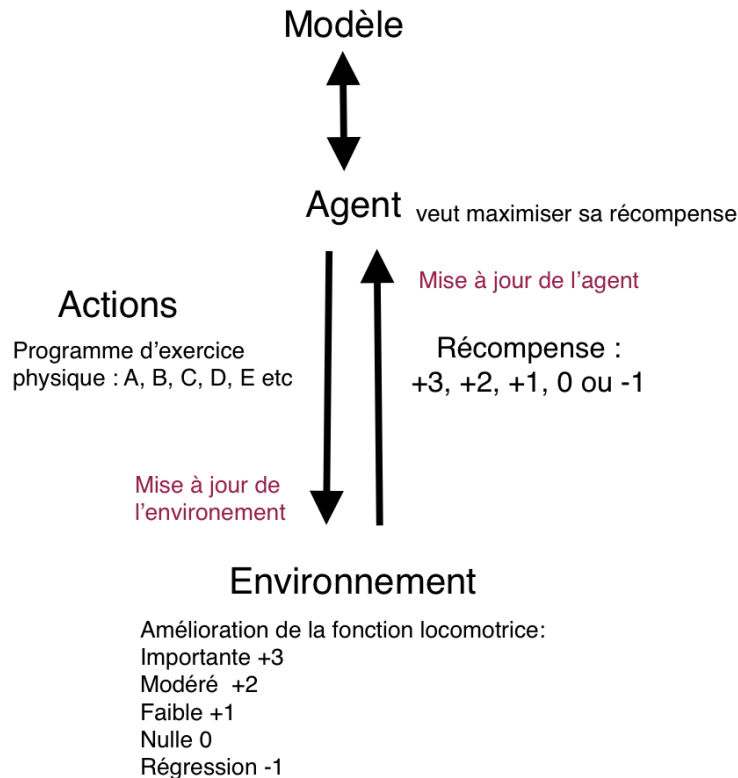


Figure 19 : Pipeline technique de machine learning par renforcement

3.2.2 La création des modèles : Différents algorithmes utilisés

La phase suivante de ce processus implique **l'entraînement** ou l'apprentissage du modèle, nécessitant la création d'un modèle à partir d'un algorithme ou d'un ensemble d'algorithmes confrontés aux données. Il existe **divers types d'algorithmes**. La sélection de l'algorithme en machine learning est guidée par la nature du problème à résoudre, les caractéristiques des données, ainsi que les contraintes propres à l'application (besoin d'interprétabilité, environnement avec des données limitées en taille et en diversité, prédictions en temps réel). Différents algorithmes suscitent un intérêt particulier dans la recherche sur le vieillissement

en bonne santé. Nous examinerons divers algorithmes, en mettant l'accent sur ceux de régression, qui sont utilisés pour prédire une valeur numérique et qui s'avèrent pertinents pour déterminer l'âge biologique. Il est important de noter l'existence d'algorithmes de classification visant à prédire des catégories spécifiques ou classes, telles que le statut de fragilité. Globalement, la première étape du pipeline est **l'échantillonnage** de population qui consiste à sélectionner un sous-groupe représentatif d'individus dans une population générale plus vaste. L'ensemble de données est ensuite divisé en un groupe d'entraînement (80% des données) et un groupe de test (20% des données). Le modèle est ensuite entraîné. La **validation croisée** est une technique essentielle où les données sont divisées en k sous-ensembles pour entraîner et évaluer le modèle successivement, permettant ainsi une estimation robuste de sa performance. Elle est essentielle pour éviter les biais et le sur-apprentissage. Il est également recommandé de **tester plusieurs algorithmes** pour comparer les performances, à travers des variables comme le F1-score, la précision (définie en début du chapitre), l'écart-type des modèles etc. Cette approche permet d'évaluer la robustesse et la fiabilité des modèles dans des conditions variées et de choisir celui qui s'adaptera le mieux à la situation. Dans la littérature, il est courant de tester au moins cinq modèles différents pour une même question biologique. Cette diversité d'approches permet d'explorer différentes stratégies de modélisation et de sélectionner celle qui offre les meilleures performances et la meilleure généralisation aux données réelles. La comparaison des performances entre plusieurs modèles permet également de mieux comprendre les forces et les limites de chaque algorithme dans le contexte spécifique de la question biologique étudiée, ce qui est essentiel pour garantir la qualité et la pertinence des résultats obtenus (Figure 20).

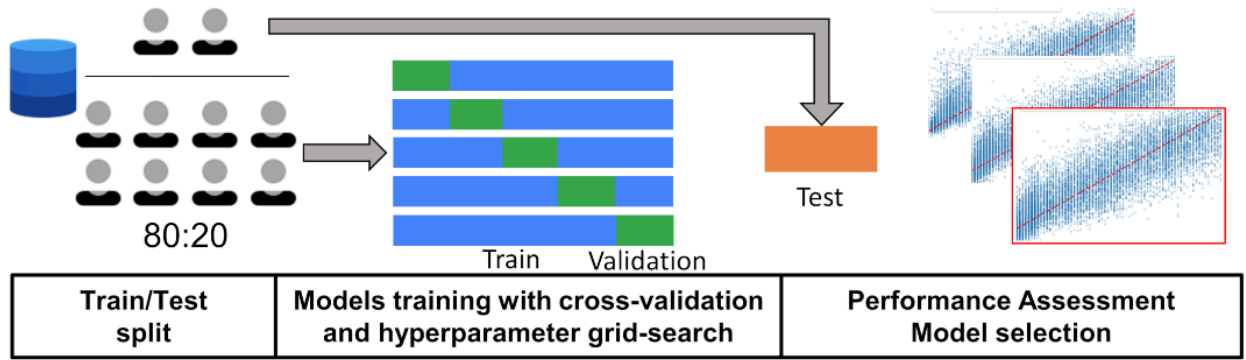


Figure 20 : Exemple de pipelines de machine learning
Bernard et al, 2023

3.2.2.1 Algorithmes pour l'apprentissage supervisé

Nous allons à présent examiner les principaux algorithmes fréquemment cités dans la littérature dédiée au vieillissement. Leur présentation technique sera complétée par les résultats pouvant être obtenus lors de leur application.

3.2.2.1.1 La régression linéaire simple ou multiple

L'algorithme de régression linéaire explore **l'influence** d'une ou de plusieurs variables sur une variable de résultat, attribuant des **poinds** (ou importance) à chacune. Il y a 2 types de régressions linéaires. La régression linéaire **simple** est utilisée pour estimer les paramètres d'une droite reliant la variable réponse à la variable prédictive. En outre, elle permet d'évaluer la significativité de cette relation. La régression linéaire simple peut par exemple définir si une caractéristique **simple** comme un biomarqueur, est liée à l'âge (Figure 21). La régression linéaire **multiple** est une méthode qui permet de déterminer les relations de corrélation entre un résultat (la variable expliquée) et plusieurs variables explicatives et indépendantes. Elle peut définir si plusieurs caractéristiques ou plusieurs biomarqueurs sont liées à l'âge. Pour être plus précis, une régression linéaire simple consiste à minimiser la disparité entre une fonction linéaire et un ensemble de données en déterminant les deux coefficients qui la

définissent ($f(x)=ax+b$). La méthode des moindres carrés, couramment utilisée, cherche à réduire la somme des carrés des écarts entre les valeurs prédites par la fonction linéaire et les valeurs observées dans les données.

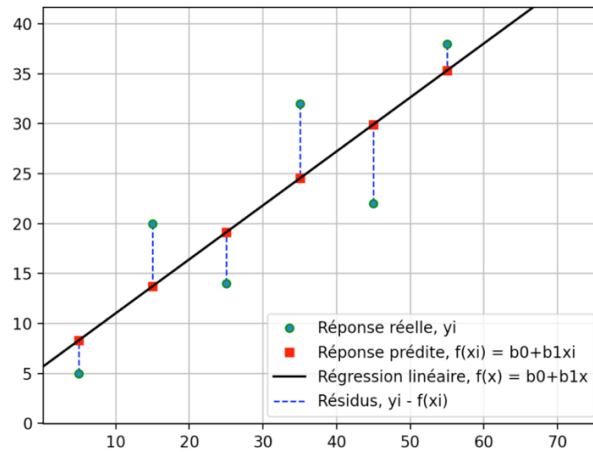


Figure 21 : Exemple de graphe de régression linéaire simple

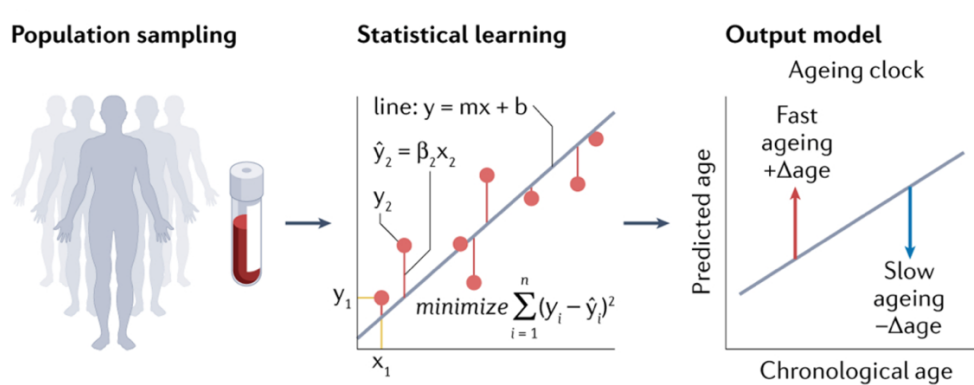


Figure 22 : Modèle d'horloge du vieillissement qui permet la création de 2 sous-groupes : « vieillissement rapide » et « vieillissement lent »
Rutledge et al, 2022

Dans l'étude présentée en figure 22 les auteurs posent la question des trajectoires de vieillissement en comparant l'âge chronologique et l'âge biologique (calculé par le modèle). Les auteurs en utilisant la régression linéaire visent à réduire au maximum les divergences entre les valeurs prédites et les valeurs réelles de l'âge chronologique. Dans le cas où un individu jeune a un âge biologique prédit plus élevé que son âge chronologique, cet individu

sera classé dans une catégorie de vieillissement accéléré permettant ainsi de situer sa trajectoire de vieillissement.

3.2.2.1.2 Le support vector machines (SVM)

Les SVM existent en tant que classifieur ou régresseur. Le SVM en tant que régresseur (Figure 23 schéma de gauche) prédit une valeur continue, comme la valeur de l'âge chronologique dans notre cas. Le principe général est de construire un **hyperplan** (ligne en pointillé sur la figure 23), pour modéliser la relation entre les variables d'entrée et la variable de sortie continue que le modèle tente de prédire (ici l'âge). Les points de données directement sur l'hyperplan (Figure 23 : triangles bleus sur le schéma de la régression) sont appelés des **vecteurs de support**, ils permettent de définir la marge. Ils sont essentiels pour définir la fonction de régression et sont sélectionnés pendant l'entraînement. L'objectif de cet hyperplan est de **minimiser « la marge »** autour de la trajectoire de la régression tout en **minimisant les erreurs de prédiction** (c'est à dire les points en dehors de la marge : point en rouge sur la figure 23). La marge correspond à l'espace virtuel pour lequel les prédictions peuvent varier tout en respectant les données d'entraînement. Les **avantages** de cette méthode sont notamment une robustesse face à un nombre important de variables (multi-omics), une résistance aux valeurs aberrantes (un point extrême a beaucoup de poids en régression linéaire) et une capacité à traiter des données non linéaires (par exemple l'expression d'un gène en fonction de l'âge dont les valeurs ne suivraient pas une droite) (Platt et al, 1999).

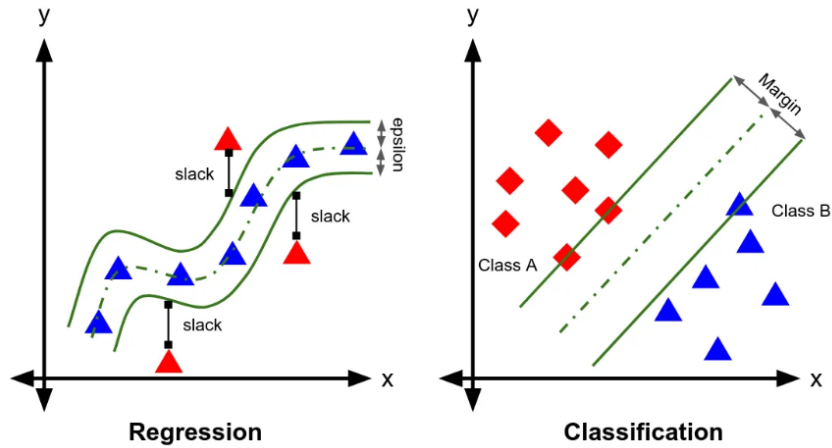


Figure 23 : Exemple de graphe SVM

Le SVM peut être particulièrement utile dans la recherche de biomarqueurs, d'informations sur la fonction d'un gène ou d'une protéine par exemple, en permettant l'analyse complexe de bases de données multi-omics (Libbrecht et al, 2015).

3.2.2.1.3 KNN ou K Nearest Neighbor

L'algorithme des plus proches voisins ou KNN est un algorithme de **classification** et de régression. Nous traiterons son intérêt en tant que classifieur. Si l'on a une nouvelle donnée que l'on souhaite classifier (Figure 24 : carré jaune) alors on **calcule la distance** qui la sépare des autres pour déterminer **ses K plus proches voisins**. Dans notre exemple, on obtient 3 voisins (entourés dans le cercle). Dans ces 3 voisins il y a 2 données de classe B et 1 donnée de classe A, notre nouvelle donnée sera donc attribuée à la **classe majoritaire** c'est-à-dire la classe B.

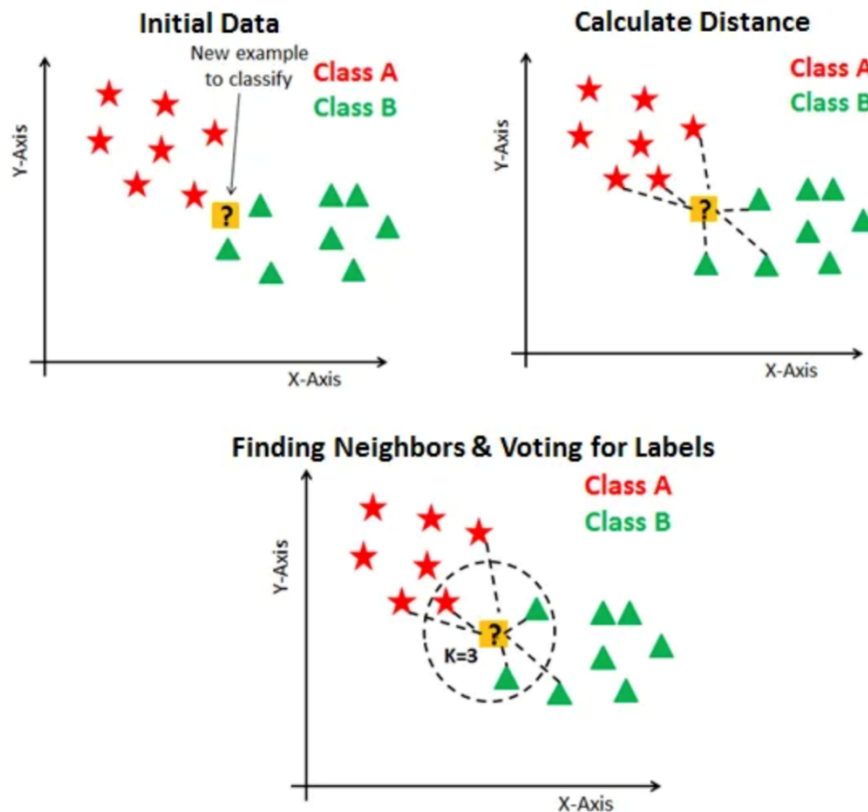


Figure 24 : Exemple de mise en place d'un algorithme KNN

3.2.2.1.4 Les forêts aléatoires (random forest et XG-boost)

Un arbre de décision est un algorithme qui peut être **classifieur** ou **régresseur**. Comme son nom l'indique, il a une structure **arborescente** où chaque nœud représente une variable explicative, chaque branche représente une règle de décision et chaque feuille représente un résultat (âge, groupe de fragilité, etc). Lors de l'apprentissage, l'algorithme choisit pour chaque nœud la caractéristique qui divise le mieux les données selon certains critères. Pour information, on parle de l'indice de Gini en classification et de l'erreur quadratique moyenne en régression. **L'indice de Gini** mesure l'impureté dans un nœud donné. Ainsi, lors de la construction d'un arbre de décision, les divisions qui minimisent l'indice de Gini sont préférées car elles conduisent à des sous-ensembles plus homogènes en termes de classes, facilitant

ainsi la prise de décision pour le modèle de classification. **L'erreur quadratique moyenne (EQM)** en régression représente la moyenne des carrés des différences entre les valeurs prédites par un modèle et les valeurs réelles. Cette mesure fournit une évaluation numérique de la précision globale du modèle de régression, où des valeurs plus faibles d'EQM indiquent un meilleur ajustement du modèle aux données observées. L'arbre est construit de manière **réursive** en divisant les données en sous-ensembles de plus en plus homogènes selon les caractéristiques choisies jusqu'à atteindre un critère d'arrêt, profondeur maximale ou nombre minimal de points dans une feuille. L'algorithme de **random forest** est une technique de « **bagging** » appliquée à l'algorithme d'arbre de décision (Figure 25). Cet algorithme crée N arbres de décision de façon indépendante et agrège, par **vote majoritaire** pour les classifications et par moyenne pour les régressions, les différents arbres afin de diminuer l'incertitude du modèle (Breiman et al, 2001). Dans notre exemple, le random forest fonctionne en créant **plusieurs arbres de décision** qui **votent** pour un âge. Afin de créer de la variabilité dans les arbres, un sous échantillon du jeu de données d'entraînement est aléatoirement sélectionné pour chaque arbre à la fois en termes de points choisis (certains donneurs) mais aussi de caractéristiques des données (certaines variables : sexe, fonctions, statut de fragilité).

