

**Faculté des sciences économiques,
sociales, politiques et de communication
École des sciences politiques et sociales (PSAD)**

A qui profite le sport ?

**Activité physique, pratique sportive
et inégalités de santé en Belgique.**

Auteur : Alice Lontie
Promoteur : Bruno Masquelier
Lecteur : Thierry Eggerickx
Année académique : 2024-2025
Master de spécialisation en méthodes quantitatives en sciences sociales

Déclaration de déontologie

« Je déclare sur l'honneur que ce mémoire a été écrit de ma plume, sans avoir sollicité d'aide extérieure illicite, qu'il n'est pas la reprise d'un travail présenté dans une autre institution pour évaluation, et qu'il n'a jamais été publié, en tout ou en partie. Toutes les informations (idées, phrases, graphes, cartes, tableaux, ...) empruntées ou faisant référence à des sources primaires ou secondaires sont référencées adéquatement selon la méthode universitaire en vigueur. Si j'ai eu recours à l'IA, j'ai scrupuleusement respecté les règles établies en la matière par la note du bureau Faculté ESPO en date du 1^{er} juillet 2024, tout particulièrement dans son point 4. Je déclare avoir pris connaissance et adhérer au Code de déontologie pour les étudiant(e)s en matière d'emprunts, de citations et d'exploitation de sources diverses en ce compris l'usage des outils d'intelligence artificielle générative et savoir qu'un plagiat ou que toute autre forme d'irrégularité est une faute grave et est susceptible de donner lieu aux procédures et sanctions prévues aux articles 107 à 114 du RGEE (Règlement général des études et des examens) ».



Signé à 17h51 le 31/07/2025
à Wavre (Belgique)

Je tiens tout d'abord à remercier mon promoteur, Bruno Masquelier, pour son aide, sa disponibilité , ses commentaires constructifs et le temps qu'il m'a consacré.

Un grand merci à tous mes amis dont le soutien a été précieux durant la rédaction de ce mémoire et plus encore pendant mon parcours universitaire.

Enfin, je tiens à remercier plus particulièrement mes parents et ma sœur d'avoir été là et de m'avoir aidée et accompagnée tout au long de mon cursus.

Table des matières

1	Introduction	1
2	Revue de la littérature	4
2.1	Pratique sportive et classe sociale	4
2.2	Effets de l'activité physique sur l'obésité et diabète selon le statut socio-économique	7
2.3	Activité physique et statut socio-économique	9
2.4	Hypothèses de recherches	11
3	Données et méthodes	14
3.1	Données	14
3.2	Variable dépendante	15
3.3	Variables indépendantes	17
3.3.1	Statut socio-économique	17
3.3.2	Activité physique	18
3.4	Variables de contrôle	18
3.4.1	Âge	18
3.4.2	Sexe	18
3.4.3	Niveau d'urbanisation	19
3.4.4	Pays d'origine	19
3.4.5	Statut matrimonial	19
3.4.6	Niveau d'études	19
3.4.7	Charge physique liée au travail	19
3.5	Méthodologie	20
4	Résultats	22
4.1	Analyse bivariable entre quintiles de revenus et risques d'obésité ou de diabète . .	24
4.1.1	Test de Kruskal-Wallis : Obésité	24
4.1.2	Test de Kruskal-Wallis : Diabète	25
4.2	Analyses factorielles et clustering	26
4.2.1	Analyse de données mixtes	26
4.2.2	Clustering	32
4.3	Modèles de régressions	38
4.3.1	Modèles de régressions : Obésité	38
4.3.2	Modèles de régressions : Diabète	41
5	Discussion et conclusion	46
6	Bibliographie	49

Liste des graphiques

1	Prévalence de l'obésité par groupes d'âges sur base des données HIS (2018)	16
2	Prévalence du diabète par groupes d'âges sur base des données HIS (2018)	17
3	Prévalence de l'obésité par quintile de revenu (avec intervalles de confiance) sur base des données HIS (2018)	25
4	Prévalence du diabète par quintile de revenu (avec intervalles de confiance) sur base des données HIS (2018)	26
5	Variance expliquée par les différentes dimensions Résultat de l'analyse factorielle de données mixtes avec pondérations	30
6	Visualisation des contributions des variables quantitatives à la FAMD Réalisé à partir des résultats FAMD	31
7	Visualisation des contributions des variables qualitatives à la FAMD Réalisé à partir des résultats FAMD	31
8	Sélection du nombre de clusters via la méthode du coude appliquée au clustering k-means Réalisé à partir des données HIS (2018)	33
9	Positionnement des clusters d'activités physiques dans l'espace factoriel (FAMD) Réalisé sur base du clustering et FAMD	34
10	Prédictions du modèle 4 pour l'obésité Réalisé sur base des données HIS (2018)	40
11	Prédictions du modèle 4 pour le diabète Réalisé sur base des données HIS (2018)	44
12	Contribution des variables aux deux dimensions de l'analyse factorielle mixte sur base des données HIS (2018)	55

Liste des tableaux

1	Statistiques descriptives des caractéristiques de l'échantillon selon le genre.	22
2	Description des variables quantitatives introduites et leurs modalités .	27
3	Description des variables catégorielles introduites et leurs modalités .	28
4	Répartition des individus selon les clusters d'activité physique, par genre. Réalisé sur base des données HIS (2018)	33
5	Répartition (%) de la charge physique liée au travail selon les clusters d'activité physique. Réalisé sur base du clustering.	37
6	Résultat du test de Wilcoxon par paires (p-values ajustées) pour l'obésité	55
7	Résultat du test de Wilcoxon par paires (p-values ajustées) pour le diabète	55
8	Résultat du test de khi carré entre la variable <i>cluster</i> et la variable <i>sexe</i>	56
9	Description statistique des clusters selon les dimensions principales : Dim.1 et Dim.2	56
10	Résultat du test de khi carré entre les clusters et l'activité physique liée au travail	56
11	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon l'activité physique	57
12	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon l'activité physique et la charge de travail.	58
13	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon le revenu Table 13	59
14	Odds ratios, erreurs standards, z, p-values et intervalles de confiance à 95% : Obésité selon l'activité physique et le revenu.	60
15	Prédictions marginales moyennes, intervalles de confiance à 95% et p-values : Obésité selon l'activité physique et le revenu.	62
16	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique	64
17	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique et la charge de travail	65
18	Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon le revenu	66
19	Odds ratios, erreurs standards, z, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique et le revenu.	67

20	Prédictions marginales moyennes, intervalles de confiance à 95% et p-values : Diabète selon l'activité physique et le revenu.	69
----	---	----

1 Introduction

En 2022, une personne sur huit dans le monde était atteinte d'obésité, un chiffre qui a plus que doublé depuis 1990 à l'échelle mondiale (OMS, 2025). Longtemps perçue comme un problème de santé réservé aux pays à revenu élevé, l'obésité touche aujourd'hui tous les pays, y compris ceux à revenu faible ou intermédiaire, devenant un enjeu de santé publique mondial.

L'obésité est un grave problème de santé publique car elle augmente considérablement le risque de maladies chroniques telles que les maladies cardiovasculaires, le diabète de type 2, l'hypertension, les maladies coronariennes et certains cancers (Eurostat, 2024). De la même manière, le diabète constitue un enjeu de santé publique majeur. En 2021, environ 537 millions d'adultes dans le monde vivaient avec un diabète, et ce nombre devrait atteindre 643 millions d'ici 2030 (Magliano et al., 2021). Le diabète de type 2, fortement lié au surpoids et à l'obésité, représente plus de 90% des cas de diabète au niveau mondial.

Ces deux problèmes de santé sont liés à un déséquilibre énergétique. Lorsque nous consommons plus de calories que l'énergie que nous dépensons dans la vie de tous les jours, notre poids corporel augmente (Ritchie & Roser, 2017). Inversement, un surplus de dépense d'énergie par rapport à la prise de calories nous permet de perdre du poids. Ainsi, l'augmentation de l'obésité peut survenir soit parce que les gens mangent plus, soit parce qu'ils bougent moins (Hill & Peters, 1998; Ritchie & Roser, 2017).

Au niveau européen, le diabète est la maladie chronique la plus courante touchant près de 64 millions de personnes (OMS, 2024). On estime qu'un tiers des personnes atteintes ne sont pas diagnostiquées. D'ici 2045, un Européen sur dix pourrait en être atteint avec le vieillissement des populations et l'augmentation des facteurs de risque tels que l'obésité (OMS, 2024).

L'obésité quant à elle a triplé dans les pays d'Europe depuis les années 1980 (OMS, 2025). Les taux les plus élevés de surcharge pondérale et d'obésité se situent dans les pays méditerranéens et d'Europe de l'Est (OMS, 2022). Les personnes les plus exposées à l'obésité sont les femmes (24%) dans la plupart des pays européens (OMS, 2022). Le risque de développer plus des comorbidités augmente également avec l'accroissement du poids. De plus, la prévalence d'obésité varie selon les groupes sociaux et l'obésité est notamment plus élevée chez les personnes ayant un faible niveau d'éducation (OMS, 2022). De même, la prévalence de l'obésité varie selon les groupes sociaux et l'obésité est notamment plus élevée chez les personnes ayant un faible niveau d'éducation (OMS, 2022).

En Belgique, entre 2001 et 2023, la prévalence de diabète chez les individus de 15 ans et plus a doublé en passant de 3,4% à 6,9% (Drieskens et al., 2025). Le diabète est en progression dans les trois régions, mais cette hausse est particulièrement marquée en Wallonie, où l'on observe une augmentation de plus de 2% depuis 2018. On peut aussi noter que les individus sans diplôme de l'enseignement secondaire supérieur ont 2,3 fois plus de risque de souffrir de diabète par rapport aux individus avec un diplôme de l'enseignement supérieur (Drieskens et al., 2025).

Par ailleurs, depuis les premières enquêtes de santé menées par Sciensano en 1997, on constate une tendance linéaire et une hausse significative de la prévalence de l'obésité en Belgique. Celle-ci est passée de 10,8 % en 1997 à 15,9 % chez les adultes en 2018 (Drieskens et al., 2018). La proportion de personnes souffrant d'obésité augmente de manière significative avec l'âge : elle est de 6,4 % chez les 18-24 ans, atteint un pic de 22,8 % chez les 65-74 ans, avant de diminuer à 13,9 % chez les individus âgés de 75 ans et plus (Drieskens et al., 2025). Au niveau régional, une analyse selon le niveau d'instruction met en avant des écarts notables : le pourcentage d'adultes qui souffrent d'obésité est significativement plus faible chez les plus diplômés (10,6% en Flandre, 10,0% à Bruxelles et 15,7% en Wallonie) que chez les moins diplômés, c'est-à-dire les diplômés du secondaire supérieur, les diplômés du secondaire inférieur et les diplômés de l'enseignement primaire et les non-diplômés avec des taux respectifs de 23,1 %, 18,9 % et 21,4 % (Drieskens et al., 2018).

L'augmentation de l'obésité au cours du temps a été notable chez les enfants de parents ayant un statut socio-économique inférieur, ce qui a entraîné une augmentation des disparités socio-économiques absolues (Drieskens et al., 2024). La prévalence de l'obésité infantile est non seulement plus élevée parmi les enfants issus de milieux défavorisés, mais les écarts avec les groupes les plus favorisés se sont accentués au cours des deux dernières décennies. Cette évolution traduit une polarisation sociale croissante face au risque d'obésité dès le plus jeune âge. La littérature nous suggèrent aussi une relation inverse entre niveau d'instruction et obésité. Autrement dit, l'obésité est plus fréquente chez les personnes peu instruites (Drieskens et al., 2018). Le niveau d'instruction est un déterminant de la santé, notamment en raison de son influence sur le statut socioéconomique (revenu, profession), mais aussi à travers la littératie en santé et les comportements sanitaires (Drieskens et al., 2018).

Face à ces enjeux de santé majeurs, l'activité physique apparaît comme un facteur protecteur par rapport à la fois à l'obésité et au diabète. Les effets d'une activité physique sur la réduction de la mortalité et de maladies chroniques comme le diabète et l'obésité sont en effet bien documentés (Rhodes et al., 2017; Warburton & Bredin, 2017). Cependant, il est important de noter que ces bienfaits ne s'appliquent pas forcément à toutes et tous, et de la même manière. Une littérature grandissante met en avant l'existence de profondes inégalités de santé au niveau de

la pratique sportive. Le statut socio-économique, en particulier, apparaît comme un facteur clé pouvant entraver les bienfaits d'une pratique sportive sur la santé. En effet, celui-ci impacte à la fois les préférences sportives, la manière de pratiquer ou même la fréquence d'activité physique.

En Belgique en 2018, 38,5 % des personnes ayant un niveau d'instruction élevé pratiquaient au moins 150 min d'activité physique aérobie modérée à vigoureuse par semaine, contre 18,8 % dans le groupe avec un niveau d'instruction faible, soit une différence absolue de 19,7 ppt (Sciensano, 2022). Les personnes ayant un faible niveau d'instruction étaient deux fois moins susceptibles de pratiquer une activité physique que les personnes ayant un niveau d'instruction élevé (Sciensano, 2022).

Les recherches montrent que selon le groupe auquel les individus appartiennent, les bénéfices en termes de santé qu'ils retirent de l'activité sportive peuvent varier. Ainsi, les groupes socio-économiques les plus aisés ont tendance à avoir accès plus facilement à des infrastructures sportives, et disposent par ailleurs de ressources économiques plus conséquentes qui leur permettent de pratiquer une activité de manière plus régulière (Guthold et al., 2018; Naik et al., 2016). En revanche, les groupes les plus défavorisés ont moins de moyens économiques et ont parfois tendance à percevoir le sport de manière plus négative (Naik et al., 2016). Le niveau d'éducation et l'accès à l'information jouent aussi un rôle sur les représentations que les différents groupes socio-économiques peuvent avoir sur l'importance de pratiquer un sport pour améliorer la santé (Naik et al., 2016; Wilson et al., 2004).

Le type de sport pratiqué produit lui aussi des effets différenciés sur la santé et est intrinsèquement lié à la position sociale. Le football par exemple est un sport très pratiqué par les classes populaires et peu pratiqué par les classes supérieures. Le genre est aussi un facteur notable de différenciation de la pratique sportive. Richard et al. (2023) montrent notamment que les sports de groupes comme le football, le handball, le judo/karaté, la gymnastique et la danse sont inégalement pratiqués en fonction du sexe. Ainsi, les inégalités en termes d'activité physique sont multiples et sont susceptibles d'induire des effets différenciés sur la santé.

Ce mémoire s'inscrit dans ce contexte spécifique d'inégalités de pratiques sportives et d'augmentation de maladies comme l'obésité et le diabète. Il vise à essayer de comprendre comment le statut socio-économique peut moduler les effets d'une pratique sportive sur la santé en Belgique. L'objectif est de déterminer s'il existe des effets différenciés et dans quel sens ils se manifestent sur les différents niveaux de statut socio-économique. Pour ce faire, nous utilisons les données de l'Enquête Santé, collectées et distribuées par l'institut de recherche belge Sciensano.

2 Revue de la littérature

2.1 Pratique sportive et classe sociale

Lorsqu'on analyse l'activité physique chez les individus, il est intéressant de se pencher sur le type d'activité et comment se pratique le sport. Comment expliquer que certains individus pratiquent un sport plutôt qu'un autre ? Que certains se tournent vers le tennis alors que d'autres vers le basket ? La notion de « classe sociale théorisée » par Bourdieu (1984) se révèle très pertinente pour analyser le degré et les modalités de participation des individus aux activités sportives.

Dans ses travaux, Bourdieu (1984) distingue 4 types de capitaux : capital économique, capital social, capital culturel et capital symbolique. Le capital économique fait référence aux ressources financières et au patrimoine d'un individu. Le capital social correspond à ses relations et réseaux sociaux, tandis que le capital culturel englobe les ressources et dispositions culturelles (Jugnot, 2019). Le capital symbolique représente lui un concept plus vague de volume de reconnaissance, de légitimité et de consécration accumulé par un individu au sein d'un champ d'appartenance (Durand, 2014). Ces capitaux, inégalement répartis dans la société, structurent les positions sociales des individus. Les classes sociales regroupent des individus proches en termes de volume et de structure de capital, notamment sur le plan économique et culturel (Claron & Lagarde, 2024).

Selon Bourdieu, la société est divisée en différents champs (qu'il soit religieux, politique, médical ou encore artistique), chacun doté de ses propres règles et hiérarchies. Les différents capitaux dont disposent les individus sont alors plus ou moins valorisés selon le champ (Jugnot, 2019). Dans chacun de ces champs, les individus occupent donc des positions différentes selon leurs capitaux et luttent symboliquement pour améliorer ou conserver leur position, en mobilisant leurs capitaux.

Au-delà des capitaux, les individus sont aussi guidés par les « habitus de classe » . Bourdieu définissait l'habitus comme “ « une loi immanente, déposée en chaque agent par la prime éducation »” (Bourdieu, 1972). L'habitus désigne un système de préférences et de croyances. Ces croyances et préférences sont acquises durant la socialisation de l'individu de manière inconsciente et permettent de s'adapter et d'intégrer un environnement social (Jugnot, 2019). Cet habitus engendre des pratiques sociales ajustées aux positions sociales (Wagner, 2012). Cela explique pourquoi, placés dans des contextes similaires, des individus issus d'un même milieu social auront tendance à partager des représentations et des normes similaires.

Enfin, Bourdieu (1984) distingue un style de vie raffiné pour la classe supérieure, un style de vie prétentieux de la classe moyenne et un style de vie prolétaire pour la classe populaire (Mutz & Mueller, 2021). Le mode de vie de la classe supérieure est dicté par la liberté de choix et la marge de possibilité alors que le style de vie de la classe populaire est dicté par les nécessités (Mutz & Mueller, 2021). En ce qui concerne la classe moyenne, le style de vie se caractérise par une idéologie de distinction : un effort d'imitation partielle des pratiques valorisées par les classes supérieures, notamment à travers des hobbies, des formes de consommation ostentatoire ou encore des loisirs dits « éducatifs » ou « cultivés », sans toutefois en maîtriser pleinement les codes ni disposer des ressources nécessaires (Bennett et al., 2009; Mutz & Mueller, 2021).

C'est dans cette perspective que le sport peut être appréhendé. En tant que champ social à part entière, il participe à la construction et à l'expression des habitus. Les préférences sportives, la manière de pratiquer ou même la fréquence d'activité physique sont étroitement liées aux capitaux détenus et à l'habitus des individus. En d'autres mots, les inégalités sociales de pratiques sportives ne relèvent pas uniquement de contraintes économiques ou de choix individuels, mais s'inscrivent dans une logique plus large de reproduction sociale (Claron & Lagarde, 2024).

À partir de données de cohorte en Suède, Lindström et al. (2001) ont mis en évidence le rôle du capital social dans la relation entre le statut socio-économique et les comportements de santé, notamment l'activité physique de loisir. Les groupes socio-économiques inférieurs auraient un risque plus élevé de se trouver dans le quartile inférieur d'activité physique de loisir par rapport aux autres groupes socio-économiques. Leurs résultats permettent aussi de mettre en avant que l'insuffisance des ressources psychosociales dans certains groupes socio-économiques est un élément important qui explique les différences socio-économiques en matière d'activité physique de loisir (Lindström et al., 2001).

Ces résultats sont cohérents avec d'autres recherches qui indiquent que le niveau socio-économique semble être associé au niveau d'activité physique. Borraccino et al. (2009) à partir des données de l'étude « Health Behaviour in School-aged Children » (HBSC) menée dans 32 pays européens, indiquent que plus le score sur l'échelle d'aisance familiale (qui est un proxy du revenu familial) est élevé, plus l'association avec une activité physique vigoureuse est forte.