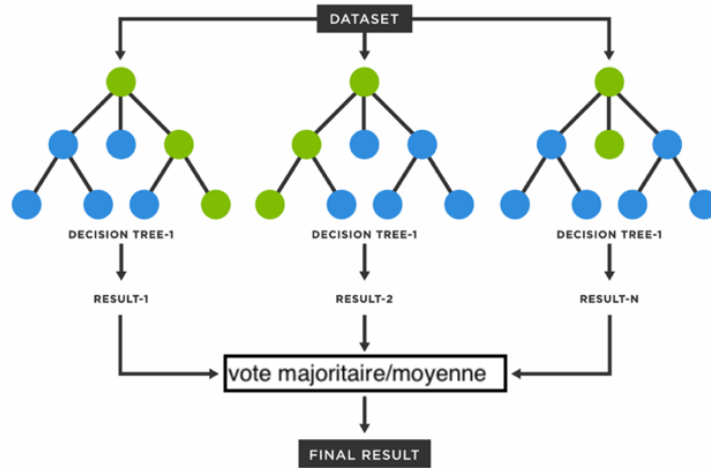


Figure 25 : Exemple d'un arbre de décision « bagging method »

Naue et al, 2017

XGBoost est quant à lui une technique de « **boosting** » appliquée à l'algorithme d'arbre de décision (Figure 26). Cette fois-ci, les arbres ne sont pas combinés de façon parallèle mais **séquentielle**. L'idée est que le prochain arbre essaie de corriger les erreurs du précédent. XGBoost utilise une **descente de gradient** afin de minimiser les erreurs commises par chaque arbre au fur et à mesure. Dans une étude récente visant à prédire l'âge physiologique, plusieurs algorithmes ont été évalués et l'algorithme XGBoost a finalement été choisi de par ses performances les plus intéressantes (Bernard et al, 2023).

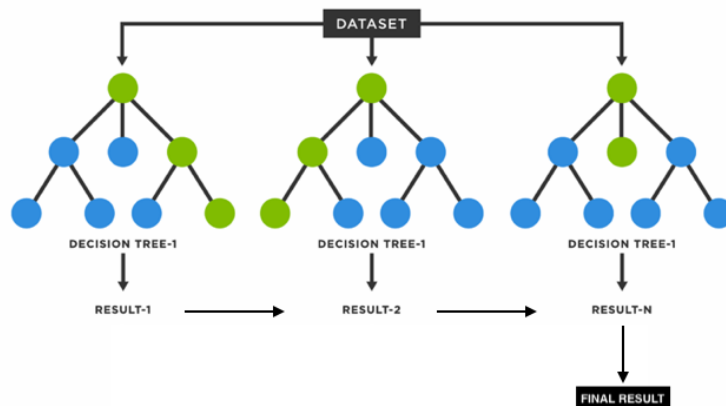


Figure 26 : Exemple d'un arbre de décision « boosting method »

Bernard et al, 2023

3.2.2.2 Algorithmes pour l'apprentissage non supervisé

3.2.2.2.1 K-means

Il s'agit d'un algorithme de regroupement « clustering » **non supervisé** qui regroupe les données en un certain nombre de groupes **définis par l'utilisateur**. Ces k groupes sont définis par leur distance à un centroïde. On commence par répartir aléatoirement les centroïdes (Figure 27 : 1). Chaque point est assigné au centroïde dont il est le plus proche (Figure 27 : 2). Chaque centroïde est ensuite recalculé comme étant le barycentre de l'ensemble des points qui lui sont assignés (Figure 27 : 3). On réitère le processus d'affectation et de calcul de barycentre jusqu'à convergence (Figure 27 : 4). Les centroïdes permettent ensuite d'attribuer nos données à des classes (Figure 27 : 5). Pour information, la distance doit être choisie par l'utilisateur mais les choix les plus communs sont la distance euclidienne ou la distance de Manhattan. La distance de Manhattan mesure la somme des déplacements horizontaux et verticaux entre deux points, tandis que la distance euclidienne mesure la longueur de la ligne droite reliant ces points dans un espace euclidien.

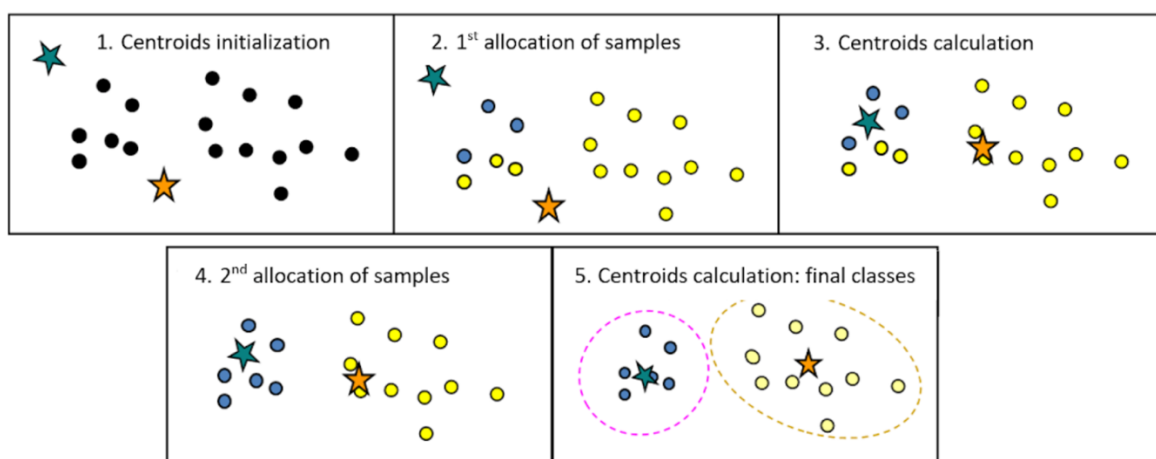


Figure 27 : Étape de l'algorithme KNN

3.2.2.2 Analyse en composantes principales (ACP)

L'Analyse en Composantes Principales (ACP) est une méthode **linéaire** datant de 1993. Elle est utilisée pour **réduire la dimension des données**, particulièrement utile lorsque les échantillons sont composés de nombreuses variables. L'idée principale est de projeter ces données dans un nouvel ensemble d'axes, appelés **composantes principales**, de manière à **maximiser la variance des variables**. Pour rappel, une composante est une combinaison linéaire des différentes variables que compose le jeu de donnée et la variance des variables se réfère à la dispersion ou à l'étendue des valeurs des variables. Les premières composantes principales représentent **l'information la plus importante** des données, tandis que les dernières composantes représentent le **bruit** ou l'information moins pertinente. Le plus souvent, en réduisant la dimension à deux, on peut visualiser ces regroupements de manière plus compréhensible. Dans l'exemple mentionné (Figure 28), on indique que la première dimension (PC1) récapitule 73% des données ce qui signifie qu'elle capture 73% de la variance totale des données. Les variables qui contribuent à cette première dimension sont **suffisamment informatives** pour permettre un regroupement (clustering) des individus en 3 groupes distincts. Il faudra ensuite analyser le poids des variables impliquées dans chaque dimension. Si par exemple un gène permet la clusterisation en 3 groupes, il faudra ensuite regarder si ces 3 groupes décrivent 3 sous-populations de vieillissement auquel cas ce gène pourrait être un marqueur du vieillissement.

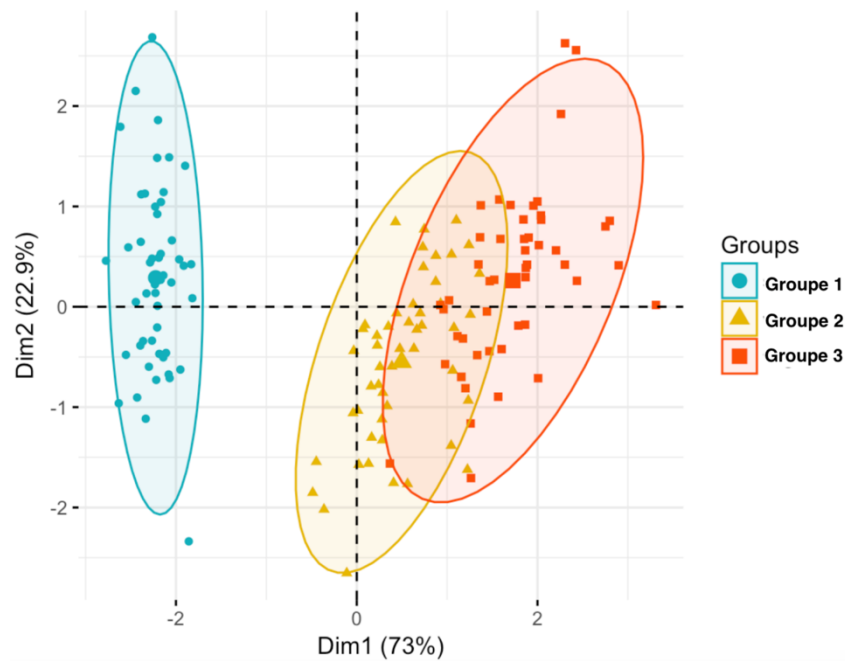


Figure 28 : Représentation d'une ACP avec 2 composantes principales

3.2.2.2.3 t-distributed Stochastic Neighbor Embedding (t-SNE)

La t-SNE (t-distributed Stochastic Neighbor Embedding) est une méthode de **réduction de dimension des données**, développée en 2008. Contrairement à la ACP, elle adopte une approche **non linéaire**. L'algorithme t-SNE consiste à créer une distribution de probabilité qui représente les **similarités entre voisins** dans un espace en **grande** dimension puis dans un espace de **plus petite** dimension. Sur le plan technique, la t-SNE opère en **trois étapes**. Dans un premier temps, les similarités entre les points dans l'espace de grande dimension seront calculées à l'aide de la distance euclidienne (Figure 29). Cette distance représente le degré de **similarité** entre les points. On crée ensuite une **matrice de similarité** pour chaque point de données (Pour les points proche (X1X2): valeur forte, pour les point éloignés (X1X10): valeur faible). Pour chaque x on aura une **distribution gaussienne** centrée autour de ce point.

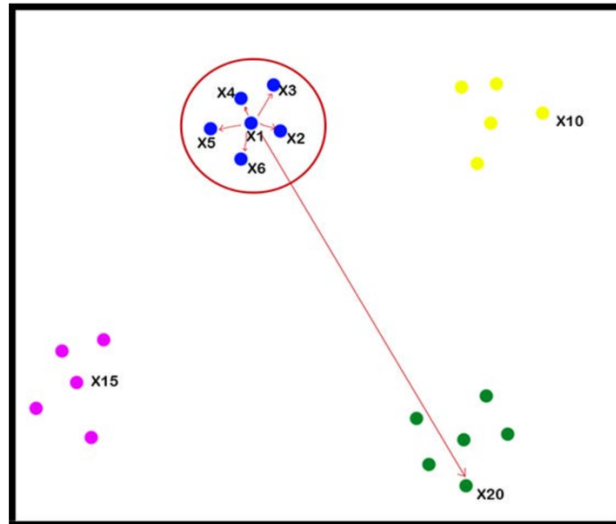


Figure 29 : Distance Euclidienne entre voisins permettant d'établir la matrice de similarité

On obtient alors une liste de probabilité conditionnelle. Dans un second temps, on crée un espace de **petite dimension** où on ne connaît pas les coordonnées exactes de nos points donc on les repartit aléatoirement dans ce nouvel espace (Figure 30). On calcule ensuite la **similarité** des points dans l'espace de petite dimension avec une **distribution de student**. Cette dernière est plus robuste face aux valeurs aberrantes. On obtient également une liste de probabilité conditionnelle.

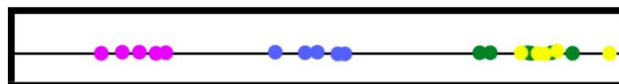


Figure 30 : Représentation de l'espace en faible dimension en 1D

Enfin, l'objectif est de représenter fidèlement nos données dans la petite dimension donc on compare les divergences entre les mesures de similarité entre la grande et la petite dimension par la **mesure de Kullback Leiber**. Afin de minimiser ces différences, on utilise la descente de gradient. L'objectif final étant de **minimiser les différences** entre la grande et la petite dimension afin d'avoir les coordonnées les plus exactes possibles dans l'espace de petite dimension. Les résultats obtenus par la t-SNE sont plus complexes à interpréter ce qui en fait

principalement un outil de **visualisation** (Figure 31). De plus, cette méthode est sensible à l'échelle des données et a un coût computationnel (logistique informatique) plus élevé.

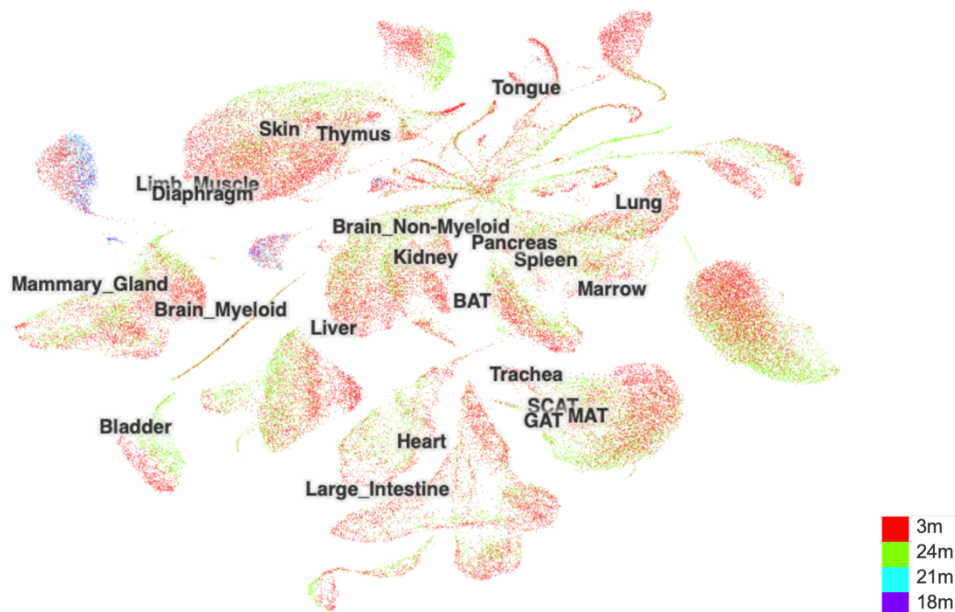


Figure 31 : T-SNE permettant de visualiser des groupes en fonction de l'âge des souris sur différents tissus de l'organisme

A partir du site Tabula Muris Senis (The Tabula Muris Consortium, 2018)

3.2.3 La particularité du deep learning ou apprentissage profond

Le deep learning est une **branche** de l'apprentissage automatique datant des années 2000, qui utilise des **réseaux de neurones profonds**. Ces modèles sont capables de détecter des caractéristiques à partir de données brutes, ce qui est particulièrement utile pour des domaines tels que le traitement d'images, le traitement du langage naturel mais aussi pour l'analyse de données complexes, comme les omics.

Concrètement, il s'agit de plusieurs neurones avec une organisation **hiérarchique** inspirée du **cerveau humain**, ces neurones interconnectés avec des poids en fonction de leurs liens pourront interagir les uns avec les autres dans des couches de profondeur différentes (Figure

32). Chaque neurone fait des calculs simples (combinaison linéaire passée par une fonction d'activation) avec les données d'entrée avant de les faire passer aux autres (Choi et al, 2020).

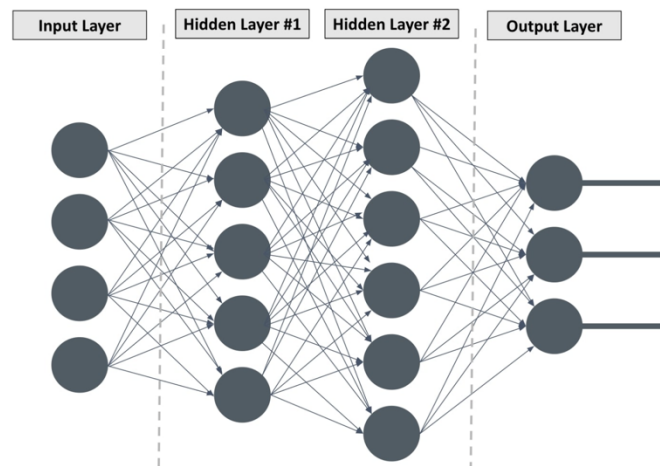


Figure 32 : Exemple de réseaux de neurones profond

Les architectures profondes peuvent nécessiter des **quantités massives** de données pour éviter le **surapprentissage** (une des limites selon laquelle le modèle apprend par cœur et ne pourra donc pas faire une bonne prédiction), et l'entraînement de ces modèles peut exiger des **ressources informatiques significatives**. Le deep learning est essentiellement retrouvé dans l'imagerie pour aider au diagnostic et au dépistage de maladies. On retrouve l'utilisation du deep learning pour l'analyse des IRM (cardiaque, pulmonaire, hépatique, cérébrale) (Tao et al, 2020 ; Khan et al, 2023) et la détection des lésions (logiciel LYNA pour Lymph Node Assistant) ou encore pour la détection de tumeurs à partir de biopsies ou dans le dépistage précoce de cancer avec des mammographies intelligentes (Logiciel Therapixel : <https://www.therapixel.com/>) etc. De façon plus technique, il s'agit d'une méthode d'approximation de fonction. En effet, des données ou « entrées » sont présentés à une fonction, puis les réponses sont comparées à un objectif. On corrige par la suite les paramètres de la fonction considérée par rapport au résultat obtenu. En pratique, une base de données constituée de milliers d'images médicales est soumise au logiciel qui apprend à reconnaître, par exemple, différents types de cancers, de métastases, ou de facteurs aggravants sur les

biopsies, ce qui permet ensuite de détecter les tumeurs sur des nouveaux patients afin de faciliter et d'accompagner le clinicien (Sechopoulos et al, 2021). Le clinicien peut définir des régions d'intérêt (ROI) et classifiez les lésions en différents sous types, ou bien le modèle peut automatiquement définir des sous-types de lésions selon des différences morphologique par exemple et ceux sans à priori, réduisant ainsi le temps d'analyse nécessaire pour le clinicien (Figure 33).

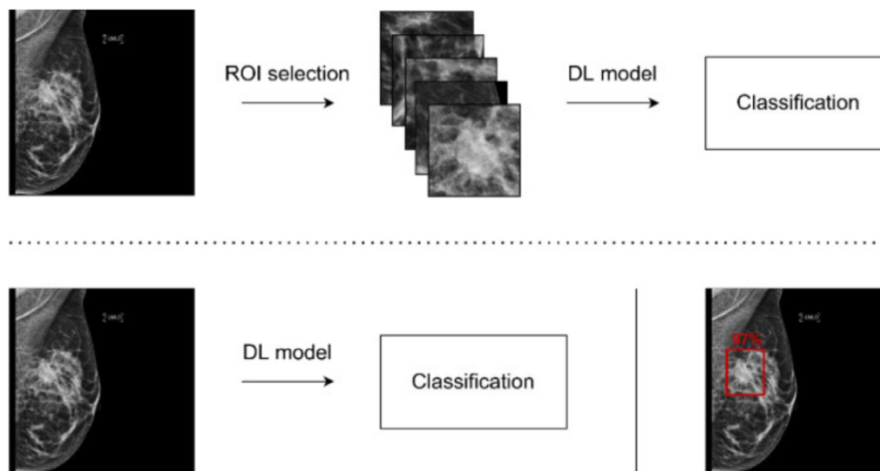


Figure 33 : Étapes de création d'un modèle de reconnaissance d'image
Balkenende et al, 2022

Une des limites du Deep learning est le phénomène de la « boîte noire » c'est-à-dire que l'entrée est fournie à l'algorithme et une sortie émerge, mais il n'est pas exactement clair quelles caractéristiques ont été identifiées ou comment elles ont informé la sortie du modèle. Des méthodes d'explicabilité de ces modèles complexes sont donc nécessaires.

3.2.4 Explicabilité des modèles ou la compréhension des facteurs qui influencent le modèle

Comme nous l'avons vu précédemment, certains modèles sont très complexes et il n'est pas toujours évident de comprendre l'impact de nos variables sur la sortie d'un modèle. Pourtant pour le chercheur la **compréhension des variables** impliquées dans un résultat est essentielle.

Il existe plusieurs méthodes d'explicabilité des données comme le LIME (Local Interpretable

Model-agnostic Explanations), Partial Dependence Plots (PDP) ou encore la méthode des SHAP-value. Nous ne parlerons que de la méthode des **Shapley values** souvent utilisée dans la littérature concernant le vieillissement. Sur le plan technique, la méthode des Shapley Values attribue une valeur à chaque variable dans la prédiction en considérant toutes les façons dont elles interagissent. Les Shapley values reflètent la **contribution** de chaque variable dans différentes combinaisons possibles. En moyennant ces contributions sur toutes les combinaisons, la méthode assure une répartition juste pour chaque variable, déterminant ainsi **leur importance relative** dans la prédiction (Lundberg et al, 2017). En raison de la simplicité du graphe à analyser, cette méthode permet de supprimer les variables ayant la moins grande contribution au modèle, c'est-à-dire celles ayant le moins d'impact sur la question biologique finale. L'exemple ci-dessous (Figure 34) comprend les données du **NHANES dataset** récupérées entre 1999 et à 2018. La méthode des Shapley values a permis d'ordonner le poids des variables concernant la prédiction de l'âge (issue finale) en se basant sur des caractéristiques à l'échelle individuelle.

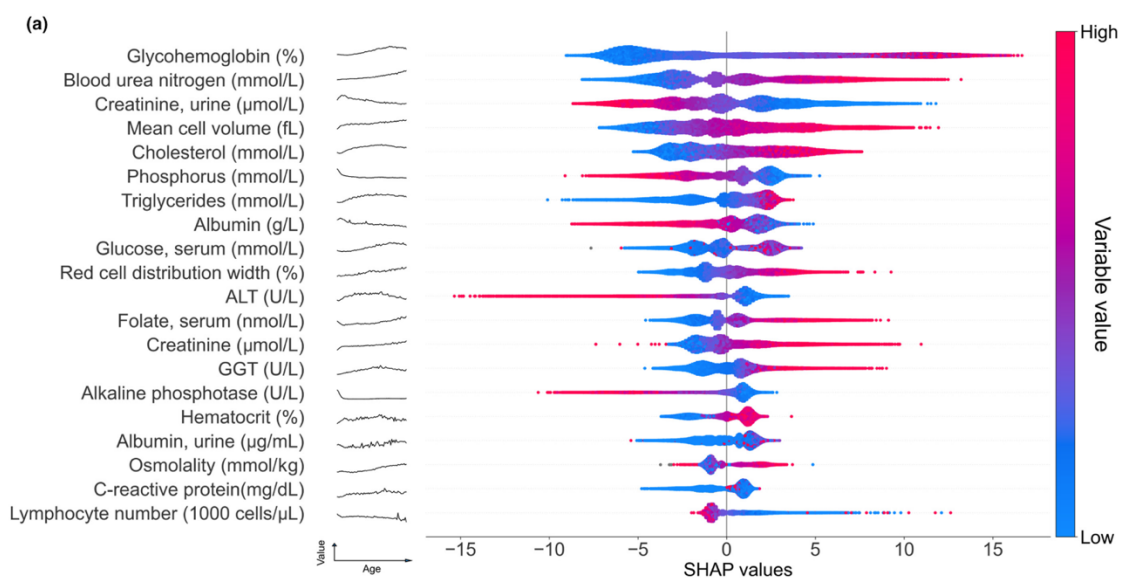


Figure 34 : Exemple d'un graphe des Shapley values
Bernard et al, 2023

Selon ce graphique, les variables à gauche sont rangées dans l'ordre d'importance pour la question biologique posée, à savoir la détermination de l'âge. La variable la plus importante dans ce modèle est donc le % d'hémoglobine glyquée (Figure 33). Les valeurs positives des Shapley values correspondent à un âge biologique élevé (facteur « de risque ») et les valeurs négatives (facteur « protecteur ») à un âge biologique plus faible. Les couleurs correspondent à la valeur de la variable avec en rouge une valeur haute et en bleu une valeur faible. Dans cet exemple, un % élevé d'hémoglobine glyquée est associé à un âge avancé (Bernard et al, 2023). Cependant, la méthode des Shapley values ne permet **pas d'évaluer de relations de causalité**, d'analyser des relations complexes entre plus de 2 variables à la fois, ni de déterminer automatiquement des profils intéressants.

3.3 Applications actuelles

Après avoir exploré le principe général du machine learning ainsi que ses différents aspects techniques, nous allons maintenant nous intéresser aux applications dans le domaine de la recherche sur le vieillissement.

3.3.1 L'analyse des « big data »

Le terme « **Big data** » a été créé dans les années 90 par un informaticien qui voulait décrire un ensemble de données particulier pour lequel les statistiques classiques ne suffisent pas. Ces données se caractérisent par cinq aspects principaux à savoir leur **volume** (grande quantité d'informations), leur **vitesse** (de création, de collecte, de mesure), leur **variété** (différences de nature, format et structure), leur **valeur** (données clés dans différents domaines générant des profits), leur **véracité** (qualité, précision et fiabilité) (Bourany et al, 2019). L'apprentissage automatique devient particulièrement utile lorsque l'ensemble de

données que l'on souhaite analyser est trop volumineux ou trop complexe. On parle de données trop **volumineuses** dans le cas, par exemple, de grandes cohortes pour lesquelles de nombreuses données sont disponibles pour chaque individu (cliniques, épidémiologiques, génétiques, mesures biologiques...) (Gialluisi, et al, 2019). Les données sont considérées comme trop **complexes** pour une analyse humaine lorsque le nombre de caractéristiques ou de variables est trop élevé. Cette complexité se retrouve dans le cas d'ensembles de données comportant de nombreuses dimensions comme des images haute résolution (exemple : IRM), des séquences temporelles volumineuses (exemple : glycémies à intervalles réguliers sur une période de temps longue) ou des données non structurées telles que du texte brut (exemples : ordonnances, comptes rendus de radiologie). Pouvoir gérer et analyser ces données constitue la première étape pour répondre ensuite aux différentes questions d'intérêt dans le cadre d'un vieillissement en bonne santé, notamment la prédiction de l'âge biologique, l'identification de biomarqueurs etc (figure 35). Parmi les « Big data » on retrouve par exemple **les grandes cohortes** comme la cohorte NHANES (National Health and Nutrition Examination Survey), NuAge (Cohorte de Quebec Longitudinal Study of Aging), ELSA (English Longitudinal Study of Ageing), BLSA (Cohorte de Baltimore Longitudinal Study of Aging) ou encore la UK biobanq contenant des informations collectées sur un demi millions d'individus depuis 2006. Par exemple, le jeu de données ou « dataset » du NHANES fournit des informations importantes sur la prévalence des maladies, les habitudes nutritionnelles, les comportements de santé, et d'autres facteurs liés à la santé. Les enquêtes NHANES sont menées périodiquement et comprennent des examens médicaux, des questionnaires et des entretiens pour recueillir des données détaillées sur la santé physique et mentale, l'alimentation, l'activité physique, l'utilisation de médicaments etc. Ces cohortes permettent ensuite de répondre à différentes problématiques par exemple la prédiction et l'évaluation

des facteurs de risque de la mortalité à 5 ans (Wei et al, 2021) en utilisant les données de 1999 à 2014 ou encore la prédiction de l'âge physiologique (Bernard et al, 2023) en utilisant les données de 2008 à 2018. Le choix de la cohorte dépend souvent des questions spécifiques de recherche et des objectifs de l'étude sur le vieillissement en question (Figure 35).

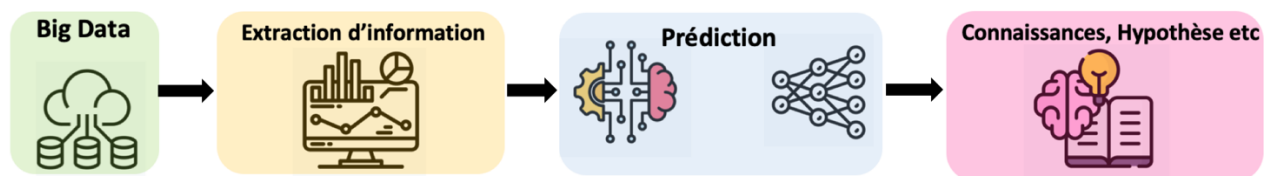


Figure 35 : Intérêt du machine learning : des « Big data » à l'amélioration des connaissances scientifiques

De plus, de nombreuses bases de données sont disponibles en libre accès en ligne comme la Gene Weaver database qui fournit des informations sur les relations entre les gènes, les maladies et les fonctions biologiques ou encore la Human Ageing Genomic Resources (HAGR). La HAGR est une plateforme qui regroupe plusieurs bases de données et outils liés au génome humain et au vieillissement tel que GenAge (gènes associés à la longévité et aux maladies liées à l'âge), AnAge (longévité animale, données sur la durée de vie de différentes espèces), GenDR (gènes liés à la restriction calorique et à son impact sur la durée de vie), LongevityMap (carte interactive permettant d'explorer des informations génétiques liées à la longévité) (Figure 36). Il est souvent simple d'acquérir une grande quantité de données, mais la clé réside ensuite dans la capacité à les exploiter efficacement.