Au-delà du simple fait de pratiquer une activité sportive, ce sont aussi le type de sport choisi, le contexte dans lequel il est pratiqué et la manière de le pratiquer qui distinguent les pratiques sportives des milieux aisés de celles des milieux plus défavorisés (Bourdieu, 1979; Claron & Lagarde, 2024; Pociello, 1987).

La distribution des pratiques sportives s'explique en partie par l'influence du milieu social dans la socialisation (Claron & Lagarde, 2024). Un individu prendra goût à une activité physique selon les codes et les normes qu'il aura reçu lors de sa socialisation en grande partie réalisée par sa famille.

Bourdieu (1978) avait établi des critères qui permettaient de distinguer les sports de la classe supérieure de ceux de la classe populaire. Premièrement, le rapport au corps, notamment la dureté et l'intensité, est un élément qui différencie la pratique sportive. Les classes populaires choisissent plus souvent des sports où le corps et la force visible sont privilégiés alors que les classes supérieures privilégient des sports, souvent instrumentés, où l'accent est mis sur l'esthétisme et l'absence de contact direct (Bodin et al., 2004; Teulon, 2024). Lorsque la pratique sportive traite le corps comme un instrument, elle est souvent moins prisee par les classes populaires (Claron & Lagarde, 2024).

Deuxièmement, les pratiques sportives se distinguent également par les valeurs qu'elles véhiculent et les motivations sous-jacentes. Les disciplines prisées par les classes aisées mettent en avant des vertus comme la santé, la rigueur, la discipline de soi et une forme d'ascétisme, valorisant ainsi le caractère (Stempel, 2005). À l'inverse, les sports populaires tendent à exalter des qualités plus directement corporelles, telles que la puissance ou l'endurance (Bodin et al., 2004; Mutz & Mueller, 2021; Stempel, 2005).

Troisièmement, l'environnement dans lequel s'exercent ces activités reflète lui aussi des clivages sociaux : les sports de prestige se pratiquent souvent dans des lieux coûteux, isolés ou sélectifs (comme le golf ou le ski alpin), alors que les disciplines populaires se déroulent dans des espaces facilement accessibles, comme les terrains de football de quartier (Claron & Lagarde, 2024; Mutz & Mueller, 2021). Certaines pratiques sportives jugées populaires comme le football ou le bodybuilding sont alors évitées ou délaissées par les classes supérieures (Claron & Lagarde, 2024).

Ces distinctions ne se réduisent donc pas à des questions économiques. Elles renvoient à des logiques symboliques plus profondes : les goûts en matière de sport sont façonnés par des formes de capital (économique, culturel, social) et participent à la construction de styles de vie socialement situés.

Richard et al. (2023) mettent en avant des inégalités de pratiques pour les sports de raquette et les sports spéciaux par rapport aux sports collectifs. Ils montrent de manière plus précise que les personnes ayant un haut niveau d'éducation et de revenu sont 4,25 fois plus susceptibles de participer à des sports comme le tennis ou le badminton que les personnes avec un niveau

d'éducation bas. Le football quant à lui semblait plus pratiqué par les classes sociales populaires. Ces résultats sont cohérents avec les autres études qui montrent que le football fait partie des sports les plus inclusifs et les sports de raquette, les activités de nature et les sports de fitness font partie d'activité plus exclusives (Breuer et al., 2010; Haut, 2018; Mutz & Mueller, 2021; Stempel, 2005; Thibaut et al., 2017). Mutz and Mueller (2021) soulignent également que, parallèlement aux sports traditionnels, des pratiques centrées sur le corps et l'esprit (comme le yoga), ainsi que des activités en lien avec la nature (telles que la randonnée), s'inscrivent aussi dans des logiques socialement structurées que les classes supérieures privilégient davantage.

Comme nous l'avons relevé précédemment, le genre est aussi un facteur notable de la pratique sportive. Richard et al. (2023) montrent notamment que les sports de groupes comme le football, le handball, le judo/karaté, la gymnastique et la danse sont inégalement pratiqués en fonction du sexe. Les femmes pratiquent plus souvent la gymnastique, la danse et les sports thérapeutiques et esthétiques (Mutz & Mueller, 2021) que les hommes qui pratiquent davantage le football et le handball. On peut donc dire que le capital culturel et le genre structurent la participation dans divers domaines classiques du sport mais avec des degrés différents.

Pour conclure, les différences observées dans la pratique sportive, qu'il s'agisse du type d'activité, des lieux de pratique, des motivations ou des modalités d'engagement, s'inscrivent dans des logiques de classe, de disparités de capital et de recherche de distinction. Le rapport au corps, au temps, à l'effort, à la compétition ou au loisir constitue autant de facteurs structurants qui influencent les préférences sportives et contribuent à différencier les individus. L'apport de Bourdieu permet de comprendre que les inégalités dans les pratiques sportives ne relèvent pas uniquement de contraintes économiques ou de choix individuels. Elles s'inscrivent dans une dynamique plus large de reproduction sociale, où le sport devient un révélateur des rapports de domination et un vecteur de distinction entre les classes sociales.

2.2 Effets de l'activité physique sur l'obésité et diabète selon le statut socio-économique

La pratique d'une activité physique est significativement associée à une baisse de la mortalité, du diabète de type 2 et de l'obésité (W. I. J. de Boer et al., 2023). Des données récentes confirment une relation curvilinéaire entre activité physique et diabète : des volumes plus élevés d'activité physique sont associés à une réduction du risque de diabète, bien que cette diminution tende à s'atténuer à des niveaux très élevés d'activité (DiPietro et al., 2019). D'autres études montrent qu'un niveau élevé d'activité physique est associé à des indicateurs d'adiposité plus favorables ainsi qu'à une limitation de la prise de poids chez l'adulte (DiPietro et al., 2019).

Cependant, les effets sur la santé varient selon les types de sport, les sports d'équipe étant souvent plus efficaces que les autres types de sport, tandis que l'effet des activités de remise en forme est nettement moins fort que celui des autres types de sport (W. I. J. de Boer et al., 2023). Ces résultats sont en contradiction avec d'autres études qui postulent que toute forme d'activité physique, quel que soit le type, peut avoir des effets bénéfiques sur la santé, pourvu qu'elle soit régulière et adaptée à l'individu (Warburton et al., 2006).

En termes de fréquence et d'intensité, il est prouvé qu'une plus grande activité physique est associée à de meilleurs résultats en matière de santé. Toutefois, ces bénéfices tendent à se stabiliser au-delà d'un certain seuil d'activité. À ce jour, les données disponibles ne permettent pas de déterminer précisément le niveau à partir duquel les effets positifs cessent d'augmenter de manière significative (OMS, 2024) ; il n'y a pas de réel consensus à ce propos (Oja et al., 2015) et ce seuil dépend sans doute des individus.

Il est également important de noter que les effets sur la santé dépendent du niveau d'activité physique pratiqué. Du côté des sports d'équipe, le football récréatif permet d'améliorer la condition aérobique, la fonction cardiovasculaire au repos, et de réduire la masse grasseuse. Une pratique régulière du rugby sur 8 semaines améliore la puissance sous-maximale, réduit la masse grasseuse et augmente la masse maigre (Mendham et al., 2014).

Concernant les sports individuels, la course à pied améliore la capacité aérobie ainsi que la fonction cardiovasculaire au repos (Oja et al., 2015). La natation semble réduire le risque de mortalité globale et l'adiposité, bien que les résultats soient moins concluants. En revanche, aucune amélioration significative n'a été observée pour l'équitation ou le cyclisme récréatif (Oja et al., 2015).

Indépendamment de ces activités physiques liées aux loisirs, on fait souvent la distinction entre l'activité physique liée au travail et celle liée aux déplacements quotidiens. Beenackers et al. (2012) ont mis en évidence que les personnes ayant un revenu plus faible pratiquent moins d'activité physique pendant leur temps libre (comme le sport ou la marche de loisir), en particulier lorsqu'elle est d'intensité vigoureuse. En revanche, elles sont davantage engagées dans une activité physique professionnelle, souvent intense (bâtiment, entretien, etc.).

Ce schéma se retrouve dans l'ensemble de l'Europe, illustrant des inégalités marquées dans les modalités de pratique de l'activité physique. Dès lors, les auteurs soulignent l'importance de

ne pas se limiter à une approche globale de l'activité physique (cumulant travail et loisir), car cela peut masquer des disparités socioéconomiques significatives. Il faudrait se concentrer sur l'augmentation de l'activité physique de loisir chez les personnes avec un statut socio-économique bas pour améliorer la santé et réduire les inégalités tout en tenant compte de la charge physique engagée sur le lieu de travail pour ces profils (Beenackers et al., 2012).

L'OMS recommande aux adultes âgés de 18 ans et plus de pratiquer au minimum 150 minutes d'activité physique d'intensité modérée par semaine, ou au moins 75 minutes d'activité physique d'intensité élevée. L'OMS réaffirme le fait que par rapport aux niveaux les plus bas d'activité physique, tous les niveaux et toutes les intensités (y compris l'intensité légère) d'activité physique sont associés à un risque de mortalité plus faible (Ekelund, Tarp, et al., 2019).

L'OMS recommande par ailleurs de limiter les comportements sédentaires, définis comme le temps passé en position assise ou allongée avec une dépense énergétique très faible. En effet, ce type de comportement est associé à des issues de santé défavorables, notamment une augmentation de la mortalité toutes causes confondues, de la mortalité liée aux maladies cardiovasculaires et au cancer, ainsi qu'à une incidence accrue des maladies cardiovasculaires, de certains cancers et du diabète de type 2 (OMS, 2024).

L'augmentation de comportements de type sédentaires a stimulé la recherche sur l'interaction potentielle entre les différents niveaux d'activité physique et les niveaux de comportement sédentaire. Les résultats globaux montrent que l'effet des comportements sédentaires est plus fort chez les personnes qui pratiquent peu d'activités physiques d'intensité modérée à vigoureuse. Inversement une pratique et une fréquence plus importante d'intensité modérée à vigoureuse peuvent atténuer les résultats défavorables pour la santé associés à des comportements sédentaires plus élevés (Ekelund, Brown, et al., 2019; Ekelund et al., 2016). En d'autres termes, pratiquer régulièrement une activité physique d'intensité modérée à élevée peut compenser, ou réduire, les risques liés à un comportement sédentaire important.

2.3 Activité physique et statut socio-économique

W. I. J. de Boer et al. (2023) mettent en avant le fait qu'au sein de chaque groupe socio-économique, la pratique du sport est associée à de meilleurs résultats en matière de santé. Leur étude, menée aux Pays-Bas à partir des données de la cohorte « Lifelines », une vaste étude longitudinale portant sur plus de 160 000 personnes, montre que l'ampleur de ces bénéfices ne varie pas selon les groupes socio-économiques ; pratiquer une activité physique est aussi bénéfique pour la santé d'une personne à faible revenu que pour une personne à revenu élevé.

Cependant, de nombreuses études ont démontré qu'en termes d'intensité, les groupes socio-économiques inférieurs sont moins susceptibles de pratiquer une activité physique que les groupes socio-économiques élevés (Beenackers et al., 2012; Borraccino et al., 2009; Lindström et al., 2001; Richard et al., 2023). Par exemple en Suède dans le comté de Skåne, le fait de vivre dans des zones socialement défavorisées multiplie par deux le risque pour les adolescents de pratiquer peu d'activités physiques et sportives (Lindström et al., 2001). En Belgique, on estime que les personnes ayant un faible niveau d'instruction sont deux fois moins susceptibles de pratiquer une activité physique que les personnes ayant un niveau d'instruction élevé (Sciensano, 2022).

Ces écarts de pratiques peuvent s'expliquer par différents facteurs. Kamphuis et al. (2008) soulignent que les personnes avec un statut socio-économique bas ont plus de probabilités de signaler que leurs quartiers sont peu sûrs, peu attrayants et ne comportent pas suffisamment d'espaces pour l'activité physique. Le manque de ressources et l'environnement dans lequel l'individu évolue semblent être des déterminants majeurs de l'engagement dans l'activité physique. En effet, les personnes privilégiées peuvent disposer de plus de ressources psychosociales, financières et de voisinage, ainsi que d'une meilleure connaissance de la santé pour adopter des comportements sains (Beenackers et al., 2012). Les facteurs individuels entrent aussi en compte pour expliquer les niveaux d'activité physique. Ce facteur regroupe principalement les croyances négatives sur le sport et la faible confiance en ses capacités. Les croyances et cognitions défavorables sont plus fréquentes chez les personnes ayant un statut socio-économique plus bas (Galobardes et al., 2003; Kamphuis et al., 2008).

Selon Chinn et al. (1999) les obstacles aux activités physiques peuvent être « internes » ou « externes ». Les obstacles internes désignent un manque de motivation ou de temps libre qui sont plus fréquents au sein des groupes socio-économiques supérieurs. En revanche, les obstacles externes tels que le manque d'argent, de moyen de transport, d'accessibilité ou encore la maladie sont plus fréquents chez individus appartenant aux groupes socio-économiques inférieurs. Richard et al. (2023) confirment ces résultats en précisant que les personnes avec un statut socio-économique plus faible font face à des obstacles spécifiques tels que la fatigue, des limitations liées à leurs santé ou encore des contraintes économiques. En revanche, ce qui pousse ces individus à maintenir leur activité physique, c'est avant le plaisir qu'ils en retirent, ou encore la dimension sociale, les interactions et le soutien reçu.

Lindström et al. (2001) ont mis en avant une influence du capital social dans le lien entre le statut socio-économique et le comportement de santé. Les auteurs montrent que l'inclusion de la participation sociale diminue les effets de la relation entre statut socio-économique bas et par-

ticipation faible aux activités physiques. Ces résultats montrent que l’insuffisance des ressources psychosociales dans certains groupes socio-économiques est un élément important qui explique les différences socio-économiques en matière d’activité physique de loisir (Lindström et al., 2001). Kamphuis et al. (2008) montrent qu’un climat de petit réseau social et de faible cohésion sociale augmentent les probabilités de ne pas faire de sport.

Enfin, le fait de ne pas rencontrer les recommandations en termes d’activité physique ne coïncide pas forcément avec une augmentation d’activités sédentaires (comme regarder la tv, jouer à des jeux vidéo, ...) comme on pourrait le penser. Cela suggère que ces comportements ne sont pas en réelle compétition. D’autres études montrent que les niveaux d’activité physique et les comportements sédentaires dépendent plus de l’environnement dans lequel les personnes évoluent (voir Biddle et al. (2004) pour une recherche sur ce thème au sein d’adolescents). Cela renforce l’idée que le capital social joue un rôle important dans les trajectoires d’activité physique des individus plus défavorisés.

2.4 Hypothèses de recherches

La littérature qui précède suggère que les individus issus de milieux défavorisés, souvent moins actifs physiquement, présentent une vulnérabilité accrue aux maladies chroniques comparativement à ceux bénéficiant d’un statut socio-économique plus élevé. Pour autant, la pratique d’une activité physique protège contre la mortalité et le risque de maladies chroniques associées quelle que soit la position socio-économique (Schnohr et al., 2004). W. I. J. de Boer et al. (2023) ont montré que la pratique sportive quel que soit le niveau d’éducation, était associée à une baisse du risque d’obésité (HR 0.77) et de diabète (HR 0.70). Au niveau de la pratique d’une activité physique, on a vu qu’en moyenne les groupes avec un statut socio-économique plus élevé sont plus engagés que les groupes de niveau socio-économique plus bas. Il est prouvé qu’un faible niveau d’éducation est associé à une inactivité physique (Nielsen et al., 2006). Les groupes socio-économiques les plus bas ont tendance à présenter des comportements sédentaires plus fréquents, ce qui peut exacerber le risque d’obésité (Hanson & Chen, 2007).

Ces différents éléments de la littérature nous amènent à formuler les hypothèses suivantes :

Hypothèse 1 : En Belgique, les individus disposant d’un statut socio-économique inférieur pratiquent en moyenne moins d’activités physiques de loisir, ce qui augmente leurs risques d’être atteints d’obésité et de diabète.

On a déjà indiqué précédemment que le manque de ressources et l’environnement dans lequel

l'individu évolue semblent être des déterminants majeurs de l'engagement d'une activité physique. Par exemple, vivre dans des zones socialement défavorisées multiplie par deux le risque pour les adolescents de pratiquer peu d'activités physiques et sportives (Lindström et al., 2001). Une étude menée en Belgique nuance ces résultats et permet de formuler une seconde hypothèse. En effet cette étude, montre qu'habiter dans un quartier à forte « marchabilité » augmente l'activité physique modérée à vigoureuse et les déplacements actifs (par marche ou vélos). De façon intéressante, les quartiers à plus faible statut socio-économique sont également associés à plus de déplacements à vélo et à moins d'utilisation de moyens de transport motorisés (Van Dyck et al., 2010) Aucun effet d'interaction entre la marchabilité du quartier et son niveau socio-économique n'a été observé, ce qui suggère que les bénéfices liés à un environnement propice à la marche sont indépendants du statut socio-économique du quartier.

Hypothèse 2 : Les individus, qui pratiquent des déplacements actifs (à vélo ou en marchant) à raison de 30 minutes par jour, ont moins de probabilités d'être atteints de diabète ou d'obésité quel que soit leur statut socio-économique.

Notre troisième hypothèse concerne l'activité physique dans un cadre professionnel. Une étude a montré que les individus ayant une dépense énergétique élevée liée à leur travail étaient plus à risque d'être inactifs pendant leur temps libre (Kaleta & Jegier, 2005). Pour les hommes qui dépensent plus de 4000 kcal par semaine au travail, le risque d'inactivité physique de loisir était 1,5 fois plus élevé que chez ceux dont la dépense était inférieure à ce seuil. Un résultat similaire a été observé chez les femmes à partir de 3500 kcal/semaine. Cela suggère que la charge physique d'un travail peut limiter l'engagement dans d'autres activités physiques de loisirs.

Dans le même ordre d'idée, combiner activité physique de loisir, activité physique liée au travail et activité physique liée au déplacement permet d'augmenter l'effet protecteur total et de réduire la mortalité totale et cardiovasculaire chez les personnes atteintes de diabète de type 2 (G. Hu et al., 2004). L'activité physique liée seulement au travail, même modérée ou vigoureuse, est indépendamment liée à une baisse du risque de mortalité chez les diabétiques (F. B. Hu et al., 2004). On peut aussi noter que cette activité physique professionnelle est souvent plus élevée chez les personnes ayant des métiers manuels (Gudnadottir et al., 2019). Notre troisième hypothèse peut donc être formulée comme suit :

Hypothèse 3 : L'association entre la fréquence de l'activité sportive de loisir et l'incidence du diabète/de l'obésité est moins forte chez les individus soumis à un plus fort engagement physique dans le cadre de leur emploi, quel que soit leur statut socio-économique.

Enfin, on a vu comment la classe sociale pouvait influencer le choix de l'activité physique via divers mécanismes (détention de capitaux, ...) et quels étaient les bienfaits de ces activités sur la santé. Il est donc intéressant de se demander comment ces effets peuvent varier selon le statut socio-économique. Pan et al. (2020) montrent qu'un niveau socio-économique faible est un facteur de risque important d'obésité, surtout chez les femmes, mais que l'activité physique peut atténuer cet effet négatif. Dans le même ordre d'idée, Warburton et al. (2006), mettent en avant que les plus grandes améliorations en matière de santé sont observées lorsque les personnes les moins en forme deviennent physiquement plus actives.

Hypothèse 4 : L'impact positif de l'activité physique de loisir sur l'obésité et le diabète est plus marqué parmi les individus de statut socio-économique faible que chez les individus avec un statut socio-économique élevé.