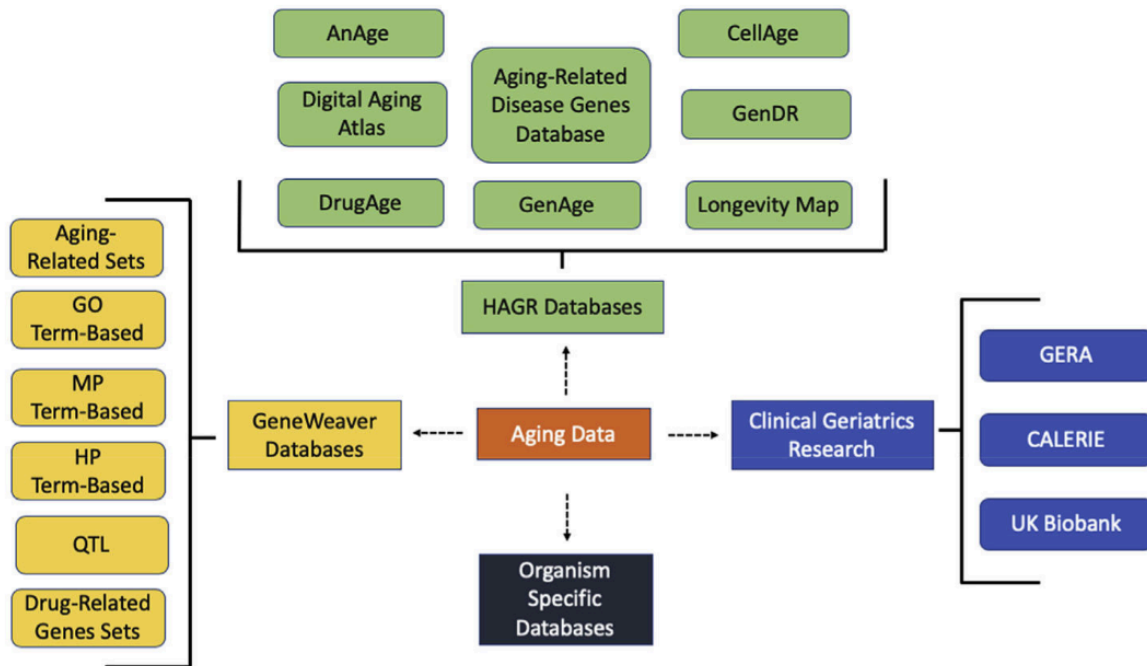


Figure 36 : Liste non exhaustive de bases de données sur le vieillissement

3.3.2 Prédications

La prédiction suscite un intérêt dans divers domaines de recherche que nous allons détailler.

3.3.2.1 La recherche clinique : L'âge biologique et les trajectoires de vieillissement

L'objectif majeur en clinique est de pouvoir prévenir le déclin fonctionnel et la fragilité afin de favoriser le vieillissement en bonne santé. Il s'agit donc de prédire l'âge biologique, meilleur informateur de l'état de santé que l'âge chronologique (Yang et al, 2022) et d'établir des trajectoires de vieillissement pour adapter les soins. Les modèles de machine learning, alimentés par des ensembles de données massives comprenant des informations génétiques, des marqueurs métaboliques et des données cliniques, peuvent fournir des estimations plus précises de l'âge biologique d'un individu. Il n'existe pas encore de consensus pour définir l'âge biologique et de nombreuses études tentent de le prédire par différentes données d'entrée et d'algorithmes. Une approche connue dans ce domaine est le développement des « **aging clocks** » ou « **horloges de vieillissement** ». Globalement, les horloges de

vieillessement sont des modèles d'apprentissage automatique qui apprennent des formules mathématiques pour estimer l'âge d'un individu à partir de caractéristiques qui varient avec l'âge tout au long de la vie (telles que les données omics : épigénomiques, métabolomiques, protéomiques, transcriptomiques, ou encore les données cliniques etc) (Rutledge et al, 2022). Par exemple les "**horloges épigénétiques**" prédisent l'âge d'un individu en fonction des modifications de la méthylation de l'ADN à l'échelle génomique (Horvath et al, 2013 ; Bocklandt et al, 2011 ; Hannum et al, 2013). Ces horloges épigénétiques exploitent des matrices de méthylation d'ADN pour évaluer l'état de méthylation de sites CpG dans le génome humain (Horvath et al, 2018). En utilisant des modèles construits via l'apprentissage automatique supervisé, elles parviennent à prédire l'âge biologique d'une source d'ADN (provenant d'un ou plusieurs tissus mais aussi de PBMCs) en fonction de l'état de méthylation des sites CpG sur des centaines de sites (site gratuit pour déterminer l'âge biologique en insérant des données de méthylation issues de sites CpG (<https://dnamage.genetics.ucla.edu/home>)). Globalement, les recherches indiquent que des patients présentant, selon leur horloge épigénétique, un âge estimé plus élevé que leur âge chronologique, ont un risque accru de problèmes de santé liés au vieillissement, tels que des maladies neurodégénératives, une diminution de la condition physique et cognitive, divers types de cancer, et une mortalité accrue. À l'inverse, un âge épigénétique inférieur à l'âge chronologique est associé au statut centenaire (Ambatipudi et al, 2017 ; Philips SS et al, 2017 ; Horvath et al, 2018) (Figure 37). Il est important de noter les limites de ces horloges épigénétiques. Tout d'abord, bien qu'elles soient souvent fortement corrélées avec l'âge chronologique, ces horloges ne fournissent pas nécessairement une explication causale des processus de vieillissement. De plus, la variabilité entre les individus, influencée par des facteurs génétiques, environnementaux ou liés au mode de vie, peut limiter la précision de la

prédiction de l'âge biologique pour certains individus. De même, ces horloges mesurent principalement des changements épigénétiques globaux sans spécifier les processus biologiques sous-jacents, ce qui peut limiter leur capacité à refléter des changements liés à des maladies spécifiques ou à des aspects particuliers du vieillissement. De plus, la plupart des études utilisant ces horloges sont basées sur des données transversales, limitant ainsi la capacité à suivre l'évolution des changements épigénétiques au fil du temps. Enfin, les signatures épigénétiques peuvent varier selon les types de tissus, ce qui nécessite des ajustements spécifiques lors de l'application de ces horloges à d'autres types de tissus ou organes.

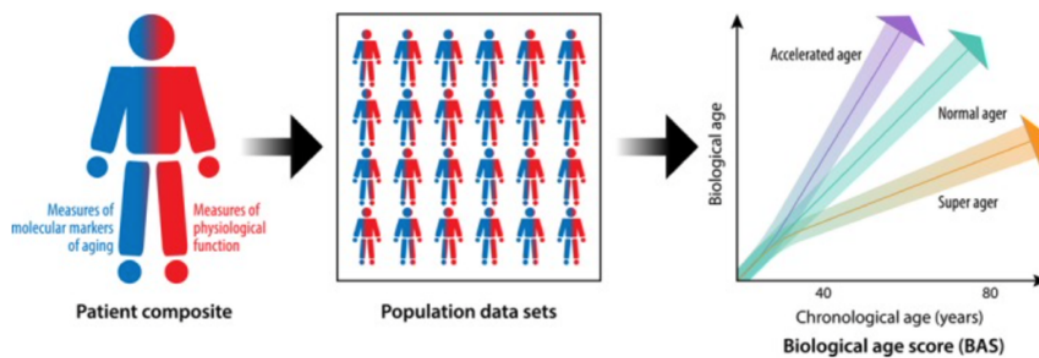


Figure 37 : Exemple de trajectoires de vieillissement
Philips SS et al, 2017

Il existe également des horloges comme iAGE basé sur des **données provenant de l'étude du système immunitaire** (concentration plasmatique en CXCL9, % de lymphocytes T mémoires etc) qui permettent de suivre la morbidité, la fragilité des individus ou encore la défaillance spécifique d'organes (Sayed et al, 2021).

L'âge biologique peut aussi être mesuré au niveau d'organes précis comme **l'âge biologique du cerveau**. Sur le même principe, il permettrait d'évaluer l'avancée ou de prédire l'évolution de certaines maladies neurodégénératives comme la maladie d'Alzheimer associée à l'âge et

liée au déclin des fonctions (Cumplido-Mayoral I, et al, 2023). L'explicabilité permet ensuite de nommer les paramètres les plus contributifs pour la prédiction de l'âge biologique, ce qui permet de déterminer des **biomarqueurs** de suivi ou de prédiction du vieillissement ou de certaines pathologies associées.

Ces modèles pourraient également être utilisés pour déterminer les **trajectoires de santé individuelles intégrées**, en prenant en compte divers facteurs tels que le mode de vie, l'environnement et les antécédents médicaux, afin de proposer des thérapeutiques adaptées. Dans la pratique clinique, le deep learning permet le développement **d'outils originaux**. Par exemple, l'indice LensAge proposé en 2023 et calculé en comparant LensAge à l'âge chronologique, serait un indicateur novateur du vieillissement biologique basé sur des photographies de l'œil humain (Figure 38). Il reflète le taux de vieillissement par rapport aux individus de même âge, surpassant significativement l'âge chronologique en termes de prévision des risques de maladies liées à l'âge. Cette approche adaptable à des photographies de smartphones suggère une utilité potentielle pour l'auto-examen régulier de l'état de vieillissement, en faisant de cet outil un indicateur idéal pour l'évaluation clinique et l'auto-surveillance de l'âge biologique (Li R et al, 2023).

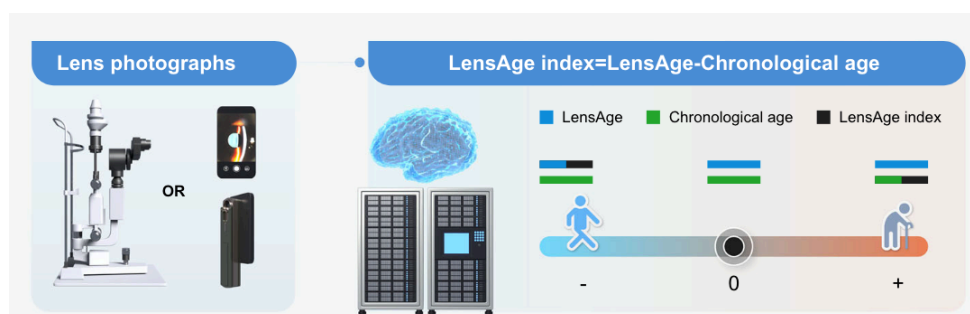


Figure 38 : Outils utilisés au cours de l'étude LensAge
Li R et al, 2023

3.3.2.2 *La recherche fondamentale : Gènes, voies dérégulées*

Un des principaux défis dans la recherche sur le vieillissement chez l'Homme réside dans l'absence d'un ensemble complet et exploitable de **biomarqueurs**, essentiels pour suivre l'évolution du vieillissement et évaluer l'efficacité d'interventions thérapeutiques futures. Pour identifier ces biomarqueurs, il est impératif de **caractériser** les voies et gènes dont la régulation est altérée, afin de comprendre les mécanismes sous-jacents à la maintenance de l'intégrité des différents piliers perturbée au cours du vieillissement.

Initialement, des études sur le ver *C. elegans* ont établi une corrélation entre certains gènes et la longévité de l'espèce (Wan et al, 2013). Chez l'Homme, le deep learning est particulièrement utilisé dans la recherche de marqueurs du vieillissement et a évolué de manière significative, offrant des avancées majeures. Une étude menée en 2016 (Putin et al), a identifié cinq marqueurs de biochimie **clinique non exhaustifs**, tels que l'albumine, le glucose, la phosphatase alcaline, l'urée et les érythrocytes, permettant de suivre le vieillissement. Aujourd'hui, l'approche est **multi-omique**, incluant par exemple la transcriptomique qui analyse les ARN messagers pour fournir des indications sur l'expression génique, et la protéomique qui examine les protéines pour comprendre les processus biologiques. Une étude récente de 2023 (Oh et al, 2023) a utilisé des modèles d'apprentissage automatique pour analyser les niveaux de protéines spécifiques de nombreux organes et présentes dans le plasma sanguin humain, mesurant ainsi l'âge organe spécifique « ageotype » et donc le vieillissement spécifique à chaque organe. Leurs découvertes ont révélé des liens entre le vieillissement accéléré d'organes et les maladies spécifiques liées à ces organes (par exemple, le vieillissement du rein associé à une insuffisance rénale). Ils ont par exemple démontré que l'ageotype rénal était le plus significativement associé aux maladies métaboliques (diabète, obésité, hypercholestérolémie et hypertension), l'ageotype

cardiaque était le plus significativement associé aux maladies cardiaques (fibrillation auriculaire et crise cardiaque) et que, l'agéotype cérébral était le plus significativement associé aux maladies cérébrovasculaires etc. Ils ont découvert que près de 20 % de la population présentait un vieillissement fortement accéléré dans un organe et seulement 1,7 % présentait un vieillissement multi-organes. Ce vieillissement accéléré des organes confèrerait un risque de mortalité 20 à 50 % plus élevé. Ils ont constaté que les personnes présentant un vieillissement cardiaque accéléré avaient un risque d'insuffisance cardiaque accru de 250 % et qu'un vieillissement cérébral et vasculaire accéléré était associé à la progression de la maladie d'Alzheimer (MA) indépendamment et aussi fortement que l'actuel seul biomarqueur sanguin de cette maladie (la protéine Tau-181 plasmatique). Cette étude est très intéressante par sa façon d'aborder l'hétérogénéité du vieillissement chez un individu et donne de nouvelles pistes sur la façon dont il faut concevoir les thérapeutiques contre le vieillissement.

Globalement, l'exploration omique à l'aide d'algorithmes permet de détecter des motifs complexes, d'identifier des relations non linéaires et de déceler des **signatures moléculaires associées au vieillissement**. Il permet d'intégrer des données provenant de différentes échelles et modèles ce qui permet une intégration toujours plus importante d'informations dans l'objectif de pouvoir récapituler la complexité des systèmes biologiques afin de comprendre les processus inhérents à l'espèce humaine comme le vieillissement.

3.3.3 Avancées thérapeutiques

3.3.3.1 *Nouveau paradigme dans l'approche de découverte thérapeutiques*

La prédiction permet d'orienter la recherche sur le vieillissement vers des interventions préventives personnalisées et des soins de santé adaptés à chaque individu. Le processus de

découverte de nouveaux médicaments **a évolué avec le temps** et les nouvelles technologies. Historiquement, les scientifiques reproduisaient les principes actifs efficaces et connus de la nature (salicylates pour l'aspirine par exemple) ou découvraient de nouveaux principes actifs par accident (pénicilline). Au 19^{ème} siècle, la chimie de synthèse a permis de développer des médicaments par rapport au cibles spécifiques impliquées dans certaines maladies. Cependant ces procédés sont **longs et coûteux** et de nombreuses molécules candidates ne parviennent pas à atteindre le stade de la commercialisation en raison de problèmes de sécurité ou encore d'efficacité (Khanna I et al, 2012). Dans la découverte de médicaments, la première étape qui dure approximativement 5 ans consiste à identifier les cibles appropriées (par exemple, les gènes, les protéines) impliquées dans la physiopathologie de la maladie. La seconde étape représente la recherche de médicaments appropriés ou de molécules analogues à des médicaments qui peuvent interférer avec ces cibles. Cette étape comprend des études *in vitro*, *in vivo* sur une durée de 1 à 3 ans (Figure 39).

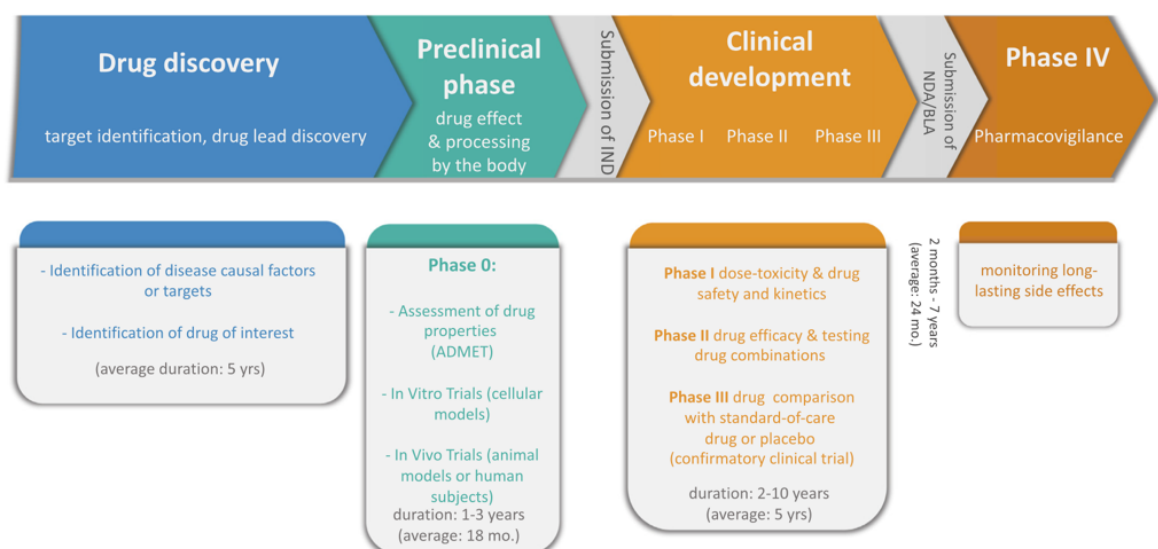


Figure 39 : Pipeline de mise en place d'un nouveau médicament
Réda C et al, 2020

L'importance de trouver de nouvelles méthodes en matière de découverte de médicaments s'est donc imposée.

3.3.3.2 Intérêt du machine learning dans la recherche de thérapeutiques

L'intégration des méthodes de machine learning dans les pipelines de développement de médicaments a différents intérêts qui convergent vers la réduction du **coût** et du **temps** de développement des médicaments (Gupta et al, 2021). La procédure de découverte de médicaments commence par un criblage de millions de composés utilisés pour identifier des composés chimiques ou des substances biologiques ayant une activité spécifique contre une cible biologique donnée, comme par exemple des anticorps générés chez la souris (Figure 39). En ce qui concerne la découverte conventionnelle, de nombreux cycles d'ingénierie sont nécessaires pour modifier les protéines afin qu'elles aient les propriétés dont un médicament a besoin. Ensuite les protéines modifiées sont rigoureusement testées pour déterminer leur sécurité d'utilisation, leur efficacité et leur activité thérapeutique. Dans 50% des cas, ce processus produit des candidats adaptés aux essais cliniques (Mock et al, Nature 2023). Les modèles d'apprentissage automatique seront bientôt utilisés pour filtrer les séquences protéiques qui se lient à la cible souhaitée. De tels outils d'IA trouvent des séquences de liaison probables dans un pool initial de millions de molécules. D'autres modèles d'apprentissage automatique seront utilisés pour prédire les propriétés qu'auront ces séquences de protéines et concevoir des candidats améliorés. Les essais automatisés testeront les candidats, puis les feedbacks permettront d'améliorer les prédictions. Ces méthodes semblent plus fiables et efficaces car 90% des candidats découverts en trois fois moins de temps passent à l'étape suivante (Figure 40).

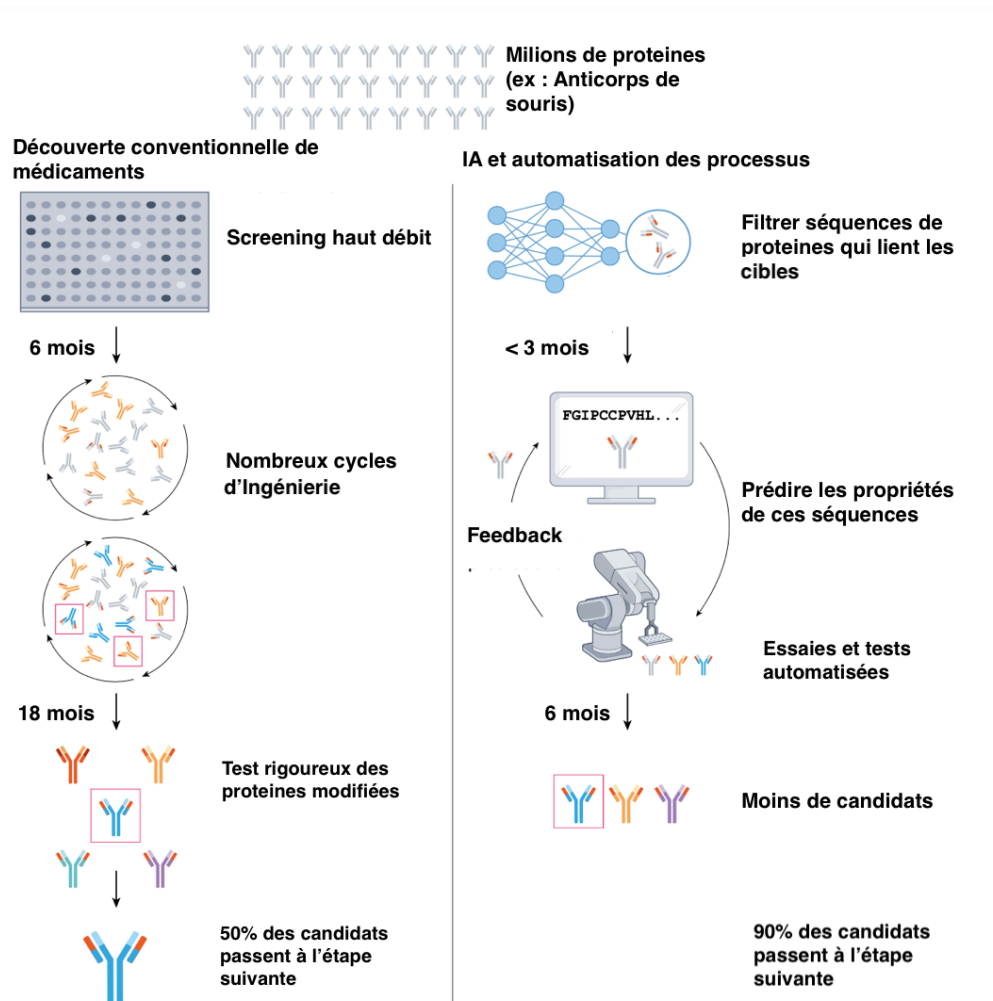


Figure 40 : Résumé de l'intérêt de l'IA dans la découverte de médicaments
Mock et al, Nature 2023

Le screening computationnel ou criblage à grande échelle permet de chercher dans une base de composés chimiques ou molécules biologiques de nouvelles molécules médicaments pour leur capacité à affecter une cible biologique spécifique ou à produire un effet pharmacologique recherché (exemple l'action de sénolytiques). Ces potentiels candidats médicaments sont très nombreux et les modèles (le plus souvent SVM et random forest) permettent ensuite de détecter des patterns cachés dans de grandes collections de données chimiques (Zhang L, et al 2021). Ces « patterns cachés » font référence à des structures ou des tendances que ne sont pas évidentes lorsqu'on examine les données brutes. Ces structures

sont par exemple des relations complexes et non linéaire entre les composés chimiques etc. Dans cet exemple, des molécules en vert et bleu sont les molécules connues ayant permis l'entraînement du modèle et les molécules en oranges sont les molécules sénolytiques nouvellement découvertes (Figure 41). Ils ont ensuite testé l'effet sénolytiques des 21 « Top Hits » candidats, sur des fibroblastes humains après induction de la senescence par oncogène ou traitement pharmacologique (Etoposide : anti-cancéreux). Ils ont pu mettre en évidence 3 composés : le ginkgetin, l'oléandrine et le périplocine avec des performances robustes et comparables à ceux déjà existant dans la littérature, sans induire d'effets indésirables significatifs. De façon intéressante ces composés sont des produits naturels que l'on trouve dans les phytothérapies traditionnelles : Ginkgo biloba (ginkgetin), Nerium oleander (oléandrin) et Periploca sepium (périplocine). Ce sont des composés dont les profils d'ADME-toxicologie sont bien connus. Cette technique a donc diminué le temps et le coût expérimental car le modèle a été entraîné sur des données déjà publiés provenant d'ensemble de données préexistants.

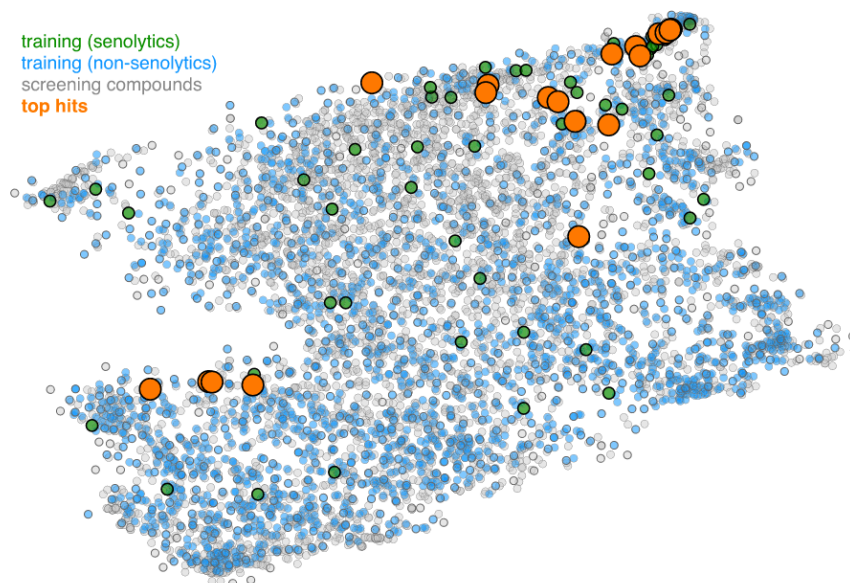


Figure 41 : Visualisation t-SNE de screening de candidats sénolytiques
Smer-Barreto, 2022

Depuis l'épidémie du Covid-19, le **repositionnement de médicaments** est devenu la méthode la plus utilisée pour développer des médicaments thérapeutiques ou des combinaisons potentielles de médicaments en raison du long cycle de conception de médicaments *de novo*. Le repositionnement de médicaments se définit par l'identification de nouvelles utilisations ou indications pour des médicaments existants, qui ont déjà été approuvés pour d'autres indications médicales. Il s'agit d'une approche basée sur l'analyse de données, où des algorithmes de machine learning sont utilisés pour extraire des informations à partir de grandes bases de données biologiques, médicales et chimiques afin de prédire de nouvelles indications pour des médicaments existants. Plutôt que de tester de nouveaux composés, le repositionnement de médicaments se concentre sur l'analyse des informations existantes pour tirer des conclusions sur de nouvelles utilisations potentielles (Yang et al, 2022). Avec l'exemple du Covid-19, un logiciel appelé DeepPurpose (Huang K et al, 2020) a été développé pour déterminer les interactions médicament-cible en intégrant des données basées sur le codage des molécules médicamenteuses et des séquences d'acides aminés afin de déterminer si les médicaments connus agissaient sur les cibles du Covid-19 (Belyaeva et al, 2021). En perspective, l'objectif primordial des thérapies consiste à utiliser ces thérapies de manière préventive, avant l'apparition des déclins, ce qui rend la **tolérance aux effets secondaires** particulièrement cruciale. Cette exigence crée une fenêtre étroite entre les bénéfices et les risques des thérapies, exigeant une précision accrue dans la prédiction des caractéristiques ADME (Administration-Distribution-Métabolisation-Elimination) de la molécule, de la dose et de la définition de la population cible. L'administration de doses minimales est privilégiée. Une stratégie potentielle pour maximiser les effets tout en minimisant les risques consiste à favoriser la synergie d'action entre les médicaments. Le recours au machine learning permet alors de déterminer les **associations médicamenteuses optimales**, tenant compte de divers

facteurs individuels, dans le but d'optimiser les bénéfices tout en réduisant les risques (Kim SK et al, 2021). Parallèlement, le machine learning joue également un rôle essentiel dans la prédiction du mode d'action et des caractéristiques ADME (absorption, distribution, métabolisation, élimination) des médicaments.

Concernant la **sélection de populations cibles**, une étude récente a suggéré l'utilisation de biomarqueurs du vieillissement, notamment l'horloge de méthylation de l'ADN, pour orienter ces sélections (Sandalova et al, 2023).

Le machine learning trouve aussi son intérêt dans **la validation de thérapeutiques**. Comme nous l'avons vu précédemment des études récentes ont souligné l'intérêt de médicaments (sénolytiques) ciblant sélectivement les cellules. Un défi majeur réside dans l'identification fiable des cellules sénescents car aucun marqueur universel n'est disponible en raison de la complexité et de l'hétérogénéité de la sénescence. Des approches d'imagerie et d'apprentissage automatique ont été développées pour relever ce défi. Des algorithmes basés sur les caractéristiques nucléaires, altérées lors de la sénescence, ont été conçus pour prédire la sénescence dans divers types cellulaires, dans le but de préserver l'intégrité des cellules saines et de limiter l'élimination à certains tissus spécifiques. Ces algorithmes peuvent de plus être facilement adaptés à d'autres contextes. En effet, dans l'exemple du cancer de tel algorithme pourrait permettre de limiter les dommages des chimiothérapies aux cellules cancéreuses seules (Duran I et al, 2024).

3.3.3.3 La médecine de précision

L'objectif ultime de ces prédictions réside dans l'avancement vers la **médecine personnalisée**, pour laquelle les traitements sont spécifiquement adaptés à chaque individu. En exploitant les résultats des prédictions, il devient envisageable de concevoir des **thérapies ciblées**,

similaires à celles déjà utilisées en oncologie, et de les étendre à la gestion des conditions liées au vieillissement. Cette approche implique la recherche de molécules directement associées aux biomarqueurs du vieillissement présents chez chaque patient. Le machine learning joue alors un rôle crucial dans la synthèse de peptides, la conception de petites molécules et le développement de médicaments *de novo*, comme l'ont souligné Dugger et al (2018). La médecine personnalisée vise à optimiser l'efficacité des interventions médicales tout en minimisant les effets indésirables associés aux traitements, offrant ainsi une approche de soins plus individualisée et adaptée à chaque patient. Par exemple, il existe des algorithmes permettant de stratifier les patients selon leur nombre de cellules sénescentes afin de proposer des traitements adaptés (Duran I et al, 2024). Ces prédictions permettent un **diagnostic précoce** en amont des déclinés offrant ainsi la possibilité d'intervenir à un stade précoce pour prévenir ou atténuer les effets irréversibles néfastes sur la santé. Cette approche de diagnostic précoce peut contribuer à réduire la progression des maladies associées au vieillissement et à améliorer la qualité de vie des individus. Enfin, la prédiction par les trajectoires de vieillissement permettra de cibler les groupes à risques et de développer des outils de prévention efficace.

4 Perspectives

L'essor de l'intelligence artificielle a renforcé le besoin de travailler de façon **pluridisciplinaire** dans le domaine des gérosiences. Les équipes transdisciplinaires, combinant des experts en intelligence artificielle, en biologie, en médecine et en sciences de données, sont devenues essentielles pour exploiter pleinement les données dans la recherche gériatrique. En collaborant de manière étroite, ces diverses disciplines peuvent intégrer les avancées de l'IA pour analyser de vastes ensembles de données biologiques, découvrir des modèles complexes

et élaborer des approches novatrices pour la compréhension et le traitement des processus liés au vieillissement. Cette approche collaborative ouvre la voie à des découvertes significatives et à des applications pratiques contribuant ainsi à améliorer la **qualité de vie des individus vieillissants**.

Cependant, **l'application éthique des nouvelles médecines anti-vieillessement** suscite des interrogations, notamment en ce qui concerne le traitement de personnes en bonne santé, soulevant des préoccupations éthiques, financières et sociales. Ces questionnements portent sur la justification morale de l'utilisation de traitements anti-vieillessement chez des individus en état de santé apparent, soulignant la nécessité d'évaluer attentivement les risques et les avantages potentiels. Sur le plan éthique, il est crucial de déterminer si ces interventions sont proportionnées aux bénéfices escomptés et si elles respectent les principes de justice, d'autonomie et de non-malfaisance. Parallèlement, les implications financières soulèvent des préoccupations quant à l'accès équitable à de telles interventions, évitant ainsi la création d'inégalités en matière de santé. Enfin, les dimensions sociales mettent en lumière les perceptions collectives du vieillissement et de la quête de l'immortalité, invitant à des réflexions sur la manière dont ces avancées médicales peuvent influencer la société et les dynamiques intergénérationnelles.