3 Données et méthodes

3.1 Données

Les données utilisées dans le cadre de ce mémoire proviennent de *l'Enquête de Santé* de 2018 de Sciensano (Drieskens et al., 2018). Sciensano est l'institut belge de recherche axé sur la santé humaine et animale. Les équipes de Sciensano réalisent régulièrement des enquêtes de santé abordant différents thèmes comme les maladies chroniques, la santé mentale, le statut nutritionnel, l'état de santé perçu, la qualité de vie, La première enquête de santé a été lancée en 1997 au niveau national. Depuis cette première enquête, une nouvelle enquête a lieu tous les 4 à 5 ans (Drieskens et al., 2025).

L'enquête de santé, aussi appelée HIS, est une enquête de type transversale dont l'échantillon compte environ 10 000 participants. Ces derniers sont sélectionnés selon un plan d'échantillonnage stratifié en grappes à plusieurs degrés. La stratification repose notamment sur des critères géographiques (province), sociodémographiques (âge, sexe, taille du ménage) et temporels (trimestre de l'année), afin d'assurer une meilleure représentativité de l'échantillon par rapport à la population belge.

La collecte de données se déroule sur une année civile et est menée par CAPI (Computer-Assisted Personal Interviewing). En complément de cet entretien, un questionnaire individuel auto-administré est proposé aux participants sélectionnés, contenant des questions plus sensibles ou personnelles.

Les données de l'enquête comprennent trois sources d'informations différentes : des données récoltées par enquête, des données basées sur un entretien CAPI et des données complétées sur la base d'un auto-questionnaire. Par exemple, certains modules comme la perception de la santé ou encore la pratique d'une activité physique proviennent uniquement du questionnaire d'auto-déclaration.

Dans le cadre de ce mémoire, une requête de données a été formulée auprès de Sciensano et la base de données analysée comprend 83 variables différentes et 12 742 observations. Différents modules sont abordés comme les informations socio-démographiques, l'éducation, les maladies chroniques, le statut nutritionnel et l'activité physique. La population cible comprend les individus âgés de 18 à 64 ans qui ont complété le questionnaire individuel lequel reprend notamment les questions relatives à la pratique sportive. Sur le total de 12 742 observations à la base, nous conservons donc 5 826 observations.

Pour tenir compte des probabilités de sélection inégales et pour garantir que les estimations produites soient représentatives de la population totale, des poids d'échantillonnage sont attribués à chaque individu. Ces poids tiennent compte à la fois de la probabilité de sélection dans le ménage et d'un facteur de post-stratification ajusté selon la province, l'âge, le sexe, la taille du ménage et le trimestre de participation.

Ces poids, fournis dans la base de données, ont été intégrés dans l'ensemble des analyses descriptives et inférentielles. Plus précisément, la variable de pondération `wfin`, la variable de stratification `provw` (province) et la variable de regroupement `hh_cluster_e1` (ménage) ont été spécifiées à l'aide de la fonction `svydesign()` du package *survey* du logiciel R, afin de respecter la structure du plan d'échantillonnage.

3.2 Variable dépendante

L'analyse porte sur deux variables dépendantes relatives aux affections suivantes : auto-diagnostique d'obésité et de diabète.

Obésité

Pour mesurer l'obésité, la variable « `NS_4` » indique si un adulte est en situation d'obésité. Elle comprend trois modalités : oui, non et pas de réponse. Cette variable est dérivée de la classification de l'indice de masse corporelle (IMC), qui distingue quatre catégories : insuffisance pondérale, poids normal, surpoids et obésité. L'IMC est calculé à partir du poids et de la taille autodéclarés par les répondants. À partir de cette mesure, une catégorisation est effectuée selon les seuils standard définis par l'OMS. La modalité « oui » dans la variable `NS_4` correspond aux individus classés dans la catégorie obésité sur base de leur IMC (plus de 30 d'IMC).

Après recodage et en tenant compte du plan d'échantillonnage de Sciensano, le graphique ci-dessous présente la prévalence d'obésité selon l'âge.

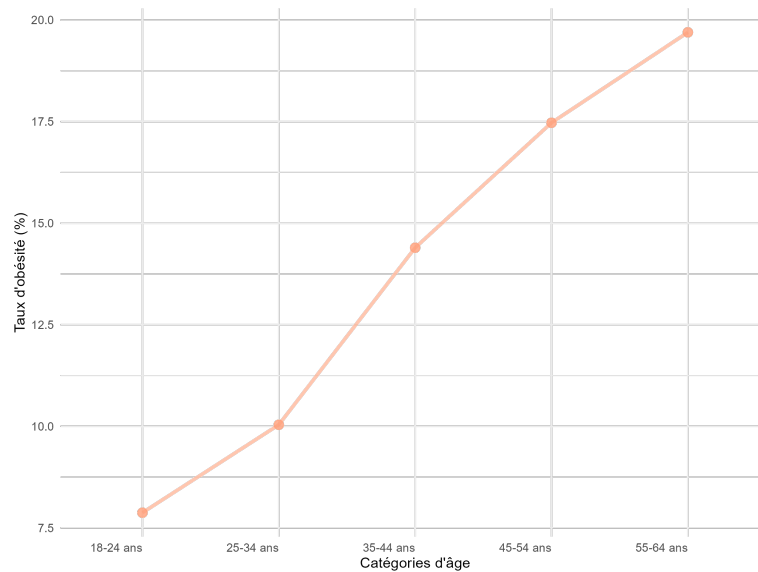


Figure 1 Prévalence de l'obésité par groupes d'âges
sur base des données HIS (2018)

Sur base de ce graphique¹, on voit une tendance linéaire d'augmentation de l'obésité avec l'âge des répondants jusqu'à 64 ans, comme attendu. Le pourcentage passe de 7,87% chez les 18-24 ans à 19.69% chez les 55-64 ans. Cette progression suggère un phénomène cumulatif lié à des facteurs biologiques, comportementaux ou environnementaux qui s'accroissent au fil du temps.

Diabète

Pour mesurer la prévalence du diabète à partir de l'HIS, on peut utiliser la variable « ma0114 » qui demande aux répondants s'ils ont souffert de diabète durant les 12 derniers mois. Cette variable est de type binomial. D'autres questions ont été posées en lien avec cette première variable pour mieux comprendre les profils de santé. Un recodage a été nécessaire car il y avait certaines valeurs négatives qui correspondent à des non-réponses.

¹Les répondants ayant omis cette question (n = 50) ont été exclus de l'analyse présentée.

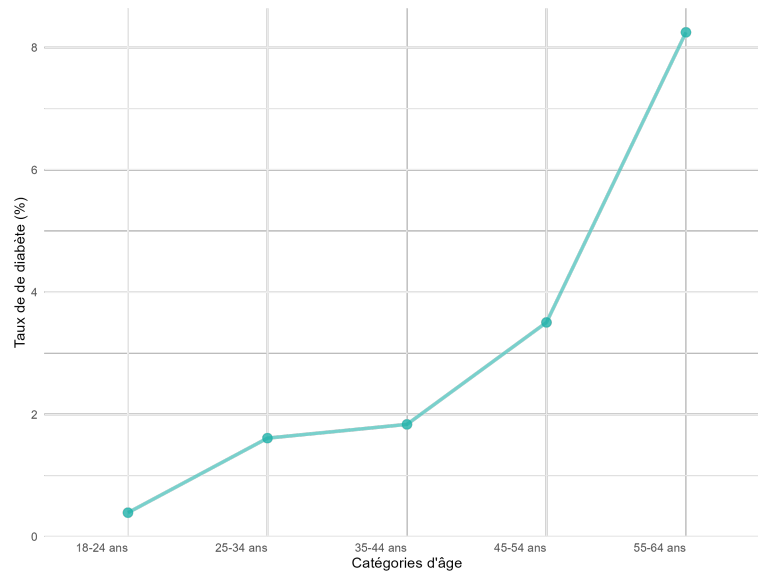


Figure 2 Prévalence du diabète par groupes d'âges
sur base des données HIS (2018)

On constate que la prévalence augmente là aussi nettement avec l'âge, passant de seulement 0,39 % chez les 18-24 ans à 8,24 % dans la tranche 55-64 ans. Cette tendance confirme que le diabète est plus fréquent chez les populations plus âgées.

3.3 Variables indépendantes

Les variables indépendantes sont le statut socio-économique et l'activité physique. Lors des analyses, ces variables vont permettre d'établir s'il existe un lien entre statut socio-économique et bienfait de l'activité physique sur la prévalence d'obésité et de diabète.

3.3.1 Statut socio-économique

La variable indépendante principale est le statut socio-économique, mesuré à l'aide de la variable « IN_1 » disponible dans les données de Sciensano. Cette variable classe les ménages en cinq quintiles de revenu équivalent, construits à partir de la distribution des revenus ajustés pour la taille du ménage (« IN02_1 ») au niveau de la population belge. Le revenu équivalent permet une comparaison standardisée des ressources économiques entre ménages de tailles différentes (Drieskens et al., 2018). Les quintiles vont du premier quintile (revenus les plus faibles) au cinquième quintile (revenus les plus élevés). Cette variable permet d'examiner dans quelle mesure le statut socio-économique influence les effets de la pratique d'une activité physique sur la santé.

3.3.2 Activité physique

L'activité physique est aussi une variable indépendante. Plusieurs variables dans la base de données, reprises lors de l'analyse en composante principale, permettent de classer le temps passé à exercer une activité physique. Cinq groupes ont été créés pour décrire la pratique d'une activité physique (voir analyse factorielle dans résultats). Le cluster 1 décrit les sédentaires et peu actifs. Le cluster 2 reprend les actifs modérés, le cluster 3 les sportifs irréguliers, le cluster 4 les sportifs complets et le dernier cluster reprend les sportifs intensifs.

3.4 Variables de contrôle

Les variables de contrôles servent dans les modèles d'analyses à tenir compte des facteurs de confusion potentiels. L'inclusion de variables de contrôles sert à augmenter la précision des modèles de prédiction en isolant la relation entre variables dépendante et indépendantes pour s'assurer que la relation observée n'est pas due à d'autres facteurs non observés.

3.4.1 Âge

L'âge est intégré comme variable de contrôle car l'impact de l'activité physique sur la santé varie selon l'âge. On a vu qu'effectivement la prévalence de diabète et d'obésité augmentait avec l'âge. De la même façon, l'âge peut favoriser l'apparition d'autres maladies. Par exemple, chez les personnes âgées on verra plus certains types d'affections comme les maladies chroniques ou dégénératives que chez les jeunes, où on verra plus souvent des maladies infectieuses ou génétiques. C'est pourquoi l'âge intégré comme variable de contrôle va permettre d'isoler l'effet de l'âge sur la relation entre la variable indépendante et la variable dépendante. Plusieurs variables présentes dans la base de données décrivent l'âge. Pour cette analyse on va utiliser la variable « age8 » qui découpe l'âge en groupe de 15 ans.

3.4.2 Sexe

Le sexe est également intégré comme variable de contrôle. Les hommes et les femmes sont différents d'un point de vue physiologique, métabolique et hormonal, ce qui influence la relation entre activité physique et santé. La variable « hc04 » reprend le sexe des répondants et est codée comme 1 pour les hommes et 2 pour les femmes.

3.4.3 Niveau d'urbanisation

Le niveau d'urbanisation est une variable pertinente à ajouter comme variable de contrôle dans le cadre de cette étude. Dans la base de données la variable « Urb » reprend 4 catégories différentes d'urbanisation : grandes villes, banlieues, communes urbanisées et rurales. Cette variable sera donc intégrée comme une variable de contrôle afin d'isoler son effet et de mieux comprendre la relation entre la pratique sportive et le statut socio-économique.

3.4.4 Pays d'origine

Il est intéressant d'intégrer le pays d'origine comme variable de contrôle car l'origine d'un individu peut influencer à la fois la santé et la pratique d'une activité physique. Dans la base de données, la variable « HC07_1 » reprend le pays d'origine de l'individu. Trois modalités sont possibles : Belgique, UE (hors Belgique) et hors UE.

3.4.5 Statut matrimonial

Le statut matrimonial est aussi une variable qui a été intégrée comme variable de contrôle. Cette variable reprend 4 catégories différentes : célibataires, mariés, divorcés et veufs. Dans la base de données, la variable « hc05 » décrit le statut matrimonial de chaque individu. Une nouvelle variable appelée « matri » a été créée sur cette base et reprend les 4 niveaux différents.

3.4.6 Niveau d'études

La dernière variable de contrôle intégrée aux modèles de régressions est le niveau d'études atteint. La variable diplôme est créée sur la base de la variable « ET_3 » qui décrit le plus haut niveau d'étude atteint. La nouvelle variable va reprendre 5 niveaux différents : « pas de diplôme », « secondaire inférieure », « secondaire supérieure », « supérieure et plus » et « autres ». La catégorie « autres » reprend surtout les catégories des jeunes qui sont encore aux études et qui n'ont pas encore terminé leurs études.

3.4.7 Charge physique liée au travail

Dans les modèles de régressions, la variable pa01 sera intégrée car elle peut capter un effet résiduel ou complémentaire que les clusters d'activité physique ne saisissent pas complètement. Cette variable décrit la charge physique liée au travail. Quatre modalités sont possibles : « Pas de tâche spécifique », « Souvent assis ou debout », « Souvent marche ou tâche d'activité physique modérée » et « Travail dur et demandant ». La manière dont cette variable est intégrée au modèle

sera expliquée plus en détail dans la partie résultat.

3.5 Méthodologie

Pour cette recherche, toutes les analyses et tests ont été réalisés en utilisant le logiciel R. Dans un premier temps, des analyses descriptives ont été réalisées pour présenter les différentes variables utilisées lors des analyses. Une analyse factorielle et un clustering ont été utilisés pour créer des groupes d'activité physique de loisir. Une analyse ANOVA a ensuite été réalisée pour établir si les différences de prévalence d'obésité et de diabète sont statistiquement significatives parmi les différentes catégories de revenus. Différents modèles ont été exécutés afin de capturer l'effet du statut socio-économique sur la relation entre l'activité physique et la prévalence d'obésité et de diabète.

Le choix des modèles d'analyse pour chaque maladie s'est avant tout effectué sur la base du type de variable. Les variables dépendantes sont toutes de type binomial et prennent soit la valeur « oui » ou « non » indiquant l'incidence de l'événement.

Étant donné que les variables sont binomiales, la régression logistique sera utilisée. Pour rappel, une variable binomiale décrit l'occurrence et la non-occurrence d'un événement (Harrell, 2015). La régression permet à la fois d'analyser dans quelle proportion les variables explicatives servent à expliquer le phénomène étudié et peut aussi servir à des fins prédictives (Legrand & Borjes, 2006). Un des avantages de la régression est la facilité d'interprétation des coefficients sous la forme d' *odds ratio* qui indiquent le rapport de probabilité de l'événement étudié (Legrand & Borjes, 2006).

Ce choix est également motivé par le type de données utilisées. Les données d'enquête de Sci-ensano sont de type transversal, tandis que la littérature scientifique privilégie généralement l'utilisation de modèles de régression de Cox, qui reposent sur des données de cohorte ou des données biographiques où la dimension temporelle et la notion de survie jouent un rôle central. Le modèle de Cox est souvent utilisé en épidémiologie et permet d'examiner l'impact de différentes variables explicatives sur le risque de survenue d'un événement au cours du temps. La plupart des résultats de la littérature scientifique utilisent donc ce modèle en se basant sur des données de cohorte où un suivi des participants est effectué sur plusieurs années (W. de Boer et al., 2021; W. I. J. de Boer et al., 2023; Paffenbarger et al., 1993; Sesso et al., 1998; Thurston et al., 2005). Étant donné que nos données sont de type transversal, sans information sur le temps écoulé depuis le changement d'état, la régression logistique a été privilégiée dans le cadre de ce mémoire.

4 Résultats

Le tableau ci-dessous reprend les différentes variables utilisées lors de cette recherche.

Table 1 Statistiques descriptives des caractéristiques de l'échantillon selon le genre.

Characteristic	Overall N = 5,826	Homme N = 2,873	Femme N = 2,953
Auto déclaration d'obésité			
Oui	887 (15%)	463 (16%)	423 (15%)
Non	4,892 (85%)	2,399 (84%)	2,493 (85%)
Auto déclaration de diabète			
Oui	207 (3.6%)	103 (3.6%)	104 (3.5%)
Non	5,614 (96%)	2,765 (96%)	2,849 (96%)
Niveau de revenu			
1 ^{er} quintile de revenu	1,059 (20%)	491 (18%)	568 (21%)
2 ^e quintile de revenu	741 (14%)	318 (12%)	423 (16%)
3 ^e quintile de revenu	970 (18%)	467 (17%)	503 (18%)
4 ^e quintile de revenu	1,318 (24%)	700 (26%)	618 (23%)
5 ^e quintile de revenu	1,309 (24%)	699 (26%)	611 (22%)
Pa01 : Charge de travail			
Souvent assis ou debout	2,940 (51%)	1,432 (50%)	1,507 (51%)
Souvent marche ou tâche physique modérée	1,704 (29%)	788 (27%)	916 (31%)
Travail dur et demandant	623 (11%)	404 (14%)	219 (7.4%)
Pas de tâche spécifique	553 (9%)	244 (8.5%)	309 (10%)
Catégories d'âges			
15–24	605 (10%)	318 (11%)	287 (9.7%)
25–34	1,186 (20%)	569 (20%)	617 (21%)
35–44	1,264 (22%)	629 (22%)	635 (21%)
45–54	1,412 (24%)	701 (24%)	711 (24%)
55–64	1,360 (23%)	656 (23%)	704 (24%)
Milieu de résidence			
Grandes villes	2,217 (38%)	1,081 (38%)	1,136 (38%)
Banlieue	1,053 (18%)	532 (19%)	521 (18%)
Villes urbanisées	1,887 (32%)	922 (32%)	964 (33%)
Rural	669 (11%)	338 (12%)	331 (11%)

Characteristic	Overall N = 5,826	Homme N = 2,873	Femme N = 2,953
Niveau de diplôme			
Autres	425 (7.4%)	215 (7.6%)	211 (7.3%)
Primaire	222 (3.9%)	97 (3.4%)	125 (4.3%)
Secondaire	692 (12%)	366 (13%)	325 (11%)
Supérieur court	1,947 (34%)	1,040 (37%)	907 (31%)
Supérieur long	2,443 (43%)	1,116 (39%)	1,327 (46%)
Origine			
Belgique	4,812 (83%)	2,385 (83%)	2,427 (82%)
UE	365 (6.3%)	180 (6.3%)	186 (6.3%)
Hors UE	646 (11%)	306 (11%)	341 (12%)
Statut matrimonial			
Célibataire	1,852 (32%)	978 (34%)	874 (30%)
Marié ou cohabitant légal	3,252 (56%)	1,591 (55%)	1,661 (56%)
Veuf	102 (1.8%)	35 (1.2%)	67 (2.3%)
Divorcé	620 (11%)	268 (9.3%)	351 (12%)

Ce tableau reprend les différentes variables utilisées lors des modèles d'analyses. Le tableau montre une répartition équilibrée entre les hommes (2873) et les femmes (2953). Par rapport à l'auto-déclaration d'obésité, on ne voit pas beaucoup de différences marquées entre les hommes (15%) et les femmes (16%). Pour le diabète, la prévalence reste faible par rapport à l'obésité et comparable entre les sexes (environ 3.6%). La distribution des revenus montre que les hommes sont légèrement plus représentés dans les quintiles plus élevées (4 et 5), alors que les femmes sont plus présentes dans les quintiles les plus bas (21% et 16% pour les deux premiers quintiles).

Pour la charge physique liée au travail, la majorité des répondants déclarent avoir un travail sédentaire ou principalement assis/debout (51%), avec une répartition similaire entre hommes (50%) et femmes (51%). On peut noter une différence importante entre les hommes qui sont engagés dans des tâches plus dures et demandantes que les femmes (14% pour 7.4%), tandis que ces dernières sont plus représentées dans des tâches de marches ou d'activités douces (31% contre 27%) ainsi que dans la catégorie sans tâches spécifiques (10% contre 8,5%).

La répartition du milieu de résidence semble équilibrée entre les sexes avec une plus grande représentation dans les grandes villes (38%) et les villes urbanisées (32%). Le niveau de diplôme varie selon le genre ; on voit que les femmes sont plus nombreuses à disposer d'un niveau d'éducation de type supérieur long (46%) par rapport aux hommes (39%).

La majorité des répondants sont d'origine belge (83%), suivis de répondants issus de hors Europe (11%)

et d'Europe (6.3%). Enfin, le statut matrimonial nous indique que les hommes sont plus souvent célibataires que les femmes (34% contre 30%) et que les femmes sont plus divorcées (12% contre 9.3%) ou veuves (2.3% contre 1.2%).

Ces résultats nous permettent de qualifier notre échantillon et d'observer déjà certaines tendances relatives à la pratique d'une activité physique de loisir.

4.1 Analyse bivariée entre quintiles de revenus et risques d'obésité ou de diabète

L'objectif est d'examiner s'il existe un lien entre l'appartenance à un groupe de revenu et le risque d'obésité et de diabète. L'ANOVA permet de comparer les moyennes entre plusieurs groupes afin de déterminer si des différences statistiques significatives existent. Cependant, le test de Levene a indiqué une p-valeur inférieure à 0,05, ce qui signifie que l'hypothèse d'homogénéité des variances n'est pas respectée. Dans ce cas, l'utilisation de l'ANOVA classique n'est pas appropriée.

Par conséquent, un test non paramétrique est privilégié : le test de Kruskal-Wallis, qui constitue une alternative à l'ANOVA à un facteur. Ce test repose sur les rangs et ne suppose pas l'égalité des variances ni la normalité des données. Il généralise le test de Wilcoxon pour comparer plus de deux groupes.

4.1.1 Test de Kruskal-Wallis : Obésité

Ce premier graphique présente le taux d'obésité au sein de chaque groupe de revenus.

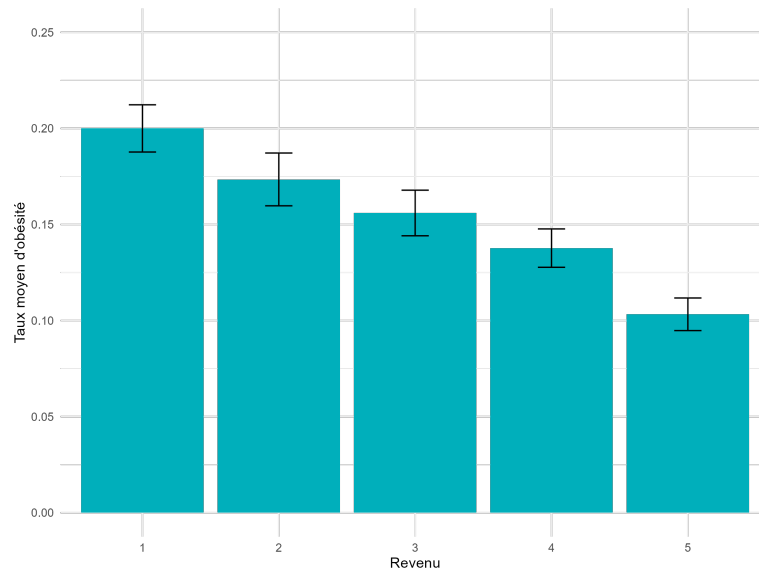


Figure 3 Prévalence de l'obésité par quintile de revenu (avec intervalles de confiance)
sur base des données HIS (2018)

Le test a révélé une différence significative d'obésité entre les groupes de revenu (*Kruskal-Wallis* ($\chi^2 = 47.72$, $df = 4$, $p < 0.001$), indiquant que la prévalence de l'obésité varie selon le niveau de revenu.

Une comparaison par paire de groupe a été réalisée sur base du test de Wilcoxon avec une correction et a montré des différences (Voir annexe tableau 6) :

- Entre les groupes de revenu 1 et 4 (p ajustée = 7.67×10^{-4})
- Entre les groupes 1 et 5 (p ajustée < 0.0001)
- Entre les groupes 2 et 5 (p ajustée < 0.0001)
- Ainsi qu'entre les groupes 3 et 5 (p ajustée = 0.002)

4.1.2 Test de Kruskal-Wallis : Diabète

Nous reproduisons maintenant la même stratégie d'analyse sur la prévalence du diabète. Ce premier graphique nous permet de voir comment se répartit la prévalence du diabète au sein des différents groupes de revenus.

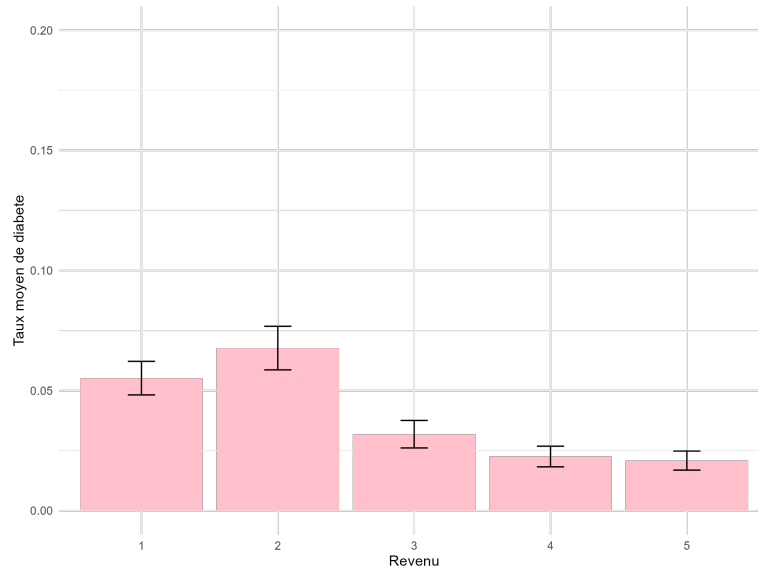


Figure 4 Prévalence du diabète par quintile de revenu (avec intervalles de confiance)
sur base des données HIS (2018)

Le test indique que les niveaux de diabète ne sont pas répartis de manière uniforme entre les groupes de revenu (*Kruskal-Wallis* $\chi^2 = 47.408$, $df = 4$, $p \leq 0.001$).

Une comparaison par paire de groupe a été réalisée sur base du test de Wilcoxon avec une correction et a montré des différences significatives (annexe tableau 7). Les différents groupes présentent des p-valeurs significatives et sont distincts les uns des autres.

La distribution du diabète varie significativement selon les groupes de revenu, avec des différences particulièrement marquées entre les groupes à faible revenu et ceux à revenu plus élevé.

Pour étudier cette relation, des modèles de régression vont être testés afin d'évaluer l'impact du revenu sur les liens déjà établis entre activité physique et l'incidence de l'obésité et du diabète. Ces modèles vont permettre de mieux comprendre si et comment le revenu modifie les effets protecteurs de l'activité physique sur ces maladies.

4.2 Analyses factorielles et clustering

4.2.1 Analyse de données mixtes

Pour simplifier l'analyse, une analyse factorielle a été réalisée. L'objectif d'une analyse factorielle est d'extraire les informations importantes, de les représenter sous la forme d'un ensemble de nouvelles variables orthogonales appelées composantes principales (Abdi & Williams, 2010). Des variables numériques et catégorielles ont été utilisées afin d'avoir une réduction de variables en dimensions les plus intéres-

santes possibles. C'est pourquoi la méthode choisie est l'analyse factorielle des données mixtes (FAMD) qui est une méthode destinée à analyser un ensemble de données contenant à la fois des variables quantitatives et qualitatives (Pagès, 2004).

Table 2 Description des variables quantitatives introduites et leurs modalités

Variable	Modalité
PA02 : Nombre de jours de marche (minimum 10 minutes) pour se rendre à différents endroits et en revenir au cours d'une semaine type.	X Valeur variable : nombre de jours -1 ² Sans réponse -3 ³ Pas applicable
PA04 : Nombre de jours à vélo (minimum 10 minutes) pour se rendre à différents endroits dans un petit village typique.	X Valeur de la variable : nombre de jours -1 Sans réponse -3 Pas applicable
PA06 : Nombre de jours consacrés à la pratique d'un sport, d'une activité physique ou récréative au cours d'une semaine type	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable
PA07h : Temps consacré aux sports, au fitness ou aux activités récréatives au cours d'une semaine type : heures	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable
PA07m : Temps consacré aux sports, au fitness ou aux activités récréatives au cours d'une semaine type : minutes	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable
PA08 : Nombre de jours d'activités visant à renforcer les muscles au cours d'une semaine type	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable
PA09h : Temps passé en position assise au cours d'une journée type : heures	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable

(Suite en page suivante)

²Ces cas n'ont pas été pris en compte dans l'analyse.

³Ces cas n'ont pas été pris en compte dans l'analyse.

Variable	Modalité
PA09m : Temps passé en position assise au cours d'une journée type : minutes	X Valeur variable : nombre de minutes -1 Sans réponse -3 Pas applicable

Table 3 Description des variables catégorielles introduites et leurs modalités

Variable	Modalités
PA_1 : Population âgée de 15 ans et plus respectant les recommandations de l'OMS en matière d'activités physiques aérobies bénéfiques pour la santé (non liées au travail)	1 Oui 2 Non -1 ⁴ Sans réponse -3 ⁵ Pas applicable
PA_2 : Population âgée de 15 ans et plus respectant les recommandations de l'OMS en matière d'activités physiques favorisant la santé et le renforcement musculaire (non liées au travail)	1 Oui 2 Non -1 Sans réponse -3 Pas applicable
PA_3 : Population âgée de 15 ans et plus respectant les recommandations de l'OMS en matière d'activités physiques aérobies renforçant la musculature et favorisant la santé (hors activités professionnelles)	1 Oui 2 Non -1 Sans réponse -3 Pas applicable
PA_4 : Population âgée de 15 ans et plus respectant les recommandations de l'OMS en matière d'activités physiques totales favorables à la santé (activités physiques aérobiques (non liées au travail) et activités physiques au travail)	1 Oui 2 Non -1 Sans réponse -3 Pas applicable
PA01 : Description du travail	1 Principalement assis ou debout 2 Principalement marche ou tâches nécessitant un effort physique modéré 3 Principalement travail pénible ou exigeant physiquement 4 Ne réalise aucune tâche professionnelle -1 Sans réponse -3 Pas applicable

(suite page suivante)

⁴Ces cas n'ont pas été pris en compte dans l'analyse.

⁵Ces cas n'ont pas été pris en compte dans l'analyse.

Variable	Modalités
PA03 : Temps passé à marcher lors d'une telle journée	1 10 à 29 minutes par jour 2 30 à 59 minutes par jour 3 1 heure à moins de 2 heures par jour 4 2 heures à moins de 3 heures par jour 5 3 heures ou plus par jour -1 Sans réponse -3 Pas applicable
PA05 : Temps passé à faire du vélo lors d'une telle journée	1 10 à 29 minutes par jour 2 30 à 59 minutes par jour 3 1 heure à moins de 2 heures par jour 4 2 heures à moins de 3 heures par jour 5 3 heures ou plus par jour -1 Sans réponse -3 Pas applicable
PA10 : Activité physique pendant les loisirs au cours des 12 derniers mois	1 Entraînement intensif et sport de compétition plus d'une fois par semaine 2 Jogging et autres sports récréatifs ou jardinage, au moins 4 heures par semaine 3 Jogging et autres sports récréatifs ou jardinage, au maximum 4 heures par semaine 4 Marche, vélo ou autres activités légères, au moins 4 heures par semaine 5 Marche, vélo ou autres activités légères, au maximum 4 heures par semaine 6 Lecture, télévision ou autres activités sédentaires -1 Sans réponse -3 Pas applicable

Les résultats obtenus mettent en avant deux dimensions qui permettent de résumer les informations disponibles sur l'activité physique des répondants.

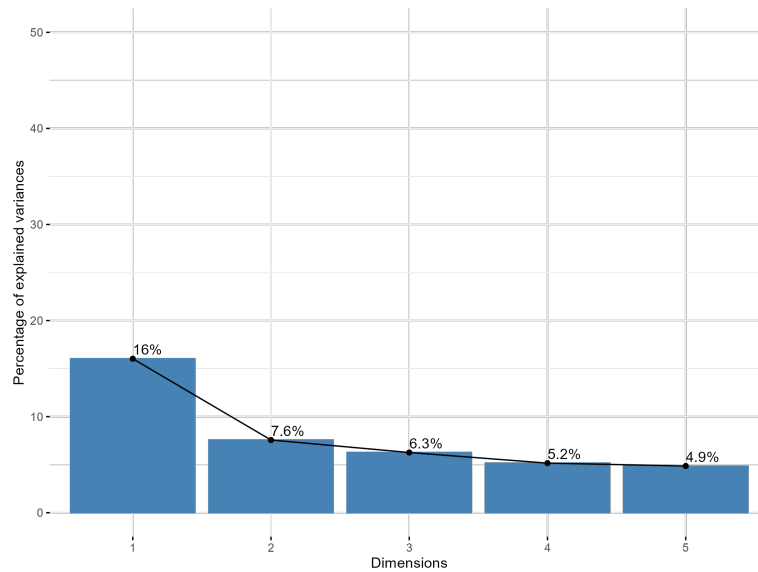


Figure 5 Variance expliquée par les différentes dimensions
 Résultat de l'analyse factorielle de données mixtes avec pondérations

Sur base de ce graphique, on peut voir une perte importante de variance expliquée entre la dimension 1 et 2. Le « coude » visible après la première dimension indique que l'ajout de dimensions supplémentaires au-delà de la deuxième apporte un gain marginal en termes de variance expliquée. Par conséquent, retenir deux dimensions permet de capturer l'essentiel de l'information tout en évitant la complexité inutile d'un modèle comptant trop de dimensions. Ce compromis entre simplicité et exhaustivité garantit une interprétation claire des résultats sans perte significative d'information.

Ensuite, en analysant la contribution de chaque variable à ces deux dimensions, il a été possible de les interpréter (annexe graphique 12). A l'aide des graphiques de contributions des variables qualitatives et quantitatives, il a été possible de catégoriser les deux dimensions en précisant la nature exacte de ce que chacune des dimensions mesurait.

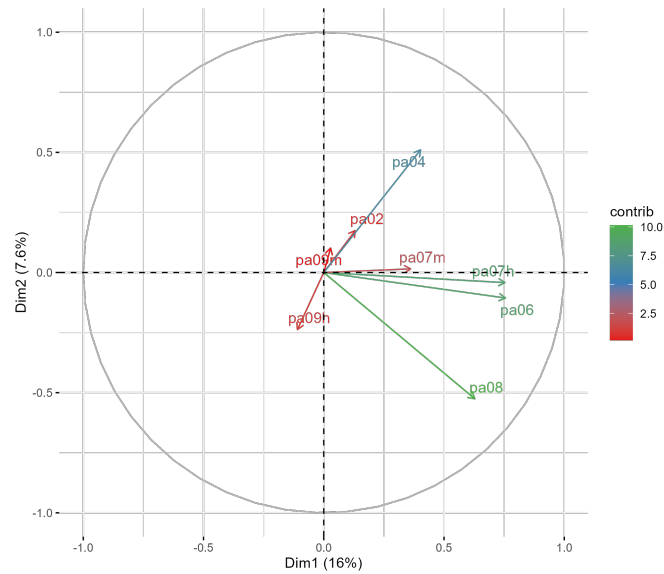


Figure 6 Visualisation des contributions des variables quantitatives à la FAMD
Réalisé à partir des résultats FAMD

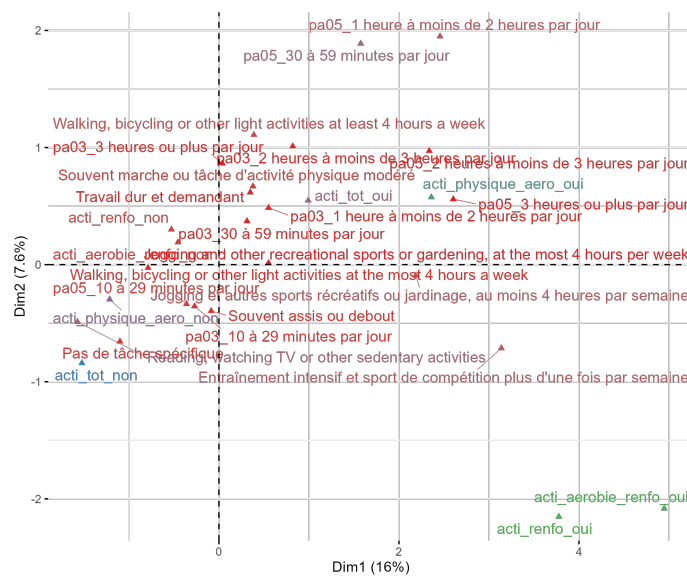


Figure 7 Visualisation des contributions des variables qualitatives à la FAMD
Réalisé à partir des résultats FAMD

Pour la dimension 1, les variables fortement corrélées sont : le temps consacré à la pratique d'une activité physique (pa06), l'heure consacré à une activité sportive ou récréative (pa07h), le type d'activité pratiquée au cours des 12 derniers mois (pa10) et le nombre de jours consacré à une activité de renforcement par semaine (pa08), ainsi que les modalités qualitatives ; le respect des recommandations de l'OMS en matière d'activités physiques aérobiques bénéfiques pour la santé non liées au travail (acti_physique_aero_oui), le respect des recommandations de l'OMS en matière d'activités physiques favorisant la santé et le renforcement musculaire non liées au travail (acti_renfo_oui) et le respect des recommandations de l'OMS en matière d'activités physiques aérobiques de renforcement musculaire

favorables à la santé non liées au travail (`acti_aerobie_renfo_oui`). La dimension semble opposer les individus actifs physiquement aux moins actifs. Elle capte l'intensité générale de l'activité physique, aussi bien d'endurance que de renforcement musculaire. Les individus avec un score élevé sont très engagés physiquement et sont probablement dans des activités structurées et régulières telles que l'entraînement en salle ou les sports encadrés.

Pour la dimension 2, les variables les plus contributrices sont : le nombre de jours passés à faire du vélo par semaine (`pa04`), le nombre de minutes passées à se déplacer à vélo (`pa05`), le nombre de jours consacrés à une activité de renforcement par semaine (`pa08`) et le type d'activité pratiquée au cours des 12 derniers mois (`pa10`), ainsi que les variables qualitatives de respect des recommandations de l'OMS en matière de renforcement musculaire non liées au travail (`acti_renfo_oui`) et de respect des recommandations de l'OMS en matière d'activités physiques aérobies de renforcement musculaire favorables à la santé non liées au travail (`acti_aerobie_renfo_oui`). Cette dimension semble davantage refléter un style de vie actif sans nécessairement faire référence à l'intensité. Les individus avec un score élevé adoptent des comportements actifs modérés, comme faire du vélo pour leurs déplacements (`pa04`, `pa05`) et pratiquer occasionnellement du renforcement musculaire (`pa08`). On est ici davantage dans une logique de routine active ou d'habitudes en faveur de la santé, plutôt que de sport intense.