L'utilisation de l'IA n'est également pas sans limites, notamment en termes de coût, de temps de développement et de nécessité d'expertise. Suite au rapport Villani sur l'intelligence artificielle commandé par le gouvernement en 2017, destiné à définir une feuille de route sur l'IA, deux grands défis ont été identifiés à savoir : « Améliorer les diagnostics médicaux à l'aide de l'intelligence artificielle » et la « Sécurité des systèmes IA : Sécuriser, certifier et fiabiliser les systèmes fondés sur l'intelligence artificielle ». Ainsi, la confiance en l'intelligence

artificielle est devenue un enjeu national avec la frugalité. La frugalité consiste à être sobre en termes de coût énergétique et de gestion de ressource. Ces challenges financiers peuvent être lié au besoin d'investissements pour acquérir les infrastructures informatiques avancées et pour former le personnel à la manipulation de ces technologies complexes. De plus, le temps de développement des modèles d'IA peut être significatif, impliquant des phases d'apprentissage et d'ajustement qui peuvent prolonger les délais avant la mise en œuvre effective des solutions. La nécessité d'une expertise spécialisée constitue également un défi, car l'efficacité des systèmes d'IA dépend souvent de la compétence de professionnels qualifiés dans le domaine. Ces obstacles soulignent la nécessité d'une approche **réfléchie et planifiée** lors de l'intégration de l'IA, avec un besoin de gestion efficace des ressources financières, une prise en compte des délais de développement et un investissement dans le développement des compétences nécessaires.

La découverte de processus biologiques grâce aux techniques de machine learning représente une ère prometteuse dans la recherche médicale. Cependant, il est essentiel de reconnaître que malgré les avancées technologiques, l'être humain reste **intrinsèquement complexe**. Notre biologie est le produit de multiples facteurs interagissant de manière dynamique, et notre compréhension de ces mécanismes est encore incomplète. Il est essentiel de garder à l'esprit que les machines ne sont pas des substituts parfaits à la compréhension humaine. Même avec l'intégration de vastes ensembles de données, les modèles générés par les algorithmes de machine learning peuvent ne pas capturer pleinement la richesse et la complexité de la biologie humaine. L'humain est bien plus qu'une simple somme de données biologiques ; nos émotions, nos expériences et nos interactions sociales jouent également un rôle crucial dans notre santé. Ainsi, bien que les techniques de machine learning soient des

outils puissants dans la recherche biomédicale, il est important de les considérer comme des aides à la compréhension plutôt que comme des solutions finales. L'approche la plus efficace consiste à utiliser les machines comme des instruments **complémentaires** à l'expertise humaine, en reconnaissant les limites de chaque approche et en cherchant à combiner les forces des deux pour avancer dans la compréhension et le traitement des maladies afin de **promouvoir le vieillissement en bonne santé.**

5 Bibliographie.

1. Adam EK, Quinn ME, Tavernier R, McQuillan MT, Dahlke KA, Gilbert KE. Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. *Psychoneuroendocrinology*. sept 2017;83:25-41.
2. Ader I, Penicaud L, Andrieu S. Healthy Aging Biomarkers: The INSPIRE's Contribution. *J Frailty Aging* 10. 24 avr 2021;313-319 (2021).
3. Ahadi S, Zhou W, Schüssler-Fiorenza Rose SM, Sailani MR, Contrepolis K, Avina M, et al. Personal aging markers and ageotypes revealed by deep longitudinal profiling. *Nat Med*. janv 2020;26(1):83-90.
4. Alalousi A, Razif R, AbuAlhaj M, Anbar M, Nizam S. A Preliminary Performance Evaluation of K-means, KNN and EM Unsupervised Machine Learning Methods for Network Flow Classification. *IJECE*. 1 avr 2016;6(2):778.
5. Ambatipudi S, Horvath S, Perrier F, Cuenin C, Hernandez-Vargas H, Le Calvez-Kelm F, et al. DNA methylome analysis identifies accelerated epigenetic ageing associated with postmenopausal breast cancer susceptibility. *European Journal of Cancer*. avr 2017;75:299-307.
6. Angelova PR, Abramov AY. Role of mitochondrial ROS in the brain: from physiology to neurodegeneration. *FEBS Letters*. mars 2018;592(5):692-702.
7. Antczak M, Michaelis M, Wass MN. Environmental conditions shape the nature of a minimal bacterial genome. *Nat Commun*. 15 juill 2019;10(1):3100.
8. Anton S, Leeuwenburgh C. Fasting or caloric restriction for Healthy Aging. *Experimental Gerontology*. oct 2013;48(10):1003-5.
9. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Visvanathan K, et al. Leisure Time Physical Activity and Mortality: A Detailed Pooled Analysis of the Dose-Response Relationship. *JAMA Intern Med*. 1 juin 2015;175(6):959.
10. Austad SN. Candidate Bird Species for Use in Aging Research. *ILAR Journal*. 1 janv 2011;52(1):89-96.

11. Badal VD, Vaccariello ED, Murray ER, Yu KE, Knight R, Jeste DV, et al. The Gut Microbiome, Aging, and Longevity: A Systematic Review. *Nutrients*. 7 déc 2020;12(12):3759.
12. Baechle JJ, Chen N, Makhijani P, Winer S, Furman D, Winer DA. Chronic inflammation and the hallmarks of aging. *Molecular Metabolism*. août 2023;74:101755.
13. Bagherniya M, Butler AE, Barreto GE, Sahebkar A. The effect of fasting or calorie restriction on autophagy induction: A review of the literature. *Ageing Res Rev*. 2018 Nov;47:183-197. doi: 10.1016/j.arr.2018.08.004. Epub 2018 Aug 30. PMID: 30172870.
14. Baker GT, Sprott RL. Biomarkers of aging. *Experimental Gerontology*. janv 1988;23(4-5):223-39.
15. Balakrishnan G, Zhao A, Sabuncu MR, Gutttag J, Dalca AV. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *IEEE Trans Med Imaging*. août 2019;38(8):1788-800.
16. Barzilai NR. TARGETING AGING WITH METFORMIN (TAME). *Innovation in Aging*. 1 juill 2017;1(suppl_1):743-743.
17. Baştanlar Y, Ozuysal M. Introduction to machine learning. *Methods Mol Biol*. 2014;1107:105-28. doi: 10.1007/978-1-62703-748-8_7. PMID: 24272434.
18. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel JP, et al. The World report on ageing and health: a policy framework for healthy ageing. *The Lancet*. mai 2016;387(10033):2145-54.
19. Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, et al. Quantification of biological aging in young adults. *Proc Natl Acad Sci USA [Internet]*. 28 juill 2015 [cité 8 déc 2023];112(30). Disponible sur: <https://pnas.org/doi/full/10.1073/pnas.1506264112>
20. Bergamini E, Cavallini G, Donati A, Gori Z. The Role of Autophagy in Aging: Its Essential Part in the Anti-Aging Mechanism of Caloric Restriction. *Annals of the New York Academy of Sciences*. oct 2007;1114(1):69-78.
21. Bernard D, Doumard E, Ader I, Kemoun P, Pagès J, Galinier A, et al. Explainable machine learning framework to predict personalized physiological aging. *Ageing Cell*. août 2023;22(8):e13872.

22. Biagi E, Rampelli S, Turrioni S, Quercia S, Candela M, Brigidi P. The gut microbiota of centenarians: Signatures of longevity in the gut microbiota profile. *Mechanisms of Ageing and Development*. juill 2017;165:180-4.
23. Bjornsson H. An integrated epigenetic and genetic approach to common human disease. *Trends in Genetics*. août 2004;20(8):350-8.
24. Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat Med*. oct 2006;12(10):1133-8.
25. Bocklandt S, Lin W, Sehl ME, Sánchez FJ, Sinsheimer JS, Horvath S, et al. Epigenetic Predictor of Age. Landsberger N, éditeur. *PLoS ONE*. 22 juin 2011;6(6):e14821.
26. Bortz J, Guariglia A, Klaric L, Tang D, Ward P, Geer M, et al. Biological age estimation using circulating blood biomarkers. *Commun Biol*. 26 oct 2023;6(1):1089.
27. Bourany T. Les 5V du big data: Regards croisés sur l'économie. 3 juin 2019;n° 23(2):27-31.
28. Breiman L. random forest. *Machine Learning*. 2001;45(1):5-32
29. Brunet A. Old and new models for the study of human ageing. *Nat Rev Mol Cell Biol*. sept 2020;21(9):491-3.
30. Burch JB, Augustine AD, Frieden LA, Hadley E, Howcroft TK, Johnson R, et al. Advances in Geroscience: Impact on Healthspan and Chronic Disease. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 1 juin 2014;69(Suppl 1):S1-3.
31. Camera DM, Smiles WJ, Hawley JA. Exercise-induced skeletal muscle signaling pathways and human athletic performance. *Free Radical Biology and Medicine*. sept 2016;98:131-43.
32. Campisi J, d'Adda Di Fagagna F. Cellular senescence: when bad things happen to good cells. *Nat Rev Mol Cell Biol*. sept 2007;8(9):729-40.
33. Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E. From discoveries in ageing research to therapeutics for healthy ageing. *Nature*. 11 juill 2019;571(7764):183-92.
34. Cannon WB. Organization for physiological homeostasis. juill 1929;

35. Cellerino A, Valenzano DR, Reichard M. From the bush to the bench: the annual *Nothobranchius* fishes as a new model system in biology. *Biological Reviews*. mai 2016;91(2):511-33.
36. Cesari M, Araujo de Carvalho I, Amuthavalli Thiyagarajan J, Cooper C, Martin FC, Reginster JY, et al. Evidence for the Domains Supporting the Construct of Intrinsic Capacity. *The Journals of Gerontology: Series A*. 10 nov 2018;73(12):1653-60.
37. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *N Engl J Med* 2007;356:2131-42
38. Chen D, Kerr C. The Epigenetics of Stem Cell Aging Comes of Age. *Trends in Cell Biology*. juill 2019;29(7):563-8.
39. Chen S, Gan D, Lin S, Zhong Y, Chen M, Zou X, Shao Z, Xiao G. Metformin in aging and aging-related diseases: clinical applications and relevant mechanisms. *Theranostics*. 2022 Mar 6;12(6):2722-2740. doi: 10.7150/thno.71360. PMID: 35401820; PMCID: PMC8965502.
40. Choi rene Y, Coyner AS. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Translational Vision Science & Technology*. févr 2020;
41. Civitarese AE, Carling S, Heilbronn LK, Hulver MH, Ukropcova B, Deutsch WA, et al. Calorie Restriction Increases Muscle Mitochondrial Biogenesis in Healthy Humans. Barzilai N, éditeur. *PLoS Med*. 6 mars 2007;4(3):e76.
42. Clark RI, Salazar A, Yamada R, Fitz-Gibbon S, Morselli M, Alcaraz J, et al. Distinct Shifts in Microbiota Composition during *Drosophila* Aging Impair Intestinal Function and Drive Mortality. *Cell Reports*. sept 2015;12(10):1656-67.
43. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *The Lancet*. mars 2013;381(9868):752-62.
44. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, et al. Caloric Restriction Delays Disease Onset and Mortality in Rhesus Monkeys. *Science*. 10 juill 2009;325(5937):201-4.

45. Cozachenco D, Ribeiro FC, Ferreira ST. Defective proteostasis in Alzheimer's disease. *Ageing Res Rev.* 2023 Mar;85:101862. doi: 10.1016/j.arr.2023.101862. Epub 2023 Jan 21. PMID: 36693451.
46. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age and Ageing.* 1 janv 2019;48(1):16-31.
47. Cumplido-Mayoral I, García-Prat M, Operto G, Falcon C, Shekari M, Cacciaglia R, et al. Biological brain age prediction using machine learning on structural neuroimaging data: Multi-cohort validation against biomarkers of Alzheimer's disease and neurodegeneration stratified by sex. *eLife.* 17 avr 2023;12:e81067.
48. Daya A, Donaka R, Karasik D. Zebrafish models of sarcopenia. *Dis Model Mech.* 1 mars 2020;13(3):dmm042689.
49. Deo RC. Machine Learning in Medicine. *Circulation.* 17 nov 2015;132(20):1920-30.
50. Dugger SA, Platt A, Goldstein DB. Drug development in the era of precision medicine. *Nat Rev Drug Discov.* mars 2018;17(3):183-96.
51. Duran I, Pombo J, Sun B, Gallage S, Kudo H, McHugh D, Bousset L, Barragan Avila JE, Forlano R, Manousou P, Heikenwalder M, Withers DJ, Vernia S, Goldin RD, Gil J. Detection of senescence using machine learning algorithms based on nuclear features. *Nat Commun.* 2024 Feb 3;15(1):1041. doi: 10.1038/s41467-024-45421-w. PMID: 38310113; PMCID: PMC10838307.
52. Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, et al. Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome. *Nature.* mai 2003;423(6937):293-8.
53. Escobar KA, Cole NH, Mermier CM, VanDusseldorp TA. Autophagy and aging: Maintaining the proteome through exercise and caloric restriction. *Aging Cell.* févr 2019;18(1):e12876.
54. Fabris F, Magalhães JPD, Freitas AA. A review of supervised machine learning applied to ageing research. *Biogerontology.* avr 2017;18(2):171-88.

55. Ferrucci L, Gonzalez-Freire M, Fabbri E, Simonsick E, Tanaka T, Moore Z, et al. Measuring biological aging in humans: A quest. *Aging Cell*. févr 2020;19(2):e13080.
56. Fieschi M., (1983), «Sphinx: un système d'aide à la décision en médecine», Thèse de doctorat en biologie humaine, faculté de médecine de Marseille
57. Fontana L, Meyer TE, Klein S, Holloszy JO. Long-term calorie restriction is highly effective in reducing the risk for atherosclerosis in humans. *Proc Natl Acad Sci USA*. 27 avr 2004;101(17):6659-63.
58. Foretz M, Hébrard S, Leclerc J, Zarrinpashneh E, Soty M, Mithieux G, et al. Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *J Clin Invest*. 1 juill 2010;120(7):2355-69.
59. Franceschi C. Inflammaging as a Major Characteristic of Old People: Can It Be Prevented or Cured? *Nutrition Reviews*. 28 juin 2008;65:S173-6.
60. Franceschi C, Capri M, Monti D, Giunta S, Olivieri F, Sevini F, et al. Inflammaging and anti-inflammaging: A systemic perspective on aging and longevity emerged from studies in humans. *Mechanisms of Ageing and Development*. janv 2007;128(1):92-105.
61. Franceschi C, Garagnani P, Morsiani C, Conte M, Santoro A, Grignolio A, et al. The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates. *Front Med*. 12 mars 2018;5:61.
62. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 1 mars 2001;56(3):M146-57.
63. Gardiner SL, Milanese C, Boogaard MW, Buijsen RAM, Hogenboom M, Roos RAC, et al. Bioenergetics in fibroblasts of patients with Huntington disease are associated with age at onset. *Neurol Genet*. oct 2018;4(5):e275.
64. Gardiner SL, Milanese C, Boogaard MW, Buijsen RAM, Hogenboom M, Roos RAC, et al. Bioenergetics in fibroblasts of patients with Huntington disease are associated with age at onset. *Neurol Genet*. oct 2018;4(5):e275.

65. Gerou M, Hall B, Woof R, Allsop J, Kolb SJ, Meyer K, et al. Amyotrophic lateral sclerosis alters the metabolic aging profile in patient derived fibroblasts. *Neurobiology of Aging*. sept 2021;105:64-77.
66. Ghosh TS, Rampelli S, Jeffery IB, Santoro A, Neto M, Capri M, et al. Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries. *Gut*. juill 2020;69(7):1218-28.
67. Gialluisi A, Di Castelnuovo A, Donati MB, de Gaetano G, Iacoviello L, the Moli-sani Study Investigators. Machine Learning Approaches for the Estimation of Biological Aging: The Road Ahead for Population Studies. *Front Med*. 3 juill 2019;6:146.
68. Giampieri E, Ostan R, Guidarelli G, Salvioli S, Berendsen AAM, Brzozowska A, et al. A Novel Approach to Improve the Estimation of a Diet Adherence Considering Seasonality and Short Term Variability – The NU-AGE Mediterranean Diet Experience. *Front Physiol*. 5 mars 2019;10:149.
69. Glossmann HH, Lutz OMD. Metformin and Aging: A Review. *Gerontology*. 2019;65(6):581-90.
70. Golbidi S, Daiber A, Korac B, Li H, Essop MF, Laher I. Health Benefits of Fasting and Caloric Restriction. *Curr Diab Rep*. 2017 Oct 23;17(12):123. doi: 10.1007/s11892-017-0951-7. PMID: 29063418.
71. Green CL, Lamming DW, Fontana L. Molecular mechanisms of dietary restriction promoting health and longevity. *Nat Rev Mol Cell Biol*. janv 2022;23(1):56-73.
72. Grillo AS, Bitto A, Kaeberlein M. The NDUFS4 Knockout Mouse: A Dual Threat Model of Childhood Mitochondrial Disease and Normative Aging. In: Weissig V, Edeas M, éditeurs. *Mitochondrial Medicine* [Internet]. New York, NY: Springer US; 2021 [cité 18 déc 2023]. p. 143-55. (Methods in Molecular Biology; vol. 2277). Disponible sur: https://link.springer.com/10.1007/978-1-0716-1270-5_10
73. Guarente L, Kenyon C. Genetic pathways that regulate ageing in model organisms. *Nature*. 9 nov 2000;408(6809):255-62.