4.2.2 Clustering

En se focalisant sur ces deux dimensions, et de ce deux dimensions seules, différents groupes ont été créés sur la base d'un clustering au départ de la distance entre observations sur ces deux axes. Afin de décider combien de groupe garder, la méthode du coude a été utilisée. Cette méthode permet de déterminer visuellement combien de dimensions garder en se basant sur les valeurs propres. Tous les facteurs situés avant le changement de pente (le pli du coude) sont pris en compte, au-delà de ce pli les facteurs supplémentaires apportent peu à l'explication.

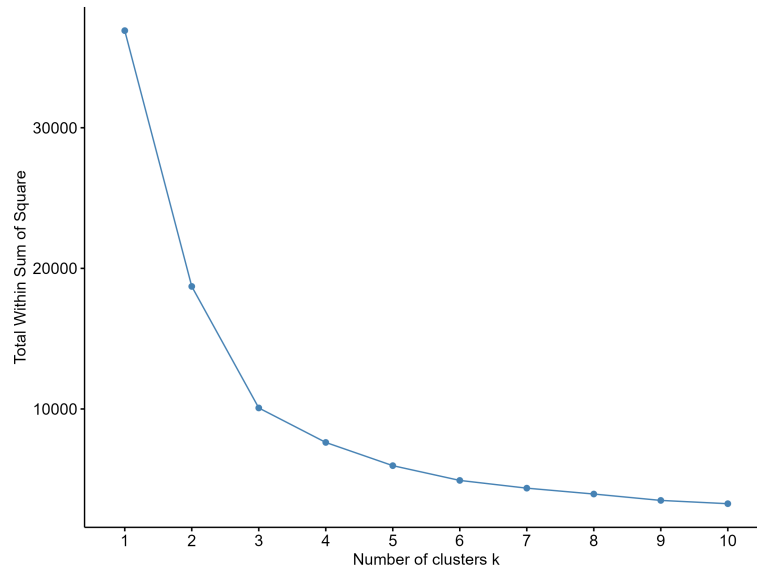


Figure 8 Sélection du nombre de clusters via la méthode du coude appliquée au clustering k-means

Réalisé à partir des données HIS (2018)

Sur la base de ce graphique, on voit que le coude se situe à 3 groupes. Au-delà, chaque cluster supplémentaire n'apporte plus beaucoup de réduction de variance intra-groupe. Toutefois, compte tenu des hypothèses formulées dans ce mémoire, il est préférable de retenir cinq groupes afin d'éviter une typologie trop simplifiée et d'obtenir des résultats plus détaillés et nuancés concernant les comportements observés.

4.2.2.1 Caractérisation des groupes

Ce tableau reprend les différents groupes sur la base du clustering et permet de se rendre compte des proportions présentes d'individus dans chaque groupe d'activité physique.

Table 4 Répartition des individus selon les clusters d'activité physique, par genre.

Réalisé sur base des données HIS (2018)

Characteristic	Overall N = 5,826	Homme N = 2,873	Femme N = 2,953
Cluster			
Sédentaires	1,850 (32%)	807 (28%)	1,043 (35%)
Actifs modérés	1,636 (28%)	749 (26%)	886 (30%)
Sportifs irréguliers	808 (14%)	435 (15%)	373 (13%)
Sportifs complets	1,003 (17%)	578 (20%)	426 (14%)
Sportifs intensifs	529 (9%)	304 (11%)	225 (7.6%)

Sur base de ce tableau, on peut dire que notre échantillon se compose principalement d'individus sédentaires (32%), suivis d'actifs modérés (28%), des sportifs complets (17%), des sportifs irréguliers (14%) et des sportifs intensifs (9%). Par rapport aux clusters d'activité physique de loisir, on constate que

les femmes sont plus souvent sédentaires (35% contre 28% chez les hommes) et actives modérées (30% contre 26%). En opposition les hommes sont plus représentés dans les catégories de sportifs irréguliers (15% contre 13%), sportifs complets (20% contre 14%) et sportifs intensifs (11% contre 7,6%). Cette répartition illustre une tendance des hommes à pratiquer davantage d'activité physique de loisir régulière et intensive comparée aux femmes. Ces différences de pourcentages entre les hommes et les femmes de l'échantillon sont même statistiquement significatives au seuil de 0.05 (annexe 8 : $p < 0.0001$) indiquant une réelle différence dans la pratique d'activité physique entre les deux sexes.

Une analyse plus détaillée va nous permettre d'affiner ces groupes et de les caractériser selon leurs pratiques et leurs intensités.

Le graphique ci-dessous permet de se rendre compte des positions des différents clusters les uns par rapport aux autres sur les deux dimensions. Il présente l'ensemble des individus enquêtés selon leur score obtenu sur les deux dimensions, et regroupés via l'analyse de clustering.



Figure 9 Positionnement des clusters d'activités physiques dans l'espace factoriel (FAMD)
Réalisé sur base du clustering et FAMD

Le point blanc représente le centre de gravité (moyenne des scores) des individus dans chaque cluster. Le graphique illustre la manière dont les groupes se positionnent selon les deux dimensions : intensité de l'activité physique (axe horizontal) et style de vie actif (axe vertical).

Cluster 1 : Sédentaires, peu actifs

Sur la première dimension, ce groupe se caractérise par une moyenne très faible (-1.81 ; annexe 9 pour les moyennes) et par un engagement très faible dans les activités physiques, qu'elles soient d'endurance ou de renforcement musculaire. Les individus de ce cluster ont aussi une moyenne très faible sur la

dimension deux (-0.91) indiquant que ces individus n'ont pas non plus de comportements actifs modérés ou de routines de mobilité douce. Ce profil s'observe également dans leurs habitudes déclarées : plus de 60% marchent moins de 30 minutes par jour, et près de 94% ne déclarent aucune activité physique régulière. L'activité professionnelle est également peu exigeante physiquement, avec plus de 75% des individus déclarant être « souvent assis ou debout ». On peut donc conclure que ce sont des individus très peu actifs et peu engagés dans des activités physiques.

Cluster 2 : Actifs modérés

Sur la première dimension, la moyenne de ce groupe est aussi basse, mais plus élevée que pour le cluster 1 (-1.02). Cela indique que ces individus ne s'engagent pas dans des activités physiques régulières ou intensives. En ce qui concerne la deuxième dimension, on voit une moyenne positive (0.6) ce qui indique une certaine routine active. Les individus de ce cluster se distinguent par une proportion relativement élevée d'individus marchant quotidiennement : plus de 30 % marchent au moins 30 minutes par jour, et environ 10 % utilisent le vélo pour leurs déplacements quotidiens. De plus, 38 % respectent les recommandations de l'OMS en matière d'activité aérobie (hors travail), un taux supérieur à celui du cluster 1. Ce cluster se caractérise aussi par une proportion de 20 % d'individu pratiquant une activité physique de loisir de manière modérée à intense, incluant le jogging, le sport récréatif ou encore le jardinage. On peut donc dire que ce groupe incarne un style de vie actif modéré, avec des habitudes saines mais sans engagement sportif intense. Ils semblent privilégier la pratique d'une activité physique comme moyen de santé et de bien-être plutôt que pour la performance ou la compétition.

Cluster 3 : Sportifs irréguliers

Ce cluster a une moyenne positive sur la première dimension (1.06) ce qui indique un engagement plus important dans des activités physiques intenses et structurées. Sur la deuxième dimension, la moyenne de ce groupe est négative (-0.39) ce qui suggère un style de vie moins marqué par des routines actives ou des habitudes de mobilité douce comme pour le cluster 2. En effet près de 50% des individus de ce cluster ne marchent pas plus de 30 minutes par jour et moins de 10% utilisent le vélo de manière régulière. En revanche, près de 40% déclarent pratiquer une activité sportive de loisir au moins 4 heures par semaine ou plus, et plus de 50% respectent les recommandations de l'OMS en matière d'activité physique aérobie (hors travail). On peut qualifier ce groupe de sportifs sans mode de vie actif dans la vie quotidienne (peu de vélo pour se déplacer, moins d'habitudes en faveur de la santé que le cluster 2 mais une activité physique tout de même importante).

Cluster 4 : Sportifs complets

Ce groupe a une moyenne très élevée sur la dimension 1 (1.61) ce qui indique une forte participation à des activités physiques. Sur la deuxième dimension, la moyenne du groupe est aussi positive, indiquant des habitudes et des comportements de pratiques modérées et de santé (mobilité douce, renforcement musculaire). Plus de 75% des individus de ce cluster respectent les recommandations de l'OMS en matière d'activité physique, aussi bien aérobie que musculaire. Environ 65% déclarent marcher plus

d'une heure par jour et la majorité pratiquent une activité physique de loisir de plus de 4 heures par semaine à travers des activités sportives régulières et encadrées. Ces individus combinent donc entraînements structurés avec une vie quotidienne active. Ils pratiquent des sports réguliers en salle ou encadrés, complétés par une routine de mobilité douce et d'habitudes santé. Ce groupe a aussi le plus faible pourcentage de sédentarité avec un temps passé assis le plus bas parmi tous les groupes (pa09h et pa09m). En somme, c'est le profil « idéal » d'un mode de vie sportif complet et équilibré.

Cluster 5 : Sportifs intensifs

Sur la première dimension, ce groupe a la moyenne la plus importante (4.75), ce qui renvoie à l'intensité physique la plus forte. Près de 90% de ces individus déclarent faire du sport au moins plusieurs fois par semaine, souvent dans un cadre structuré comme un entraînement ou une compétition. Sur la deuxième dimension, ils ont une valeur moyenne très négative (-2.15), indiquant qu'ils n'adoptent pas du tout un style de vie actif routinier ou des habitudes modérées. On peut dire que ces individus sont très investis dans une pratique sportive intense, probablement compétitive ou professionnelle. A côté de ça, ils n'adoptent pas de comportements de mobilité douce et d'habitudes santé comme par exemple un déplacement à vélo ou en marchant. Ces individus semblent être axés sur une pratique compétitive, centrée sur la performance sans traduire cet engagement dans la vie quotidienne (contrairement au cluster 4).

4.2.2.2 Clusters et variable de charge de travail (pa01)

Les deux premières dimensions de l'ACP captent principalement la variabilité liée à la pratique d'activité physique de loisir et aux déplacements actifs, aspects importants pour évaluer l'engagement physique des individus indépendamment du travail. La variable pa01, décrivant la nature de l'activité physique au travail, est peu contributive sur ces dimensions mais joue un rôle clé sur une dimension supérieure (dimension 3) qui distingue les individus selon la charge physique liée à leur travail. Cette distinction est importante pour notre hypothèse selon laquelle un travail physiquement exigeant peut limiter l'activité physique de loisir, notamment chez les individus à statut socio-économique plus faible. Ainsi, bien que la synthèse en deux dimensions rende compte de l'essentiel de la variabilité liée à l'activité physique volontaire, il est important de considérer la nature du travail comme un facteur complémentaire influençant le comportement physique global.

Ce tableau croisé permet de mettre en perspective les clusters avec les différentes modalités de pa01.

Cluster d'activité physique	Souvent assis ou debout	Marche ou activité modérée	Travail dur et demandant	Pas de tâche spécifique
Sédentaires	76.7	1.9	0.6	20.8
Actifs modérés	18.7	57.0	20.8	3.5
Sportifs irréguliers	64.8	21.2	7.9	6.1
Sportifs réguliers	45.1	38.7	12.2	4.0
Sportifs intensifs	51.3	33.5	10.5	4.6

Table 5 Répartition (%) de la charge physique liée au travail selon les clusters d'activité physique.

Réalisé sur base du clustering.

Sur base de ce tableau, on peut voir que le groupe sédentaire se caractérise par une charge de travail qui nécessite d'être souvent assis ou debout (76%) tandis que les actifs modérés sont soumis à des charges modérées (57%) ou exigeantes (20.8%). Les sportifs réguliers et intensifs, quant à eux, se situent entre ces deux extrêmes : une part non négligeable exerce un travail physiquement modéré (38,7% et 33,5% respectivement), mais une proportion notable est aussi souvent assise (45 à 51%). Enfin, les sportifs irréguliers montrent un profil oscillant entre des tâches assises ou debout (64,8%) et des tâches modérées (21.2%).

Un test de khi carré a aussi été effectué afin de voir si l'association entre les deux variables était bien significative (annexe tableau 10). Le résultat de la p-valeur nous indique qu'il existe bien un lien statistique entre les clusters d'activité physique et la charge physique liée au travail. Cette analyse croisée permet d'affiner les clusters en détaillant l'activité physique liée au travail et de justifier son intégration dans les modèles, car elle peut capter un effet résiduel ou complémentaire que les clusters d'activité physique ne saisissent pas complètement.

Cette première série d'analyses a permis d'identifier des groupes d'activités physiques de loisir distincts, différenciés selon des dimensions de fréquence et d'intensité. Le test du khi-deux réalisé entre ces clusters et la variable de charge physique liée au travail a confirmé l'existence d'un lien statistiquement significatif, ce qui est crucial pour la suite des analyses. Par ailleurs, l'ANOVA a révélé des différences notables de prévalence d'obésité et de diabète selon les quintiles de revenus. Ces résultats établissent ainsi une base solide pour approfondir l'analyse des facteurs explicatifs à travers des modèles de régression.

4.3 Modèles de régressions

4.3.1 Modèles de régressions : Obésité

Modèle 1 : Obésité selon l'activité physique.

Dans ce premier modèle, on teste comment l'activité physique impacte la probabilité d'être en situation d'obésité. Les résultats nous montrent que tous les clusters se distinguent significativement du cluster de référence ($p < 0.05$). Les groupes d'actifs modérés, de sportifs irréguliers, de sportifs complets et de sportifs intensifs ont des effets significativement négatifs par rapport au groupe des sédentaires. Les groupes ayant une activité physique plus élevée sont associés à moins d'obésité.

Les odds ratio nous indiquent que l'effet protecteur augmente avec l'intensité de l'activité physique (annexe tableau 11). Les actifs modérés ont 35% de chances en moins d'être en situation d'obésité, en comparaison aux sédentaires (OR=0.77 ; $p = 0.02$). Ce risque chute à 54% chez les sportifs irréguliers (OR = 0,46 ; $p < 0,001$), à 59% chez les sportifs complets (OR = 0,41 ; $p < 0,001$), et atteint 68% de réduction chez les sportifs intensifs (OR = 0,32 ; $p < 0,001$). Ces résultats soulignent un gradient d'effet selon l'intensité de l'activité physique : plus cette dernière est soutenue, plus le risque d'obésité diminue.

Modèle 2 : Obésité selon l'activité physique et la charge physique liée au travail

Dans un second modèle (annexe tableau 12), la variable pa01, décrivant le niveau d'activité physique au travail, a été introduite en complément des clusters d'activité physique de loisir. Les résultats nous indiquent qu'aucune des modalités d'activité physique liées travail (pas de tâche spécifique, souvent assis ou debout, souvent marche modérée et travail dur et demandant) n'a montré d'effet statistiquement significatif sur l'obésité par rapport à la catégorie de référence. L'ajout de cette variable n'a pas modifié l'association protectrice des clusters de loisir, qui restent significativement liés à un moindre risque d'obésité. Ces résultats suggèrent que, pour l'obésité, l'activité physique liée aux loisirs est un facteur protecteur important, alors que l'activité physique associée à la charge de travail ne semble pas exercer d'effet clair.

Modèle 3 : Obésité selon le revenu

Dans ce modèle, on examine comment le revenu impacte la prévalence de l'obésité au sein de la population. Les résultats indiquent que les groupes de revenu 3, 4 et 5 sont statistiquement significatifs par rapport au groupe de référence de revenu 1 (annexe tableau 13). Les groupes de revenu 4 ($p = 0.025$) et 5 ($p = 0.0001$) sont significativement associées à une baisse du risque d'obésité. Le niveau de revenu 3 ($p = 0.069$) diminue aussi le risque d'obésité (au seuil de 0.01). Cela nous indique que le revenu élevé est associé à une moindre prévalence de l'obésité.

Sur base des odds ratio, on observe que les individus de revenu 3 ont 29% de chances en moins d'être obèses comparés aux individus aux revenus les plus bas (OR = 0.71 ; IC [0,52 ; 0,97] ; $p = 0,03$) (annexe tableau 13). Cette association est aussi significative pour les revenus de catégorie 4, avec une réduction

du risque d'obésité de 34% (OR = 0,661 ; IC95 % [0,485 ; 0,898] ; $p = 0,008$). Le groupe de revenu 5 montre un effet protecteur encore plus marqué, avec une diminution du risque d'obésité de 54,5% (OR = 0,456 ; IC95 % [0,325 ; 0,637] ; $p < 0,001$).

Modèle 4 : Obésité selon l'activité physique et le revenu

Dans ce dernier modèle, on intègre un terme d'interaction entre les différents clusters d'activité physique et le revenu. Les résultats montrent que les effets pour les groupes de sportifs irréguliers et de sportifs complets restent statistiquement significatifs au seuil de 0.05 indépendamment du revenu (annexe tableau 14). Le cluster 5, des sportifs intensifs, n'est lui plus significatif.

Seul le cinquième quintile de revenus et le cluster comportant les sportifs intensifs se distinguent significativement du cluster des sédentaires du premier quintile de revenu. Cela nous indique que ceux qui ont le revenu le plus élevé et qui font partie des sportifs intensifs connaissent une réduction marquée du risque d'obésité, au-delà des effets principaux (OR = 0.08 ; $p = 0.03$). En d'autres termes, le revenu renforce l'effet protecteur de l'activité physique pour ce groupe. La plupart des autres interactions sont non significatives, indiquant que l'effet du revenu est similaire pour les autres groupes ; l'effet combiné du revenu varie peu pour les autres groupes. Toutefois, une autre interaction est apparue significative entre le cluster 3 et le revenu 3 ($p = 0.0232$), mais avec un odd ratio supérieur à 1 (OR=4.02 ; IC95%[1.25 ; 14.31] ; $p = 0.023$), cette association doit être interprétée avec prudence, car l'intervalle de confiance est large (annexe tableau 14). Ces résultats méritent donc d'être confirmés par des analyses complémentaires ou sur un échantillon plus important.

En résumé, un revenu élevé produit un effet protecteur indépendant, ce qui signifie que quels que soient les niveaux d'activité physique, avoir un revenu élevé réduit le risque d'obésité. De même, adopter une activité physique de type sportifs irréguliers ou sportifs réguliers diminue les risques d'obésité. Ces effets sont robustes : même en tenant compte de l'interaction entre activité physique et revenu, ces facteurs gardent leur influence propre sur l'obésité. Enfin, pour les revenus les plus élevés et les sportifs intensifs, le revenu renforce l'effet protecteur de l'activité physique. Autrement dit, le revenu amplifie l'effet protecteur dont les individus bénéficient au travers de leurs activités physiques par rapport à l'obésité.

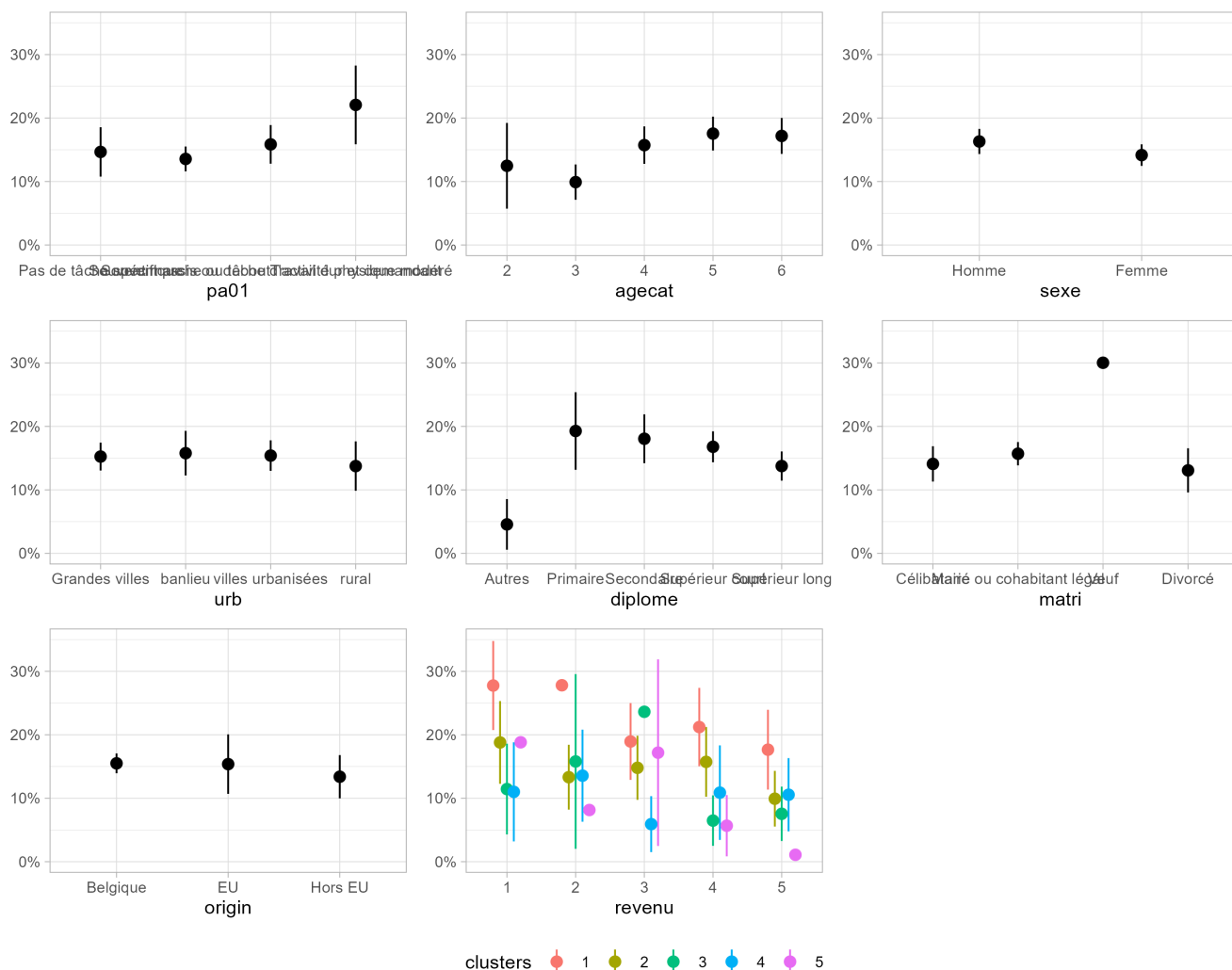


Figure 10 Prédictions du modèle 4 pour l'obésité
Réalisé sur base des données HIS (2018)

Les prédictions marginales peuvent nous servir à affiner nos résultats (annexe tableau 15). On peut observer plusieurs éléments au niveau des variables socio-démographiques. Les probabilités d'obésité augmentent avec l'âge et atteignent un pic entre 45 et 54 ans (17,56% ; IC[14,9–20,2%] ; $p < 0,001$) avant de baisser légèrement pour les 55-64 ans (17,2% ; IC [14,4–20%] ; $p < 0,001$). Toutes les associations sont hautement significatives ($p < 0,001$). Les hommes sont plus à risque comparativement aux femmes (16,3% contre 14,2% pour les femmes). La probabilité d'obésité varie peu selon le type de milieu de vie, bien que légèrement plus faible en zone rurale (13,7% ; IC [9,8–17,6%]) par rapport aux grandes villes (15,3% ; IC [13,1–17,65%]). Les différences restent significatives ($p < 0,001$). Le niveau d'éducation montre une relation inverse avec l'obésité. Les individus ayant un diplôme primaire ou secondaire affichent des taux respectifs de 19,3% (I C[13,2–25,4%]) et 18,1% (IC [14,2–22%]), nettement plus élevés que ceux ayant un diplôme supérieur long (13,7%, IC [11,4–16,0%]). Les veufs présentent la probabilité la plus élevée d'obésité (30% ; $p < 0,001$) comparativement aux mariés, célibataires et divorcés, ce qui suggère qu'une dimension sociale pourrait jouer un rôle dans le risque d'obésité.

Au niveau de l'activité physique liée au travail, on voit que les individus exerçant un travail dur et exigeant physiquement présentent le risque le plus élevé d'obésité (22,1% ; IC [15,9–28,3%] ; $p < 0,001$), comparé à ceux sans tâche spécifique (14,8% ; IC [10,9–18,7%]). Ceux dont l'activité implique d'être souvent assis ou debout ou de marcher affichent des probabilités intermédiaires (respectivement 13,6% et 15,9%). Ces résultats peuvent paraître surprenants mais pour autant notre 2e modèle a montré qu'une fois les clusters d'activités physiques contrôlés, l'association significative au niveau de l'activité physique liée au travail devenait nulle.

Chez les individus sédentaires et peu actifs, la prévalence d'obésité passe de 27,8% pour les revenus faibles à 17,6% pour les revenus élevés, illustrant clairement un effet protecteur. Pour les actifs modérés, cette baisse est également notable, passant de 18,8% à 10%. Chez les sportifs irréguliers, la tendance est non linéaire puisqu'on observe un pic à 23,6% pour les catégories de revenu moyenne (revenu 3). Pour les sportifs complets, le risque reste relativement bas allant de 11% pour les catégories de revenu les plus basses à 10,5% pour les catégories les plus aisées, excepté pour les revenus moyen pour qui le risque semble relativement bas (5,9%). Enfin, chez les sportifs intensifs, le risque d'obésité est très faible pour les revenus élevés (1,1%), mais reste incertain pour les revenus faibles (18,8% ; IC [-0,5-38,1%] ; $p = 0.056$).

Ces résultats montrent que le revenu joue un rôle protecteur, surtout chez les moins actifs. L'effet conjugué d'une activité physique régulière et d'un revenu élevé aboutit aux taux d'obésité les plus bas observés dans l'échantillon.

Nous reproduisons maintenant la même stratégie d'analyse sur la prévalence du diabète.

4.3.2 Modèles de régressions : Diabète

Pour cette partie sur le diabète, les quintiles de revenu, utilisés initialement sous leur forme catégorielle (cinq groupes représentant 5 quintiles), ont été regroupés en deux catégories : *revenu bas à moyen* (quintiles 1 à 3) et *revenu élevé* (quintiles 4 et 5). En effet, l'utilisation des quintiles initiaux dans les modèles de régression a généré des résultats instables et parfois aberrants, en particulier dans le cadre du quatrième modèle intégrant des interactions entre les catégories de revenu et les clusters d'activité physique (soit 25 termes d'interaction). Les résultats suspects suggéraient un petit nombre de cas dans ces termes d'interaction et faussaient les modèles.

Ce regroupement vise donc à garantir des résultats plus robustes en augmentant la taille des effectifs dans chaque catégorie de revenu, ce qui permet d'obtenir des estimations plus stables. La nouvelle variable "revcat" comprend 2 modalités : *revenu bas à moyen* (quintiles 1 à 3) et *revenu élevé* (quintiles 4 et 5). Cette dichotomisation se justifie sur le plan théorique car on distingue un groupe potentiellement plus vulnérable en matière de santé et un groupe à revenu élevé souvent privilégié par rapport aux ressources et à l'activité physique.

Modèle 1 : Diabète selon l'activité physique

Ce premier modèle évalue l'impact de l'activité physique sur la probabilité d'être atteint de diabète. Les résultats nous montrent que le cluster 4 et le cluster 5 se distinguent significativement du cluster 1 de référence (annexe tableau 16). Les groupes de sportifs complets et intensifs ont des effets significativement négatifs par rapport au groupe des sédentaires. Les groupes ayant une activité physique plus élevée sont associés à moins de diabète.

Selon les odds ratio, le groupe de sportifs complets ont une réduction de 66% du risque de diabète par rapport aux sédentaires et peu actifs ($OR = 0.34$; $p = 0.0019$). Les sportifs intensifs ont une réduction de 79% de souffrir de diabète, en comparaison aux sédentaires ($OR = 0.21$; $p = 0.007$). Les profils les plus actifs en matière d'activité physique sont ceux qui sont le moins à risque de diabète.

Le modèle indique aussi que l'âge est une variable fortement liée au diabète ; plus les individus sont âgés, plus ils sont à risque de diabète. On peut aussi noter que les individus vivant en milieu rural ont 2.2 fois plus de chances d'avoir du diabète par rapport aux individus des grandes villes ($OR = 2.21$; $p = 0.0078$). Les individus originaires de hors Union européenne ont 2.9 fois plus de chances (odds) de souffrir de diabète que les natifs du pays ($OR = 2.98$; $p < 0.001$). Les odds ratios pour le statut matrimonial et le niveau de diplôme ne montrent pas d'association significative avec le risque de diabète, indiquant un risque relativement uniforme entre les groupes.

Modèle 2 : Diabète selon l'activité physique et la charge physique liée au travail

Dans un second modèle, la variable pa01, décrivant le niveau d'activité physique au travail, a été introduite en complément des clusters. Les résultats montrent qu'un travail physiquement exigeant est associé à une diminution significative du risque de diabète ($OR = 0.35$; $IC95\% [0.129 ; 0.957]$; $p = 0.04$) par rapport aux individus qui n'ont pas de tâche spécifique (annexe tableau 17).

Les individus qui sont souvent assis ou debout ont une réduction du risque de diabète de 37% par rapport aux individus qui n'ont pas de tâche spécifique ($OR = 0.62$; $IC95\% [0.37977 ; 1.024]$; $p = 0.06$). Pour autant cette association n'est significative qu'au seuil de 0.1.

Cela suggère que les comportements d'activité physique liée au travail ont un effet protecteur indépendant des activités de loisirs. Cependant, cet effet, étant marginalement significatif, mériterait d'être vérifié.

Modèle 3 : Diabète selon le revenu

Ce modèle examine comment le revenu impacte l'incidence de diabète. Les résultats suggèrent qu'il

existe une tendance vers une réduction du risque de diabète chez les personnes ayant un revenu élevé par rapport au revenu bas et moyen, mais cette relation n'est pas statistiquement significative (annexe tableau 18). Cela signifie que les données ne permettent pas de conclure avec certitude qu'il existe un effet du revenu sur la prévalence du diabète.

Modèle 4 : Diabète selon l'activité physique et le revenu

Les résultats de ce modèle suggèrent que l'effet protecteur associé à l'activité physique n'est pas uniforme selon le niveau de revenu (annexe tableau 19). Les individus actifs modérés ayant un haut revenu ont un risque de diabète 2.42 fois plus élevé que les sédentaires à bas revenu ($OR = 2.42$; $p = 0.05$). De même, les sportifs irréguliers avec un haut revenu ont un risque de souffrir de diabète significativement plus élevé que les sédentaires à bas revenu ($OR = 5.76$; $p = 0.019$). Ces résultats indiquent que chez les personnes ayant un haut revenu, les profils d'activité modérée ou irrégulière sont moins protecteurs qu'attendu vis-à-vis du diabète (voir modèle 1). Cela atténue ou contrebalance partiellement l'effet bénéfique attendu de l'activité physique.

En revanche, pour les profils les plus actifs, c'est-à-dire les sportifs complets et intensifs, l'interaction avec le revenu n'est pas significative, ce qui suggère que chez ces individus, l'effet protecteur de l'activité physique demeure stable quel que soit le revenu. Cela suggère que l'activité physique joue un rôle central dans la prévention du diabète, comme les individus les plus actifs présentent un risque réduit quel que soit leur revenu alors que ceux qui ont un haut revenu mais une activité physique modérée ou irrégulière sont paradoxalement plus exposés au diabète.

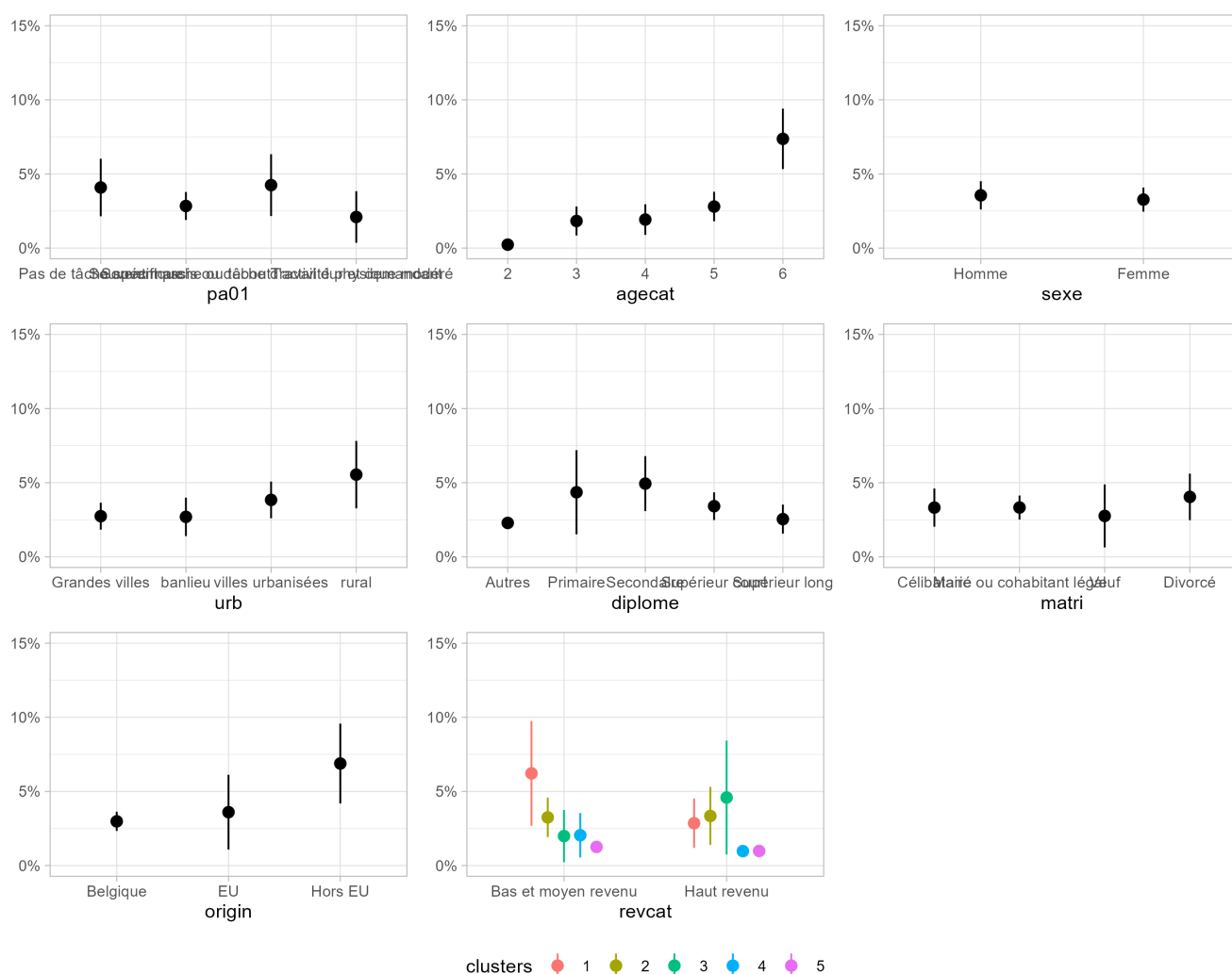


Figure 11 Prédiction du modèle 4 pour le diabète
Réalisé sur base des données HIS (2018)

Ce graphique permet d'illustrer les prédictions marginales du modèle. Ces prédictions confirment certains résultats obtenus avec les modèles de régression : l'âge, l'origine hors UE et le milieu rural restent des déterminants significatifs du diabète (annexe tableau 20). Les différences selon le sexe et le niveau d'étude demeurent également marqués, avec une prévalence plus faible du diabète chez les individus ayant un niveau d'éducation supérieur (2.5%).

Par rapport à la charge physique liée au travail, on observe un effet protecteur global de l'activité physique. Le risque de diabète est plus élevé chez les individus sans tâches spécifiques (4.1% ; IC [2.1 ; 6%] ; $p < 0.001$) et ceux qui marchent ou qui effectuent une tâche d'activité physique modéré (4.2% ; IC [2.2 ; 6.3%] ; $p < 0.001$). Les individus où la prévalence de diabète est la plus faible est chez ceux qui effectuent un travail dur et demandant (2.1% ; $p = 0.018$) et ceux qui sont souvent assis ou debout (2.8% ; $p < 0.001$).

En somme, certaines formes d'activité physique modérée ou irrégulière semblent moins protectrices chez les individus à haut revenu. En effet, les personnes appartenant au cluster 2 (actifs modérés) et au cluster 3 (sportifs irréguliers) avec un haut revenu présentent un risque de diabète respectivement de 3,3% et 4,6%, comparé à 6,2% chez les sédentaires à revenu bas et moyen. Ces taux sont même plus élevés que ceux observés dans les mêmes clusters à revenu bas ou moyen (respectivement 3,2% et 2,0%), suggérant une moindre efficacité de l'activité physique pour ces clusters lorsqu'elle est combinée à un haut revenu.

En revanche, pour les profils les plus actifs (cluster 4 et 5) le risque de diabète reste faible quel que soit le revenu. Les prédictions marginales indiquent des taux de diabète très faibles dans ces groupes (1.0% à 2.0%), sans différence significative entre niveaux de revenu. Ces résultats suggèrent que l'activité physique intense ou régulière constitue un facteur de protection robuste et indépendant du statut socio-économique, ce qui semble en ligne avec les résultats des modèles de régressions.

5 Discussion et conclusion

L'objectif de ce mémoire était de voir dans quelle mesure le statut socio-économique impactait les effets de la pratique d'une activité physique et ses bienfaits sur la santé. Les aspects bénéfiques du sport sur la santé ont vite été établis tant il existe des liens positifs entre les deux au niveau de la littérature scientifique. Pour autant, l'enjeu était de voir si selon l'appartenance à un certain groupe socio-économique les effets sur l'obésité et le diabète variaient. Pour ce faire, différentes hypothèses ont été formulées à la suite de la revue de littérature.

Les résultats des modèles de régressions confirment l'effet bénéfique de la pratique de l'activité physique de loisir sur l'obésité et le diabète. Une pratique plus intensive et régulière exerce un effet protecteur significatif. Les profils d'actifs modérés, de sportifs irréguliers, réguliers et intensifs présentent les taux d'obésité et de diabète les plus faibles. En somme, plus l'intensité de l'activité physique est élevée, plus le risque diminue, ce qui valide nos hypothèses 1 et 2. Ces effets sont d'autant plus robustes qu'ils se maintiennent même en tenant compte d'autres variables sociodémographiques ou de l'activité physique liée au travail.

En ce qui concerne l'hypothèse 3, les analyses montrent que l'activité physique exercée dans le cadre professionnel n'a pas d'effet protecteur clair contre l'obésité, bien qu'elle soit associée à une réduction modérée du risque de diabète. Ce résultat suggère que l'activité physique liée au travail n'offre pas le même effet protecteur que l'activité physique de loisir par rapport à l'obésité et au diabète. Les pratiques d'activité physique volontaires et de plaisir semblent donc jouer un rôle central, indépendant et potentiellement plus bénéfique.