74. Guerville F, De Souto Barreto P, Ader I, Andrieu S, Casteilla L, Dray C, et al. REVISITING THE HALLMARKS OF AGING TO IDENTIFY MARKERS OF BIOLOGICAL AGE. *J Prev Alz Dis.* 2019;1-9.
75. Guo J, Huang X, Dou L, Yan M, Shen T, Tang W, et al. Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. *Sig Transduct Target Ther.* 16 déc 2022;7(1):391.
76. Guo Y, Guan T, Shafiq K, Yu Q, Jiao X, Na D, Li M, Zhang G, Kong J. Mitochondrial dysfunction in aging. *Ageing Res Rev.* 2023 Jul;88:101955. doi: 10.1016/j.arr.2023.101955. Epub 2023 May 15. PMID: 37196864.
77. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers.* août 2021;25(3):1315-60.
78. Hamczyk MR, Nevado RM, Baretino A, Fuster V, Andrés V. Biological Versus Chronological Aging. *Journal of the American College of Cardiology.* mars 2020;75(8):919-30.
79. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell.* janv 2013;49(2):359-67.
80. Harman D. Aging: A Theory Based on Free Radical and Radiation Chemistry. *Journal of Gerontology.* 1 juill 1956;11(3):298-300.
81. Hartmann C, Herling L, Hartmann A, Köckritz V, Fuellen G, Walter M, et al. Systematic estimation of biological age of in vitro cell culture systems by an age-associated marker panel. *Front Aging.* 15 févr 2023;4:1129107.
82. Hartmann C, Herling L, Hartmann A, Köckritz V, Fuellen G, Walter M, et al. Systematic estimation of biological age of in vitro cell culture systems by an age-associated marker panel. *Front Aging.* 15 févr 2023;4:1129107.
83. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. *Experimental Cell Research.* déc 1961;25(3):585-621.

84. Heilbronn LK, Ravussin E. Calorie restriction and aging: review of the literature and implications for studies in humans. *The American Journal of Clinical Nutrition*. sept 2003;78(3):361-9.
85. Hekimi S, Lapointe J, Wen Y. Taking a “good” look at free radicals in the aging process. *Trends in Cell Biology*. oct 2011;21(10):569-76.
86. Hernandez-Segura A, Nehme J, Demaria M. Hallmarks of Cellular Senescence. *Trends in Cell Biology*. juin 2018;28(6):436-53.
87. Hertel J, Friedrich N, Wittfeld K, Pietzner M, Budde K, Van der Auwera S, et al. Measuring Biological Age via Metabonomics: The Metabolic Age Score. *J Proteome Res*. 5 févr 2016;15(2):400-10.
88. Hickson LJ, Langhi Prata LGP, Bobart SA, Evans TK, Giorgadze N, Hashmi SK, et al. Senolytics decrease senescent cells in humans: Preliminary report from a clinical trial of Dasatinib plus Quercetin in individuals with diabetic kidney disease. *EBioMedicine*. sept 2019;47:446-56.
89. Hood L, Heath JR, Phelps ME, Lin B. Systems Biology and New Technologies Enable Predictive and Preventative Medicine. *Science*. 22 oct 2004;306(5696):640-3.
90. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biology*. 2013;14(10):R115.
91. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biology*. 2013;14(10):R115.
92. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. juin 2018;19(6):371-84.
93. Huang K, Fu T, Glass LM, Zitnik M, Xiao C, Sun J. DeepPurpose: a deep learning library for drug-target interaction prediction. *Bioinformatics*. 2021 Apr 1;36(22-23):5545-5547. doi: 10.1093/bioinformatics/btaa1005. PMID: 33275143; PMCID: PMC8016467.
94. Hunter RW, Hughey CC, Lantier L, Sundelin EI, Peggie M, Zeqiraj E, et al. Metformin reduces liver glucose production by inhibition of fructose-1-6-bisphosphatase. *Nat Med*. sept 2018;24(9):1395-406.

95. Jang JY, Blum A, Liu J, Finkel T. The role of mitochondria in aging. *Journal of Clinical Investigation*. 31 août 2018;128(9):3662-70.
96. Junnila RK, List EO, Berryman DE, Murrey JW, Kopchick JJ. The GH/IGF-1 axis in ageing and longevity. *Nat Rev Endocrinol*. juin 2013;9(6):366-76.
97. Justice JN, Ferrucci L, Newman AB, Aroda VR, Bahnson JL, Divers J, et al. A framework for selection of blood-based biomarkers for geroscience-guided clinical trials: report from the TAME Biomarkers Workgroup. *GeroScience*. déc 2018;40(5-6):419-36.
98. Justice JN, Nambiar AM, Tchkonina T, LeBrasseur NK, Pascual R, Hashmi SK, et al. Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study. *EBioMedicine*. févr 2019;40:554-63.
99. Jylhävä J, Pedersen NL, Hägg S. Biological Age Predictors. *EBioMedicine*. juill 2017;21:29-36.
100. Kaeberlein M, Creevy KE, Promislow DEL. The dog aging project: translational geroscience in companion animals. *Mamm Genome*. août 2016;27(7-8):279-88.
101. Koblan LW, Erdos MR, Wilson C, Cabral WA, Levy JM, Xiong ZM, Tavarez UL, Davison LM, Gete YG, Mao X, Newby GA, Doherty SP, Narisu N, Sheng Q, Krilow C, Lin CY, Gordon LB, Cao K, Collins FS, Brown JD, Liu DR. In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice. *Nature*. 2021 Jan;589(7843):608-614. doi: 10.1038/s41586-020-03086-7. Epub 2021 Jan 6. PMID: 33408413; PMCID: PMC7872200.
102. Klaips CL, Jayaraj GG, Hartl FU. Pathways of cellular proteostasis in aging and disease. *J Cell Biol*. 2018 Jan 2;217(1):51-63. doi: 10.1083/jcb.201709072. Epub 2017 Nov 10. PMID: 29127110; PMCID: PMC5748993.
103. Kemoun Ph, Ader I, Planat-Benard V, Dray C, Fazilleau N, Monsarrat P, et al. A gerophysiology perspective on healthy ageing. *Ageing Research Reviews*. janv 2022;73:101537.
104. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: Linking Aging to Chronic Disease. *Cell*. nov 2014;159(4):709-13.
105. Kerber RA, O'Brien E, Cawthon RM. Gene expression profiles associated with aging and mortality in humans. *Aging Cell*. juin 2009;8(3):239-50.

106. Kevin HP. Mechanisms of Muscle Denervation in Aging: Insights from a Mouse Model of Amyotrophic Lateral Sclerosis. *Aging and disease*. 2015;6(5):380.
107. Khanna I. Drug discovery in pharmaceutical industry: productivity challenges and trends. *Drug Discovery Today*. oct 2012;17(19-20):1088-102.
108. Kim EC, Kim JR. Senotherapeutics: emerging strategy for healthy aging and age-related disease. *BMB Rep*. 31 janv 2019;52(1):47-55.
109. Kim SK, Goughnour PC, Lee EJ, Kim MH, Chae HJ, Yun GY, et al. Identification of drug combinations on the basis of machine learning to maximize anti-aging effects. Rishi A, éditeur. *PLoS ONE*. 28 janv 2021;16(1):e0246106.
110. Kirkwood TBL. Understanding the Odd Science of Aging. *Cell*. févr 2005;120(4):437-47.
111. Koblan LW, Erdos MR, Wilson C, Cabral WA, Levy JM, Xiong ZM, et al. In vivo base editing rescues Hutchinson–Gilford progeria syndrome in mice. *Nature*. 28 janv 2021;589(7843):608-14.
112. Kruempel JCP, Howington MB, Leiser SF. Computational tools for geroscience. *Translational Medicine of Aging*. 2019;3:132-43.
113. Larson-Meyer DE, Heilbronn LK, Redman LM, Newcomer BR, Frisard MI, Anton S, et al. Effect of Calorie Restriction With or Without Exercise on Insulin Sensitivity, β -Cell Function, Fat Cell Size, and Ectopic Lipid in Overweight Subjects. *Diabetes Care*. 1 juin 2006;29(6):1337-44.
114. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*. 18 avr 2018;10(4):573-91.
115. Li Q, Milenkovic T. Supervised Prediction of Aging-Related Genes From a Context-Specific Protein Interaction Subnetwork. *IEEE/ACM Trans Comput Biol and Bioinf*. 1 juill 2022;19(4):2484-98.
116. Li R, Chen W, Li M, Wang R, Zhao L, Lin Y, et al. LensAge index as a deep learning-based biological age for self-monitoring the risks of age-related diseases and mortality. *Nat Commun*. 6 nov 2023;14(1):7126.

117. Li X, Li C, Zhang W, Wang Y, Qian P, Huang H. Inflammation and aging: signaling pathways and intervention therapies. *Signal Transduct Target Ther.* 2023 Jun 8;8(1):239. doi: 10.1038/s41392-023-01502-8. PMID: 37291105; PMCID: PMC10248351.
118. Li Z, Zhang W, Duan Y, Niu Y, Chen Y, Liu X, et al. Progress in biological age research. *Front Public Health.* 12 avr 2023;11:1074274.
119. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet.* juin 2015;16(6):321-32.
120. Lindsell CJ, Pomann GM, Oster RA, Mooney SD, Enders FT. Data science in clinical and translational research: Improving the health of the data to knowledge pipeline. *J Clin Trans Sci.* 2021;5(1):e77.
121. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The Hallmarks of Aging. *Cell.* juin 2013;153(6):1194-217.
122. Lozano-Torres B, Estepa-Fernández A, Rovira M, Orzáez M, Serrano M, Martínez-Máñez R, et al. The chemistry of senescence. *Nat Rev Chem.* 13 juin 2019;3(7):426-41.
123. Lu WH, Rolland Y, Guyonnet S, De Souto Barreto P, Vellas B. Reference centiles for intrinsic capacity throughout adulthood and their association with clinical outcomes: a cross-sectional analysis from the INSPIRE-T cohort. *Nat Aging.* 9 nov 2023;3(12):1521-8.
124. Lujambio A. To clear, or not to clear (senescent cells)? That is the question. *BioEssays* [Internet]. juill 2016 [cité 8 déc 2023];38(S1). Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/bies.201670910>
125. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. 2017 [cité 19 déc 2023]; Disponible sur: <https://arxiv.org/abs/1705.07874>
126. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci.* sept 2018;21(9):1281-9.
127. Maynard C, Weinkove D. The Gut Microbiota and Ageing. In: Harris JR, Korolchuk VI, éditeurs. *Biochemistry and Cell Biology of Ageing: Part I Biomedical Science* [Internet]. Singapore: Springer Singapore; 2018 [cité 8 déc 2023]. p. 351-71. (Subcellular

Biochemistry; vol. 90). Disponible sur: http://link.springer.com/10.1007/978-981-13-2835-0_12

128. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity Prediction using Deep Learning. *Front Environ Sci* [Internet]. 2 févr 2016 [cité 8 déc 2023];3. Disponible sur: <http://journal.frontiersin.org/Article/10.3389/fenvs.2015.00080/abstract>
129. McNeil JJ, Nelson MR, Woods RL, Lockery JE, Wolfe R, Reid CM, et al. Effect of Aspirin on All-Cause Mortality in the Healthy Elderly. *N Engl J Med*. 18 oct 2018;379(16):1519-28.
130. McNeil JJ, Wolfe R, Woods RL, Tonkin AM, Donnan GA, Nelson MR, et al. Effect of Aspirin on Cardiovascular Events and Bleeding in the Healthy Elderly. *N Engl J Med*. 18 oct 2018;379(16):1509-18.
131. Memme JM, Erlich AT, Phukan G, Hood DA. Exercise and mitochondrial health. *J Physiol*. 2021 Feb;599(3):803-817. doi: 10.1113/JP278853. Epub 2019 Dec 9. PMID: 31674658.
132. Mertens J, Reid D, Lau S, Kim Y, Gage FH. Aging in a Dish: iPSC-Derived and Directly Induced Neurons for Studying Brain Aging and Age-Related Neurodegenerative Diseases. *Annu Rev Genet*. 23 nov 2018;52(1):271-93.
133. Mitchell SJ, Scheibye-Knudsen M, Longo DL, De Cabo R. Animal Models of Aging Research: Implications for Human Aging and Age-Related Diseases. *Annu Rev Anim Biosci*. 16 févr 2015;3(1):283-303.
134. Moaddel R, Ubaida-Mohien C, Tanaka T, Lyashkov A, Basisty N, Schilling B, et al. Proteomics in aging research: A roadmap to clinical, translational research. *Aging Cell*. avr 2021;20(4):e13325.
135. Moqri M, Herzog C, Poganik JR, Justice J, Belsky DW, Higgins-Chen A, et al. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell*. août 2023;186(18):3758-75.
136. Moskalev AA, Shaposhnikov MV, Plyusnina EN, Zhavoronkov A, Budovsky A, Yanai H, et al. The role of DNA damage and repair in aging through the prism of Koch-like criteria. *Ageing Research Reviews*. mars 2013;12(2):661-84.

137. Muñoz-Espín D, Serrano M. Cellular senescence: from physiology to pathology. *Nat Rev Mol Cell Biol.* juill 2014;15(7):482-96.
138. Naue J, Hoefsloot HCJ, Mook ORF, Rijlaarsdam-Hoekstra L, van der Zwalm MCH, Henneman P, et al. Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression. *Forensic Science International: Genetics.* nov 2017;31:19-28.
139. Niccoli T, Partridge L. Ageing as a Risk Factor for Disease. *Current Biology.* sept 2012;22(17):R741-52.
140. Oh HSH, Rutledge J, Nachun D, Pálovics R, Abiose O, Moran-Losada P, et al. Organ aging signatures in the plasma proteome track health and disease. *Nature.* 7 déc 2023;624(7990):164-72.
141. Onken B, Driscoll M. Metformin Induces a Dietary Restriction–Like State and the Oxidative Stress Response to Extend *C. elegans* Healthspan via AMPK, LKB1, and SKN-1. Hart AC, éditeur. *PLoS ONE.* 18 janv 2010;5(1):e8758.
142. Opalach K, Rangaraju S, Madorsky I, Leeuwenburgh C, Notterpek L. Lifelong Calorie Restriction Alleviates Age-Related Oxidative Damage in Peripheral Nerves. *Rejuvenation Research.* févr 2010;13(1):65-74.
143. Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience & Biobehavioral Reviews.* avr 2012;36(4):1140-52.
144. Paez HG, Pitzer CR, Alway SE. Age-Related Dysfunction in Proteostasis and Cellular Quality Control in the Development of Sarcopenia. *Cells.* 2023 Jan 7;12(2):249. doi: 10.3390/cells12020249. PMID: 36672183; PMCID: PMC9856405.
145. Palliyaguru DL, Moats JM. Frailty index as a biomarker of lifespan and healthspan: focus on pharmacological interventions. *Mech Ageing Dev [Internet].* juin 2019; Disponible sur: doi: [10.1016/j.mad.2019.03.005](https://doi.org/10.1016/j.mad.2019.03.005)
146. Pal S, Tyler JK. Epigenetics and aging. *Sci Adv.* 2016 Jul 29;2(7):e1600584. doi: 10.1126/sciadv.1600584. PMID: 27482540; PMCID: PMC4966880.

147. Pfefferli C, Jaźwińska A. The art of fin regeneration in zebrafish: The Art of Fin Regeneration. *Regeneration*. avr 2015;2(2):72-83.
148. Phillips SS, Singer BD, Vaughan DE. Molecular and physiological manifestations and measurement of aging in humans. *Aging Cell*. août 2017;16(4):624-33.
149. Phillip JM, Wu PH, Gilkes DM, Williams W, McGovern S, Daya J, et al. Biophysical and biomolecular determination of cellular age in humans. *Nat Biomed Eng*. 11 juill 2017;1(7):0093.
150. Phillip JM, Wu PH, Gilkes DM, Williams W, McGovern S, Daya J, et al. Biophysical and biomolecular determination of cellular age in humans. *Nat Biomed Eng*. 11 juill 2017;1(7):0093.
151. Platt JC. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. MIT press. 1999;
152. Platzer A. Visualization of SNPs with t-SNE. Brusic V, éditeur. *PLoS ONE*. 15 févr 2013;8(2):e56883.
153. Poeschla M, Valenzano DR. The turquoise killifish: a genetically tractable model for the study of aging. Dickinson MH, Voshall LB, Dow JAT, éditeurs. *Journal of Experimental Biology*. 1 févr 2020;223(Suppl_1):jeb209296.
154. Powers ET, Morimoto RI, Dillin A, Kelly JW, Balch WE. Biological and Chemical Approaches to Diseases of Proteostasis Deficiency. *Annu Rev Biochem*. 1 juin 2009;78(1):959-91.
155. Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, et al. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. *Aging*. 18 mai 2016;8(5):1021-33.
156. Réda C, Kaufmann E, Delahaye-Duriez A. Machine learning applications in drug development. *Computational and Structural Biotechnology Journal*. 2020;18:241-52.
157. Ren R, Ocampo A, Liu GH, Izpisua Belmonte JC. Regulation of Stem Cell Aging by Metabolism and Epigenetics. *Cell Metabolism*. sept 2017;26(3):460-74.
158. Ricklefs RE. Insights from comparative analyses of aging in birds and mammals. *Aging Cell*. avr 2010;9(2):273-84.

159. Ristow M, Schmeisser S. Extending life span by increasing oxidative stress. *Free Radical Biology and Medicine*. juill 2011;51(2):327-36.
160. Rockwood K, Howlett SE. Age-related deficit accumulation and the diseases of ageing. *Mechanisms of Ageing and Development*. juin 2019;180:107-16.
161. Rodier F, Campisi J. Four faces of cellular senescence. *Journal of Cell Biology*. 21 févr 2011;192(4):547-56.
162. Rothwell PM, Fowkes FGR, Belch JFF, Ogawa H, Warlow CP, Meade TW. Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. *Lancet* 2011;377:31–41.
163. Rubinsztein DC, Mariño G, Kroemer G. Autophagy and Aging. *Cell*. sept 2011;146(5):682-95.
164. Rutledge J, Oh H, Wyss-Coray T. Measuring biological age using omics data. *Nat Rev Genet*. déc 2022;23(12):715-27.
165. Safdar A, Bourgeois JM, Ogborn DI, Little JP, Hettinga BP, Akhtar M, et al. Endurance exercise rescues progeroid aging and induces systemic mitochondrial rejuvenation in mtDNA mutator mice. *Proc Natl Acad Sci USA*. 8 mars 2011;108(10):4135-40.
166. Sameri S, Samadi P, Dehghan R, Salem E, Fayazi N, Amini R. Stem Cell Aging in Lifespan and Disease: A State-of-the-Art Review. *Curr Stem Cell Res Ther*. 2020;15(4):362-378. doi: 10.2174/1574888X15666200213105155. PMID: 32053079.
167. Sandalova E, Maier AB. Targeting the epigenetically older individuals for geroprotective trials: the use of DNA methylation clocks. *Biogerontology* [Internet]. 16 nov 2023 [cité 5 janv 2024]; Disponible sur: <https://link.springer.com/10.1007/s10522-023-10077-4>
168. Santin Y, Lopez S, Ader I, Andrieu S, Blanchard N, Carrière A, et al. Towards a large-scale assessment of the relationship between biological and chronological aging: The INSPIRE Mouse Cohort. *J Frailty Aging*. 2020;1-11.
169. Sayed N, Huang Y, Nguyen K, Krejciova-Rajaniemi Z, Grawe AP, Gao T, et al. An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity,

- immunosenescence, frailty and cardiovascular aging. *Nat Aging*. 12 juill 2021;1(7):598-615.
170. Schmauck-Medina T, Molière A, Lautrup S, Zhang J, Chlopicki S, Madsen HB, et al. New hallmarks of ageing: a 2022 Copenhagen ageing meeting summary. *Aging*. 31 août 2022;14(16):6829-39.
171. Schmoldt A, Benthe HF, Haberland G. Digitoxin metabolism by rat liver microsomes. *Biochem Pharmacol*. 1 sept 1975;24(17):1639-41.
172. Schnohr P, O’Keefe JH, Marott JL, Lange P, Jensen GB. Dose of Jogging and Long-Term Mortality. *Journal of the American College of Cardiology*. févr 2015;65(5):411-9.
173. Schumacher B, et al. The central role of DNA damage in the ageing process. *Nature*. 2021;592(7856):695–703. doi: 10.1038/s41586-021-03307-7.
174. Schwalm C, Jamart C, Benoit N, Naslain D, Prémont C, Prévét J, et al. Activation of autophagy in human skeletal muscle is dependent on exercise intensity and AMPK activation. *The FASEB Journal*. août 2015;29(8):3515-26.
175. Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Seminars in Cancer Biology*. juill 2021;72:214-25.
176. sene FC, Ruchaya PJ, Domenjo-Vila E, Shin Teoh T, Prata L, Cottle BJ, et al. Aged-senescent cells contribute to impaired heart regeneration. *Aging Cell*. juin 2019;18(3):e12931.
177. Shay JW. Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov*. 2016 Jun;6(6):584-93. doi: 10.1158/2159-8290.CD-16-0062. Epub 2016 Mar 30. PMID: 27029895; PMCID: PMC4893918.
178. Shen P, Yue Y, Park Y. A living model for obesity and aging research: *Caenorhabditis elegans*. *Critical Reviews in Food Science and Nutrition*. 24 mars 2018;58(5):741-54.
179. Shortliffe EH, Davis R, Axline SG, Buchanan BG, Green CC, Cohen SN. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. *Comput Biomed Res*. 1975 Aug;8(4):303-20. doi: 10.1016/0010-4809(75)90009-9. PMID: 1157471.

181. Sierra F. The Emergence of Geroscience as an Interdisciplinary Approach to the Enhancement of Health Span and Life Span. *Cold Spring Harb Perspect Med.* avr 2016;6(4):a025163.
182. Smer-Barreto V, Quintanilla A, Elliott RJR, Dawson JC, Sun J, Campa VM, et al. Discovery of senolytics using machine learning. *Nat Commun.* 10 juin 2023;14(1):3445.
183. Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev.* 15 oct 2010;24(20):2239-63.
184. Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev.* 15 oct 2010;24(20):2239-63.
185. Steve. DNA methylation age of human tissues and cell types. *Genome Biology.* 2013;14(10):R115.
186. Sutton RS, Barto AG. Reinforcement learning: an introduction. Second edition. Cambridge, Massachusetts: The MIT Press; 2018. 526 p. (Adaptive computation and machine learning series).
187. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell.* août 2006;126(4):663-76.
188. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell.* août 2006;126(4):663-76.
189. Taormina G, Ferrante F, Vieni S, Grassi N, Russo A, Mirisola MG. Longevity: Lesson from Model Organisms. *Genes.* 9 juill 2019;10(7):518.
190. Travin DY, Feniouk BA. Aging in Birds. *Biochemistry (Mosc).* 2016 Dec;81(12):1558-1563. doi: 10.1134/S0006297916120178. PMID: 28259133.
191. The Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature.* oct 2018;562(7727):367-72.
192. Tigges J, Krutmann J, Fritsche E, Haendeler J, Schaal H, Fischer JW, et al. The hallmarks of fibroblast ageing. *Mechanisms of Ageing and Development.* juin 2014;138:26-44.

193. Triggler CR, Mohammed I, Bshesh K, Marei I, Ye K, Ding H, MacDonald R, Hollenberg MD, Hill MA. Metformin: Is it a drug for all reasons and diseases? *Metabolism*. 2022 Aug;133:155223. doi: 10.1016/j.metabol.2022.155223. Epub 2022 May 29. PMID: 35640743.
194. Unnikrishnan A, Freeman WM, Jackson J, Wren JD, Porter H, Richardson A. The role of DNA methylation in epigenetics of aging. *Pharmacol Ther*. 2019 Mar;195:172-185. doi: 10.1016/j.pharmthera.2018.11.001. Epub 2018 Nov 9. PMID: 30419258; PMCID: PMC6397707.
195. Valenzano DR, Benayoun BA, Singh PP, Zhang E, Etter PD, Hu CK, et al. The African Turquoise Killifish Genome Provides Insights into Evolution and Genetic Architecture of Lifespan. *Cell*. déc 2015;163(6):1539-54.
196. Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Müller F. Influence of TOR kinase on lifespan in *C. elegans*. *Nature*. déc 2003;426(6967):620-620.
197. Visser M, Schaap LA. Consequences of Sarcopenia. *Clinics in Geriatric Medicine*. août 2011;27(3):387-99.
198. Wan C, Freitas A. Prediction of the pro-longevity or anti-longevity effect of *Caenorhabditis Elegans* genes based on Bayesian classification methods. In: 2013 IEEE International Conference on Bioinformatics and Biomedicine [Internet]. Shanghai, China: IEEE; 2013 [cité 21 déc 2023]. p. 373-80. Disponible sur: <http://ieeexplore.ieee.org/document/6732521/>
199. Wang K, Klionsky DJ. Mitochondria removal by autophagy. *Autophagy*. mars 2011;7(3):297-300.
200. Wheaton WW, Weinberg SE, Hamanaka RB, Soberanes S, Sullivan LB, Anso E, et al. Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *eLife*. 13 mai 2014;3:e02242.
201. Witham MD, Granic A, Pearson E, Robinson SM, Sayer AA. Repurposing Drugs for Diabetes Mellitus as Potential Pharmacological Treatments for Sarcopenia – A Narrative Review. *Drugs Aging*. août 2023;40(8):703-19.

202. Xu M, Bradley EW, Weivoda MM, Hwang SM, Pirtskhalava T, Decklever T, et al. Transplanted Senescent Cells Induce an Osteoarthritis-Like Condition in Mice. *GERONA*. 10 août 2016;glw154.
203. Yaksic E, Lecky V, Sharnprapai S, Tungkhar T, Cho K, Driver JA, et al. DEFINING FRAILITY IN RESEARCH ABSTRACTS: A SYSTEMATIC REVIEW AND RECOMMENDATIONS FOR STANDARDIZATION. *J Frailty Aging*. 2019;1-5.
204. Yang Q, Gao S, Lin J, Lyu K, Wu Z, Chen Y, et al. A machine learning-based data mining in medical examination data: a biological features-based biological age prediction model. *BMC Bioinformatics*. 3 oct 2022;23(1):411.
205. Yang WH, Zheng B, Wu MN, Zhu SJ, Fei FQ, Weng M, et al.. An evaluation system of fundus photograph-based intelligent diagnostic technology for diabetic retinopathy and applicability for research. *Diabetes Ther* (2019) 10:1811–22. doi: 10.1007/s13300-019-0652-0
206. Zhang L, Pitcher LE, Prahalad V, Niedernhofer LJ, Robbins PD. Recent advances in the discovery of senolytics. *Mechanisms of Ageing and Development*. déc 2021;200:111587.
207. Zhang L, Pitcher LE, Yousefzadeh MJ, Niedernhofer LJ, Robbins PD, Zhu Y. Cellular senescence: a key therapeutic target in aging and diseases. *J Clin Invest*. 2022 Aug 1;132(15):e158450. doi: 10.1172/JCI158450. PMID: 35912854; PMCID: PMC9337830.
208. Zhang L, Pitcher LE, Prahalad V, Niedernhofer LJ, Robbins PD. Targeting cellular senescence with senotherapeutics: senolytics and senomorphics. *FEBS J*. 2023 Mar;290(5):1362-1383. doi: 10.1111/febs.16350. Epub 2022 Feb 1. PMID: 35015337.

Title : Machine learning, a tool of interest in promoting healthy aging: from understanding mechanisms to therapeutic developments

Abstract

The current challenge linked to aging arises from the observation that, despite the increase in life expectancy, **the healthspan does not increase significantly**. Aging is a complex and multifactorial process, and the underlying mechanisms are not fully understood. This thesis examines current knowledge, reveals gaps in understanding aging and explores the benefit of a **multidisciplinary approach**, particularly with the growing use of **machine learning**. Examples will be presented to illustrate the relevance of some strategies compared to other models. In the aging research field, the recent establish wish is to overcome the limits that exist between pathologies associated with aging. Indeed, while each pathology is often studied and treated each at a time, the new paradigme acknowledge that new therapeutics should be target **common mechanisms** of aging at the origin of these multiple diseases. The complexity of validating therapies for aging, which is a natural biological process, rely on the need to optimize benefits while minimizing risks. Knowing these challenges, machine learning plays a crucial role in the repositioning of marketing drugs authorizations and in the development of new therapies. It helps optimize screening procedures, making the process more efficient in terms of cost and time. The objective of this thesis is to explain the foundations of this discipline and to provide keys to facilitate its understanding.

Key words : Aging, Artificial Intelligence, Machine Learning, Prediction, Senotherapeutics, biomarkers

AUTEUR : Camille Joly

TITRE : Le machine learning, outil d'intérêt dans la promotion du vieillissement en bonne santé : de la compréhension des mécanismes aux développements thérapeutiques

Directrices de thèse : Pr. Douin Victorine et Dr. Ader Isabelle

RESUME en français

Le défi actuel lié au vieillissement découle du constat que, malgré l'augmentation de l'espérance de vie, **la durée de vie en bonne santé ne progresse pas** de manière significative. Le vieillissement est un processus complexe et multi-factoriel, dont les mécanismes sous-jacents ne sont pas tous connus.

Cette thèse se penche sur les connaissances actuelles, révèle les lacunes de compréhension du vieillissement et explore l'intérêt d'une **approche pluridisciplinaire**, notamment avec l'utilisation croissante du **machine learning**, un domaine en plein essor. Des exemples concrets seront présentés pour illustrer la pertinence des modèles d'analyse déjà développés. Une dimension essentielle de la recherche sur le vieillissement réside dans la nécessité de dépasser les limites qui existent entre les différentes pathologies associées au vieillissement. En effet, alors que chaque pathologie est souvent étudiée de manière isolée, le vieillissement partage des **mécanismes communs à l'origine de diverses affections**, qui conditionnent le développement des approches thérapeutiques. La complexité de la validation des thérapies pour le vieillissement, en tant que processus biologique naturel, est accentuée par la nécessité d'optimiser les bénéfices tout en minimisant les risques. Face à ces défis, le machine learning joue un rôle crucial dans le repositionnement des autorisations de mise sur le marché (AMM) et dans le développement de nouvelles thérapies. Il contribue à optimiser les procédures de criblage, rendant le processus plus efficient en termes de coût et de temps.

Ainsi, l'objectif de ce travail de thèse est d'explicitier les fondements de cette discipline et de fournir des clés pour en faciliter la compréhension.

MOTS-CLES : Vieillissement, Intelligence Artificielle, Machine Learning, Prédiction, Senothérapeutiques, biomarqueurs

DISCIPLINE administrative : Pharmacie

INTITULE ET ADRESSE DE L'UFR OU DU LABORATOIRE : Université Paul Sabatier Toulouse III
– Faculté de santé – Département de Sciences Pharmaceutiques – 31062 Toulouse Cedex 9