Enfin, le revenu apparaît comme un déterminant indépendant du risque d'obésité, avec un effet protecteur croissant à mesure que le niveau de revenu augmente. Ce lien semble plus marqué chez les individus les moins actifs, pour qui un revenu plus élevé atténue les effets négatifs de la sédentarité et d'une faible activité physique. Par rapport au diabète, les résultats sont plus nuancés et moins clairs. Certaines formes d'activité physique modérée ou irrégulière semblent moins protectrices chez les individus à haut revenu, les exposant même à des risques plus élevés que pour les individus exerçant une même pratique mais avec un revenu plus bas. Ces résultats suggèrent une moindre efficacité de l'activité physique pour ces clusters lorsqu'elle est combinée à un haut revenu. En revanche, l'activité physique intense ou régulière constitue un facteur de protection robuste et indépendant du statut socio-économique. Notre hypothèse 4 est donc partiellement confirmée.

En somme, cette recherche a permis de mettre en évidence les effets de l'activité physique sur l'obésité et le diabète, en tenant compte des différences selon le statut socio-économique. Pour autant, il est important de revenir sur les limites de cette recherche concernant les données utilisées et la méthode d'analyse.

La plupart des recherches étudiant le lien entre sport et obésité et diabète se basent sur des données de type longitudinal ce qui permet notamment d'utiliser des modèles de Cox. Ce type de modèle est particulièrement utilisé en épidémiologie pour analyser les facteurs qui favorisent l'apparition d'une maladie. Il est d'autant plus utile car celui-ci permet de modéliser les risques compétitifs dans l'incidence d'un événement. Il est donc particulièrement pertinent dans le cadre d'une étude qui étudie l'impact du sport sur la santé. Avec les données Sciensano, on dispose uniquement de données transversales et c'est pourquoi des modèles de régression logistique ont été utilisés alors que le modèle de Cox est plus pertinent pour étudier ce sujet. Les analyses proposées restent limitées à des associations statistiques ponctuelles.

Par ailleurs, la mesure utilisée pour le niveau de revenu se base uniquement sur le revenu du ménage. Cette approche peut masquer d'importantes disparités, par exemple entre des ménages de taille très différente disposant d'un même revenu brut puisqu'au-delà de la taille du ménage, on ne tient pas compte de la composition du foyer ni d'autres indicateurs économiques. Une mesure plus robuste aurait pu inclure le revenu équivalent (ajusté à la fois selon la taille et la composition du ménage), ainsi que d'autres dimensions telles que le patrimoine, la stabilité des revenus ou encore les aides sociales perçues. L'intégration de ces variables aurait permis de mieux capturer les inégalités économiques et de nuancer l'impact réel du statut socio-économique sur les deux conditions étudiées.

De plus, certaines variables comme l'alimentation, le stress, la consommation d'alcool ou de tabac auraient pu être intéressantes à ajouter comme variables de contrôle afin d'avoir des modèles encore plus précis. L'absence de ces variables limite la capacité à isoler la relation qu'on veut étudier et peut biaiser l'interprétation des résultats.

La causalité est aussi une limite importante à prendre en compte. On ne peut pas éloigner la possibilité d'une causalité bidirectionnelle entre pratique sportive et revenu. L'hypothèse de base est que le statut socio-économique module les effets de la pratique sportive sur la santé. Cependant, l'inverse pourrait également être vrai. Le fait d'être engagé dans une activité physique peut favoriser le maintien ou même l'amélioration de la position socioéconomique. Il est alors difficile de déterminer quelle est la cause de quoi et surtout quelle est la cause antérieure. Ainsi, cette limite nous invite à interpréter les résultats avec prudence.

Enfin, une autre limite importante concerne la taille de l'échantillon, qui, bien que suffisante pour des analyses globales, devient restreinte dès lors que l'on croise plusieurs variables catégorielles, notamment les clusters d'activité physique et les quintiles de revenu (qui donnait 25 termes d'interactions). Cette configuration engendre des effectifs faibles dans certaines combinaisons, ce qui limite la robustesse statistique de certaines estimations. Cela a notamment conduit à regrouper certaines catégories de revenu afin d'éviter des résultats instables comme dans le cas des modèles de régression de diabète. Disposer d'un échantillon plus important aurait pu renforcer la robustesse des résultats et améliorer leur généralisation à l'ensemble de la population cible.

Malgré ces limites, ce mémoire suggère quelques pistes pour comprendre comment le statut socio-économique peut moduler les effets d'une pratique sportive sur la santé en Belgique. Il apparaît évident que la pratique d'une activité physique de loisir est un levier majeur de protection contre l'obésité et le diabète. Les modèles d'analyses ont permis de souligner l'importance d'agir simultanément sur les comportements de santé et les inégalités sociales pour prévenir efficacement le diabète et l'obésité. Favoriser un cadre de vie qui rende l'activité physique accessible, soutenable et surtout désirable pour l'ensemble de la population semble constituer un enjeu pour la prévention de l'obésité et du diabète.

6 Bibliographie

- Abdi, H., & Williams, L. J. (2010). Principal component analysis [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/wics.101>] *WIREs Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>
- Beenackers, M. A., Kamphuis, C. B., Giskes, K., Brug, J., Kunst, A. E., Burdorf, A., & van Lenthe, F. J. (2012). Socioeconomic inequalities in occupational, leisure-time, and transport related physical activity among European adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 116. <https://doi.org/10.1186/1479-5868-9-116>
- Bennett, T., Savage, M., Silva, E. B., Warde, A., Gayo-Cal, M., & Wright, D. (2009, January). *Culture, Class, Distinction*. Routledge. <https://doi.org/10.4324/9780203930571>
- Biddle, S. J. H., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of Sports Sciences*, 22(8), 679–701. <https://doi.org/10.1080/02640410410001712412>
- Bodin, D., Héas, S., & Robène, L. (2004). Les goûts sportifs : Entre distinction et pratique électorale raisonnée. *Sociologie et Sociétés*, 36(1), 187. Retrieved June 17, 2025, from <https://shs.hal.science/halshs-00346164>
- Borraccino, A., Lemma, P., Iannotti, R., Zambon, A., Dalmasso, P., Lazzeri, G., Giacchi, M., & Cavallo, F. (2009). Socio-economic effects on meeting PA guidelines: Comparisons among 32 countries. *Medicine and science in sports and exercise*, 41(4), 749–756. <https://doi.org/10.1249/MSS.0b013e3181917722>
- Bourdieu, P. (1972). *Esquisse d'une théorie de la pratique. Précédé de « Trois études d'ethnologie kabyle »* [ISSN: 1424-6201]. Librairie Droz. <https://doi.org/10.3917/droz.bourd.1972.01>
- Bourdieu, P. (1978). Sport and social class [Publisher: SAGE Publications Ltd]. *Social Science Information*, 17(6), 819–840. <https://doi.org/10.1177/053901847801700603>
- Bourdieu, P. (1979). *La distinction : Critique sociale du jugement*. Paris : Éditions de Minuit. Retrieved July 31, 2025, from <http://archive.org/details/ladistinctioncri0000bour>
- Bourdieu, P. (1984). Espace social et genèse des "classes" [Publisher: Persée - Portail des revues scientifiques en SHS]. <https://doi.org/10.3406/arss.1984.3327>
- Breuer, C., Hallmann, K., Wicker, P., & Feiler, S. (2010). Socio-economic patterns of sport demand and ageing [Number: 2 Publisher: BioMed Central]. *European Review of Aging and Physical Activity*, 7(2), 61–70. <https://doi.org/10.1007/s11556-010-0066-5>
- Chinn, D. J., White, M., Harland, J., Drinkwater, C., & Raybould, S. (1999). Barriers to physical activity and socioeconomic position: Implications for health promotion. *Journal of Epidemiology and Community Health*, 53(3), 191–192. <https://doi.org/10.1136/jech.53.3.191>
- Claron, C., & Lagarde, M. (2024). Sports et classes sociales : Les inégalités en pratique [Publisher: La Découverte]. *Regards croisés sur l'économie*, 35(2), 165–171. <https://doi.org/10.3917/rce.035.0165>
- de Boer, W., Corpeleijn, E., Dekker, L., Mierau, J., & Koning, R. (2021). How is sport participation related to mortality, diabetes and prediabetes for different body mass index levels? [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/sms.13940>] *Scandinavian Journal of Medicine & Science in Sports*, 31(6), 1342–1351. <https://doi.org/10.1111/sms.13940>

- de Boer, W. I. J., Mierau, J. O., & Koning, R. H. (2023). Do differences in sport participation contribute to socioeconomic health inequalities? Evidence from the Lifelines cohort study on all-cause mortality, diabetes and obesity. *Preventive Medicine Reports*, 36, 102479. <https://doi.org/10.1016/j.pmedr.2023.102479>
- DiPietro, L., Buchner, D. M., Marquez, D. X., Pate, R. R., Pescatello, L. S., & Whitt-Glover, M. C. (2019). New scientific basis for the 2018 U.S. Physical Activity Guidelines. *Journal of Sport and Health Science*, 8(3), 197–200. <https://doi.org/10.1016/j.jshs.2019.03.007>
- Drieskens, S., Charafeddine, R., & Boone, C. (2025). *Statut pondéral. bruxelles, belgique : Sciensano* (No. D/2025.14.440/70). www.enquetesante.be
- Drieskens, S., Charafeddine, R., Vandevijvere, S., De Pauw, R., & Demarest, S. (2024). Rising socioeconomic disparities in childhood overweight and obesity in Belgium. *Archives of Public Health*, 82(1), 98. <https://doi.org/10.1186/s13690-024-01328-y>
- Drieskens, S., Charafeddine, R., & Gisele, L. (2018). *Enquête de santé 2018 : Etat nutritionnel. bruxelles, belgique : Sciensano* (No. D/2019/14.440/62.). www.enquetesante.be
- Durand, P. (2014). Capital symbolique.
- Ekelund, U., Brown, W. J., Steene-Johannessen, J., Fagerland, M. W., Owen, N., Powell, K. E., Bauman, A. E., & Lee, I.-M. (2019). Do the associations of sedentary behaviour with cardiovascular disease mortality and cancer mortality differ by physical activity level? A systematic review and harmonised meta-analysis of data from 850 060 participants. *British Journal of Sports Medicine*, 53(14), 886–894. <https://doi.org/10.1136/bjsports-2017-098963>
- Ekelund, U., Steene-Johannessen, J., Brown, W. J., Fagerland, M. W., Owen, N., Powell, K. E., Bauman, A., Lee, I.-M., Lancet Physical Activity Series 2 Executive Committee, & Lancet Sedentary Behaviour Working Group. (2016). Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet (London, England)*, 388(10051), 1302–1310. [https://doi.org/10.1016/S0140-6736\(16\)30370-1](https://doi.org/10.1016/S0140-6736(16)30370-1)
- Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., Whincup, P., Diaz, K. M., Hooker, S. P., Chernofsky, A., Larson, M. G., Spartano, N., Vasan, R. S., Dohrn, I.-M., Hagströmer, M., Edwardson, C., Yates, T., Shiroma, E., Anderssen, S. A., & Lee, I.-M. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: Systematic review and harmonised meta-analysis. *BMJ (Clinical research ed.)*, 366, l4570. <https://doi.org/10.1136/bmj.l4570>
- Eurostat. (2024). Overweight and obesity - BMI statistics. Retrieved July 30, 2025, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Overweight_and_obesity_-_BMI_statistics
- Galobardes, B., Costanza, M. C., Bernstein, M. S., Delhumeau, C., & Morabia, A. (2003). Trends in Risk Factors for Lifestyle-Related Diseases by Socioeconomic Position in Geneva, Switzerland, 1993–2000: Health Inequalities Persist. *American Journal of Public Health*, 93(8), 1302–1309. <https://doi.org/10.2105/ajph.93.8.1302>

- Gudnadottir, U., Cadmus-Bertram, L., Spicer, A., Gorzelitz, J., & Malecki, K. (2019). The relationship between occupational physical activity and self-reported vs measured total physical activity. *Preventive Medicine Reports*, 15, 100908. <https://doi.org/10.1016/j.pmedr.2019.100908>
- Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: A pooled analysis of 358 population-based surveys with 1.9 million participants [Publisher: Elsevier]. *The Lancet Global Health*, 6(10), e1077–e1086. [https://doi.org/10.1016/S2214-109X\(18\)30357-7](https://doi.org/10.1016/S2214-109X(18)30357-7)
- Hanson, M. D., & Chen, E. (2007). Socioeconomic status and health behaviors in adolescence: A review of the literature. *Journal of Behavioral Medicine*, 30(3), 263–285. <https://doi.org/10.1007/s10865-007-9098-3>
- Harrell, F. E. (2015). Binary logistic regression. In J. Harrell Frank E. (Ed.), *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis* (pp. 219–274). Springer International Publishing. https://doi.org/10.1007/978-3-319-19425-7_10
- Haut, J. (2018). Sport und soziale Ungleichheit. In *Sport in Kultur und Gesellschaft* (pp. 1–11). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53385-7_17-1
- Hill, J. O., & Peters, J. C. (1998). Environmental contributions to the obesity epidemic. *Science (New York, N.Y.)*, 280(5368), 1371–1374. <https://doi.org/10.1126/science.280.5368.1371>
- Hu, F. B., Willett, W. C., Li, T., Stampfer, M. J., Colditz, G. A., & Manson, J. E. (2004). Adiposity as compared with physical activity in predicting mortality among women [Publisher: Massachusetts Medical Society _eprint: <https://www.nejm.org/doi/pdf/10.1056/NEJMoa042135>]. *New England Journal of Medicine*, 351(26), 2694–2703. <https://doi.org/10.1056/NEJMoa042135>
- Hu, G., Eriksson, J., Barengo, N. C., Lakka, T. A., Valle, T. T., Nissinen, A., Jousilahti, P., & Tuomilehto, J. (2004). Occupational, Commuting, and Leisure-Time Physical Activity in Relation to Total and Cardiovascular Mortality Among Finnish Subjects With Type 2 Diabetes [Publisher: American Heart Association]. *Circulation*, 110(6), 666–673. <https://doi.org/10.1161/01.CIR.0000138102.23783.94>
- Jugnot, M. (2019, November). L’Habitus, Pierre Bourdieu (Fiche concept). Retrieved June 16, 2025, from <https://partageonsleco.com/2019/11/06/lhabitus-pierre-bourdieu-fiche-concept/>
- Kaleta, D., & Jegier, A. (2005). Occupational energy expenditure and leisure-time physical activity. *International Journal of Occupational Medicine and Environmental Health*, 18(4), 351–356.
- Kamphuis, C. B. M., Van Lenthe, F. J., Giskes, K., Huisman, M., Brug, J., & Mackenbach, J. P. (2008). Socioeconomic Status, Environmental and Individual Factors, and Sports Participation. *Medicine & Science in Sports & Exercise*, 40(1), 71–81. <https://doi.org/10.1249/mss.0b013e318158e467>
- Legrand, P., & Bories, D. (2006). (PDF) le choix des variables explicatives dans les modèles de régression logistique. *ResearchGate*. Retrieved April 14, 2025, from https://www.researchgate.net/publication/281834969_Le_choix_des_variables_explicatives_dans_les_modeles_de_regression_logistique
- Lindström, M., Hanson, B. S., & Östergren, P.-O. (2001). Socioeconomic differences in leisure-time physical activity: The role of social participation and social capital in shaping health related

- behaviour. *Social Science & Medicine*, 52(3), 441–451. [https://doi.org/10.1016/S0277-9536\(00\)00153-2](https://doi.org/10.1016/S0277-9536(00)00153-2)
- Magliano, D. J., Boyko, E. J., & Committee, I. D. A. 1. e. s. (2021). Global picture. In *IDF DIABETES ATLAS [Internet]. 10th edition*. International Diabetes Federation. Retrieved July 30, 2025, from <https://www.ncbi.nlm.nih.gov/books/NBK581940/>
- Mendham, A. E., Duffield, R., Marino, F., & Coutts, A. J. (2014). Small-sided games training reduces CRP, IL-6 and leptin in sedentary, middle-aged men. *European Journal of Applied Physiology*, 114(11), 2289–2297. <https://doi.org/10.1007/s00421-014-2953-3>
- Mutz, M., & Mueller, J. (2021). Social stratification of leisure time sport and exercise activities: Comparison of ten popular sports activities. *Leisure Studies*, 40(5), 597–611. <https://doi.org/10.1080/02614367.2021.1916834>
- Naik, H., Qiu, X., Brown, M., Eng, L., Pringle, D., Mahler, M., Hon, H., Tiessen, K., Thai, H., Ho, V., Gonos, C., Charow, R., Pat, V., Irwin, M., Herzog, L., Ho, A., Xu, W., Jones, J., Howell, D., & Liu, G. (2016). Socioeconomic status and lifestyle behaviours in cancer survivors: Smoking and physical activity. *Current Oncology*, 23(6), e546–e555. <https://doi.org/10.3747/co.23.3166>
- Nielsen, T. L., Wraae, K., Brixen, K., Hermann, A. P., Andersen, M., & Hagen, C. (2006). Prevalence of overweight, obesity and physical inactivity in 20- to 29-year-old, Danish men. Relation to sociodemography, physical dysfunction and low socioeconomic status: The Odense Androgen Study. *International Journal of Obesity (2005)*, 30(5), 805–815. <https://doi.org/10.1038/sj.ijo.0803197>
- Oja, P., Titze, S., Kokko, S., Kujala, U. M., Heinonen, A., Kelly, P., Koski, P., & Foster, C. (2015). Health benefits of different sport disciplines for adults: Systematic review of observational and intervention studies with meta-analysis. *British Journal of Sports Medicine*, 49(7), 434–440. <https://doi.org/10.1136/bjsports-2014-093885>
- OMS. (2022). *WHO european regional obesity report 2022*. World Health Organization.
- OMS. (2024). Diabetes. Retrieved July 30, 2025, from <https://www.who.int/europe/news-room/fact-sheets/item/diabetes>
- OMS. (2025). Obésité et surpoids. Retrieved July 30, 2025, from <https://www.who.int/fr/news-room/fact-sheets/detail/obesity-and-overweight>
- Paffenbarger, R. S., Hyde, R. T., Wing, A. L., Lee, I.-M., Jung, D. L., & Kampert, J. B. (1993). The association of changes in physical-activity level and other lifestyle characteristics with mortality among men [Publisher: Massachusetts Medical Society _eprint: <https://www.nejm.org/doi/pdf/10.1056/NEJM199302253280804>]. *New England Journal of Medicine*, 328(8), 538–545. <https://doi.org/10.1056/NEJM199302253280804>
- Pagès, J. (2004). Analyse factorielle de données mixtes [Publisher: Société française de statistique]. *Revue de Statistique Appliquée*, 52(4), 93–111. Retrieved May 6, 2025, from <https://eudml.org/doc/106558>
- Pan, M., Tu, R., Gu, J., Li, R., Liu, X., Chen, R., Yu, S., Wang, X., Mao, Z., Huo, W., Hou, J., & Wang, C. (2020). Associations of Socioeconomic Status and Physical Activity With Obesity Measures in Rural Chinese Adults. *Frontiers in Public Health*, 8, 594874. <https://doi.org/10.3389/fpubh.2020.594874>

- Pociello, C. (1987). *Sport et société* [Google-Books-ID: 57Opz2mKWJEC]. Université de Saint-Etienne.
- Rhodes, R. E., Janssen, I., Bredin, S. S., Warburton, D. E., & Bauman, A. (2017). Physical activity: Health impact, prevalence, correlates and interventions [Publisher: Routledge _eprint: <https://doi.org/10.1080/08870446.2017.1325486>]. *Psychology & Health*, 32(8), 942–975. <https://doi.org/10.1080/08870446.2017.1325486>
- Richard, V., Piumatti, G., Pullen, N., Lorthe, E., Guessous, I., Cantoreggi, N., & Stringhini, S. (2023). Socioeconomic inequalities in sport participation: Pattern per sport and time trends – a repeated cross-sectional study. *BMC Public Health*, 23(1), 785. <https://doi.org/10.1186/s12889-023-15650-7>
- Ritchie, H., & Roser, M. (2017). Obesity. *Our World in Data*. Retrieved June 19, 2025, from <https://ourworldindata.org/obesity>
- Schnohr, C., Højbjørre, L., Riegels, M., Ledet, L., Larsen, T., Schultz-Larsen, K., Petersen, L., Prescott, E., & Grønbaek, M. (2004). Does educational level influence the effects of smoking, alcohol, physical activity, and obesity on mortality? A prospective population study. *Scandinavian Journal of Public Health*, 32(4), 250–256. <https://doi.org/10.1080/14034940310019489>
- Sciensano. (2022, February). Inégalités dans les déterminants de santé. Retrieved June 15, 2025, from <https://www.belgiqueenbonnesante.be/fr/etat-de-sante/inegalites-de-sante/inegalites-dans-les-determinants-de-sante>
- Sesso, H. D., Lee, I.-M., & Paffenbarger, R. S. (1998). Physical activity and breast cancer risk in the college alumni health study (united states). *Cancer Causes & Control*, 9(4), 433–439. <https://doi.org/10.1023/A:1008827903302>
- Stempel, C. (2005). Adult Participation Sports as Cultural Capital: A Test of Bourdieu’s Theory of the Field of Sports [Publisher: SAGE Publications Ltd]. *International Review for the Sociology of Sport*, 40(4), 411–432. <https://doi.org/10.1177/1012690206066170>
- Teulon, F. J. (2024, September). *Sociologie de la pratique sportive*. Retrieved June 17, 2025, from <https://hal.science/hal-04714896>
working paper or preprint.
- Thibaut, E., Eakins, J., Vos, S., & Scheerder, J. (2017). Time and money expenditure in sports participation: The role of income in consuming the most practiced sports activities in flanders. *Sport Management Review*, 20(5), 455–467. <https://doi.org/10.1016/j.smr.2016.12.002>
- Thurston, R. C., Kubzansky, L. D., Kawachi, I., & Berkman, L. F. (2005). Is the association between socioeconomic position and coronary heart disease stronger in women than in men? *American Journal of Epidemiology*, 162(1), 57–65. <https://doi.org/10.1093/aje/kwi159>
- Van Dyck, D., Cardon, G., Deforche, B., Sallis, J. F., Owen, N., & De Bourdeaudhuij, I. (2010). Neighborhood SES and walkability are related to physical activity behavior in Belgian adults. *Preventive Medicine*, 50 Suppl 1, S74–79. <https://doi.org/10.1016/j.ypmed.2009.07.027>
- Wagner, A.-C. (2012). Habitus [Publisher: PUF]. *Sociologie*. Retrieved June 16, 2025, from <https://journals.openedition.org/sociologie/1200>
- Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: A systematic review of current systematic reviews. *Current Opinion in Cardiology*, 32(5), 541. <https://doi.org/10.1097/HCO.0000000000000437>

- Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: The evidence. *CMAJ : Canadian Medical Association Journal*, 174(6), 801–809. <https://doi.org/10.1503/cmaj.051351>
- Wilson, D. K., Kirtland, K. A., Ainsworth, B. E., & Addy, C. L. (2004). Socioeconomic status and perceptions of access and safety for physical activity. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 28(1), 20–28. https://doi.org/10.1207/s15324796abm2801_4

7 Annexes

Table 6 Résultat du test de Wilcoxon par paires (p-values ajustées) pour l'obésité

Groupe_1	1	2	3	4	5
2	0.1921				
3	0.0176	0.3336			
4	0.0003	0.0454	0.2618		
5	0.0000	0.0000	0.0005	0.0167	

Table 7 Résultat du test de Wilcoxon par paires (p-values ajustées) pour le diabète

Groupe_1	1	2	3	4	5
2	0.2966				
3	0.0185	0.0011			
4	0.0001	0.0000	0.2319		
5	0.0000	0.0000	0.1484	0.7723	

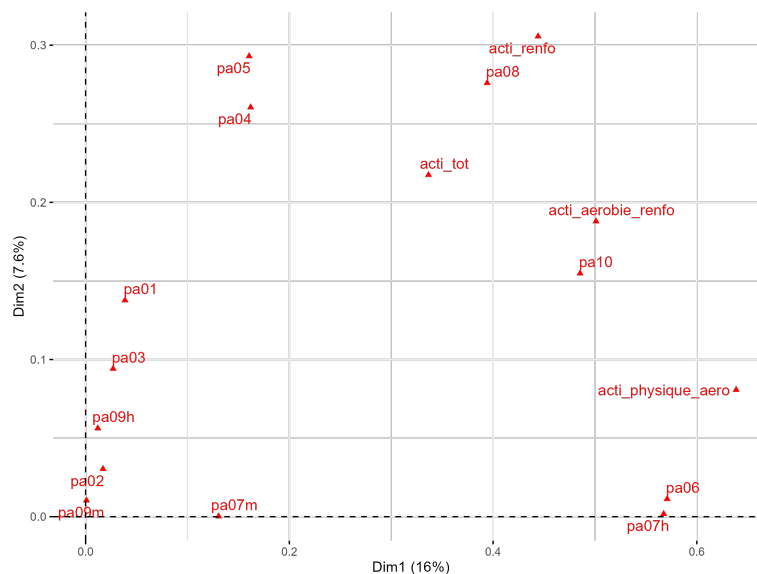


Figure 12 Contribution des variables aux deux dimensions de l'analyse factorielle mixte sur base des données HIS (2018)

Table 8 Résultat du test de khi carré entre la variable *cluster* et la variable *sexe*

Test	Valeur
Statistique χ^2 (X-squared)	65.703
Degrés de liberté (df)	4
p-value	1.83×10^{-13}

Table 9 Description statistique des clusters selon les dimensions principales : Dim.1 et Dim.2

Cluster 1 : Sédentaires					
Variable	v.test	Moy. cat.	Moy. glob.	Écart-type	p-value
Dim.2	-33.457	-0.925	-0.054	0.463	2.05×10^{-245}
Dim.1	-44.387	-1.807	-0.095	0.520	0.000
Cluster 2 : Actifs modérés					
Variable	v.test	Moy. cat.	Moy. glob.	Écart-type	p-value
Dim.2	22.360	0.596	-0.054	0.551	9.59×10^{-111}
Dim.1	-21.449	-1.019	-0.095	0.458	4.65×10^{-102}
Cluster 3 : Sportifs irréguliers					
Variable	v.test	Moy. cat.	Moy. glob.	Écart-type	p-value
Dim.1	16.911	1.065	-0.095	0.691	3.75×10^{-64}
Dim.2	-7.160	-0.385	-0.054	1.087	8.06×10^{-13}
Cluster 4 : Sportifs complets					
Variable	v.test	Moy. cat.	Moy. glob.	Écart-type	p-value
Dim.2	48.899	2.043	-0.054	0.849	0.000
Dim.1	26.941	1.616	-0.095	0.919	7.26×10^{-160}
Cluster 5 : Sportifs intensifs					
Variable	v.test	Moy. cat.	Moy. glob.	Écart-type	p-value
Dim.1	53.949	4.746	-0.095	1.319	0.000
Dim.2	-34.538	-2.146	-0.054	1.093	2.19×10^{-261}

Table 10 Résultat du test de khi carré entre les clusters et l'activité physique liée au travail

Test	Valeur
Statistique χ^2 (X-squared)	140.92
Degrés de liberté (df)	12
p-value	$< 2.2 \times 10^{-16}$

Table 11 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon l'activité physique

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0631	0.3453	-7.9999	0.0000	0.0300	0.1180
Clusters						
Sédentaires (réf.)	—	—	—	—	—	—
Actifs modérés	0.7772	0.1114	-2.2624	0.0237	0.6243	0.9664
Sportifs irréguliers	0.4676	0.1691	-4.4948	0.0000	0.3326	0.6463
Sportifs complets	0.4155	0.1562	-5.6212	0.0000	0.3038	0.5610
Sportifs intensifs	0.3233	0.2348	-4.8102	0.0000	0.1990	0.5015
Catégories d'âges						
18–24 (réf.)	—	—	—	—	—	—
25–34	0.7246	0.2885	-1.1164	0.2643	0.4181	1.3011
35–44	1.2303	0.2852	0.7266	0.4675	0.7162	2.1998
45–54	1.4982	0.2846	1.4205	0.1555	0.8736	2.6767
55–64	1.4830	0.2880	1.3683	0.1713	0.8585	2.6657
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	0.8134	0.0956	-2.1608	0.0308	0.6741	0.9808
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	1.0810	0.1364	0.5709	0.5681	0.8255	1.4097
Ville urbanisée	1.0537	0.1169	0.4476	0.6545	0.8377	1.3253
Rural	0.9355	0.1656	-0.4028	0.6871	0.6722	1.2877
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	5.9380	0.4473	3.9828	0.0001	2.5255	14.7801
Secondaire	4.8680	0.4221	3.7498	0.0002	2.1852	11.5974
Supérieur court	4.2368	0.4089	3.5309	0.0004	1.9552	9.8645
Supérieur long	2.7936	0.4141	2.4806	0.0132	1.2754	6.5644
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	1.0471	0.1270	0.3624	0.7171	0.8182	1.3467
Veuf	2.4834	0.3008	3.0238	0.0025	1.3608	4.4450
Divorcé	0.8895	0.1826	-0.6412	0.5214	0.6196	1.2685
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	0.9959	0.1975	-0.0207	0.9835	0.6670	1.4495
Hors UE	1.0253	0.1550	0.1614	0.8718	0.7520	1.3817

Table 12 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon l'activité physique et la charge de travail.

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0762	0.3660	-7.0342	0.0000	0.0350	0.1491
Clusters						
Sédentaires (réf.)	—	—	—	—	—	—
Actifs modérés	0.6361	0.1578	-2.8660	0.0042	0.4657	0.8648
Sportifs irréguliers	0.4286	0.1765	-4.7992	0.0000	0.3006	0.6012
Sportifs complets	0.3598	0.1766	-5.7894	0.0000	0.2527	0.5054
Sportifs intensifs	0.2839	0.2467	-5.1033	0.0000	0.1709	0.4514
Charge de travail						
Pas de tâche spécifique (réf.)	—	—	—	—	—	—
Souvent assis ou debout	0.8376	0.1620	-1.0943	0.2739	0.6121	1.1557
Marche ou tâche d'activité physique modéré	1.0074	0.1996	0.0372	0.9704	0.6836	1.4956
Travail dur et demandant	1.5040	0.2220	1.8384	0.0661	0.9748	2.3287
Catégories d'âges						
18–24 (réf.)	—	—	—	—	—	—
25–34	0.7460	0.2889	-1.0141	0.3106	0.4301	1.3405
35–44	1.2741	0.2855	0.8485	0.3962	0.7413	2.2792
45–54	1.5458	0.2847	1.5300	0.1261	0.9012	2.7622
55–64	1.5434	0.2888	1.5030	0.1329	0.8920	2.7779
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	0.8277	0.0962	-1.9658	0.0494	0.6853	0.9993
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	1.0531	0.1367	0.3784	0.7052	0.8037	1.3741
Ville urbanisée	1.0190	0.1176	0.1600	0.8729	0.8091	1.2832
Rural	0.9120	0.1661	-0.5546	0.5792	0.6547	1.2567
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	5.2203	0.4491	3.6799	0.0002	2.2115	13.0326
Secondaire	4.3261	0.4238	3.4560	0.0006	1.9345	10.3350
Supérieur court	3.7816	0.4106	3.2397	0.0012	1.7385	8.8276
Supérieur long	2.7103	0.4147	2.4044	0.0162	1.2358	6.3732
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	1.0547	0.1274	0.4183	0.6758	0.8235	1.3575
Veuf	2.5075	0.3014	3.0502	0.0023	1.3725	4.4932
Divorcé	0.8814	0.1826	-0.6913	0.4894	0.6138	1.2571
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	0.9830	0.1976	-0.0870	0.9307	0.6581	1.4312
Hors UE	1.0017	0.1551	0.0107	0.9914	0.7346	1.3502

Table 13 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Obésité selon le revenu Table 13

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0474	0.3732	-8.1728	0.0000	0.0210	0.0926
Quintile de revenu						
Revenu 1 (réf.)	—	—	—	—	—	—
Revenu 2	0.8730	0.1621	-0.8379	0.4022	0.6340	1.1977
Revenu 3	0.7152	0.1582	-2.1191	0.0342	0.5237	0.9743
Revenu 4	0.6607	0.1567	-2.6448	0.0082	0.4858	0.8983
Revenu 5	0.4561	0.1712	-4.5842	0.0000	0.3254	0.6371
Catégories d'âges						
18–24 (réf.)	—	—	—	—	—	—
25–34	0.8049	0.2989	-0.7261	0.4679	0.4558	1.4780
35–44	1.4277	0.2951	1.2067	0.2276	0.8168	2.6092
45–54	1.7061	0.2953	1.8092	0.0705	0.9760	3.1200
55–64	1.6110	0.2997	1.5913	0.1116	0.9130	2.9688
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	0.8654	0.1002	-1.4419	0.1494	0.7108	1.0531
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	1.0679	0.1455	0.4513	0.6518	0.8005	1.4168
Villes urbanisées	1.0415	0.1218	0.3340	0.7384	0.8201	1.3224
Rural	0.8844	0.1745	-0.7039	0.4816	0.6238	1.2377
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	6.2837	0.4752	3.8676	0.0001	2.5454	16.6869
Secondaire	5.2748	0.4481	3.7112	0.0002	2.2665	13.3768
Supérieur court	4.7013	0.4343	3.5638	0.0004	2.0803	11.6456
Supérieur long	3.3103	0.4407	2.7161	0.0066	1.4454	8.2911
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	1.1050	0.1337	0.7473	0.4549	0.8525	1.4402
Veuf	2.5824	0.3082	3.0786	0.0021	1.3946	4.6900
Divorcé	0.9132	0.1888	-0.4809	0.6306	0.6282	1.3183
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	0.9982	0.2091	-0.0087	0.9930	0.6522	1.4838
Hors UE	0.8576	0.1683	-0.9131	0.3613	0.6124	1.1855

Table 14 Odds ratios, erreurs standards, z, p-values et intervalles de confiance à 95% :
Obésité selon l'activité physique et le revenu.

Variable	OR	Erreur Std.	z	p-value	IC 2.5%	IC 97.5%
(Intercept)	0.0780	0.4120	-6.1929	0.0000	0.0325	0.1660
Cluster						
Sédentaires (réf.)	-	-	-	-	-	-
Actifs modérés	0.6159	0.2667	-1.8170	0.0693	0.3634	1.0355
Sportifs irréguliers	0.3387	0.4907	-2.2066	0.0274	0.1160	0.8220
Sportifs complets	0.3202	0.3891	-2.9271	0.0034	0.1421	0.6617
Sportifs intensifs	0.6104	0.4737	-1.0420	0.2975	0.2256	1.4761
Quintile de revenu						
revenu1 (réf.)	-	-	-	-	-	-
revenu2	1.0130	0.2429	0.0531	0.9577	0.6269	1.6273
revenu3	0.5980	0.2613	-1.9671	0.0493	0.3552	0.9920
revenu4	0.7044	0.2451	-1.4292	0.1531	0.4343	1.1370
revenu5	0.5397	0.2573	-2.3968	0.0166	0.3240	0.8901
Charge de travail						
Pas de tâche spécifique (réf.)	-	-	-	-	-	-
Souvent assis ou debout	0.9039	0.1766	-0.5721	0.5673	0.6421	1.2840
Souvent marche ou tâche modérée	1.0689	0.2150	0.3101	0.7565	0.7041	1.6366
Travail dur et demandant	1.6509	0.2405	2.0846	0.0372	1.0323	2.6519
Catégorie d'âge						
18-24 ans (réf.)	-	-	-	-	-	-
25-34 ans	0.7624	0.3075	-0.8822	0.3777	0.4245	1.4240
35-44 ans	1.3277	0.3031	0.9353	0.3497	0.7477	2.4649
45-54 ans	1.5312	0.3027	1.4073	0.1594	0.8631	2.8414
55-64 ans	1.4905	0.3076	1.2974	0.1946	0.8315	2.7899
Sexe						
Homme (réf.)	-	-	-	-	-	-
Femme	0.8350	0.1028	-1.7534	0.0796	0.6823	1.0213
Milieu de résidence						
Grande ville (réf.)	-	-	-	-	-	-
Banlieue	1.0477	0.1488	0.3134	0.7540	0.7804	1.3991
Ville urbanisée	1.0064	0.1246	0.0511	0.9593	0.7881	1.2848
Rural	0.8747	0.1789	-0.7482	0.4544	0.6118	1.2351
Niveau de diplôme						
Autres (réf.)	-	-	-	-	-	-
Primaire	5.3112	0.4862	3.4346	0.0006	2.1019	14.3771
Secondaire	4.8736	0.4587	3.4530	0.0006	2.0465	12.5848
Supérieur court	4.4325	0.4449	3.3468	0.0008	1.9164	11.1770
Supérieur long	3.4531	0.4499	2.7547	0.0059	1.4779	8.7840
Statut matrimonial						
Célibataire (réf.)	-	-	-	-	-	-
Marié ou cohabitant légal	1.1421	0.1362	0.9756	0.3293	0.8768	1.4962
Veuf	2.8328	0.3201	3.2532	0.0012	1.4973	5.2763
Divorcé	0.9105	0.1918	-0.4888	0.6250	0.6228	1.3222
Origine						
Belgique (réf.)	-	-	-	-	-	-

Table 14 – suite

Variable	OR	Erreur Std.	z	p-value	IC 2.5%	IC 97.5%
EU	0.9876	0.2125	-0.0585	0.9533	0.6413	1.4788
Hors EU	0.8336	0.1716	-1.0605	0.2890	0.5917	1.1602
Interaction clusters : revenu						
Sédentaires : revenu1 (réf.)	-	-	-	-	-	-
Actifs modérés : revenu2	0.6444	0.3776	-1.1636	0.2447	0.3057	1.3466
Sportifs irréguliers : revenu2	1.3100	0.6757	0.3996	0.6895	0.3471	5.1122
Sportifs complets : revenu2	1.3234	0.5327	0.5261	0.5989	0.4664	3.8166
Sportifs intensifs : revenu2	0.3553	0.9583	-1.0799	0.2803	0.0377	1.9792
Actifs modérés : revenu3	1.2473	0.3675	0.6014	0.5476	0.6078	2.5719
Sportifs irréguliers : revenu3	4.0226	0.6127	2.2718	0.0232	1.2595	14.3137
Sportifs complets : revenu3	0.8324	0.5670	-0.3236	0.7462	0.2684	2.5306
Sportifs intensifs : revenu3	1.4892	0.6351	0.6270	0.5307	0.4282	5.2837
Actifs modérés : revenu4	1.1631	0.3403	0.4441	0.6570	0.5975	2.2715
Sportifs irréguliers : revenu4	0.7948	0.6381	-0.3599	0.7190	0.2295	2.9147
Sportifs complets : revenu4	1.1988	0.5012	0.3617	0.7176	0.4528	3.2729
Sportifs intensifs : revenu4	0.3593	0.7416	-1.3801	0.1677	0.0752	1.4742
Actifs modérés : revenu5	0.8943	0.3930	-0.2841	0.7764	0.4101	1.9215
Sportifs irréguliers : revenu5	1.2505	0.5960	0.3750	0.7077	0.4038	4.3133
Sportifs complets : revenu5	1.5900	0.5163	0.8982	0.3691	0.5806	4.4559
Sportifs intensifs : revenu5	0.0829	1.1268	-2.2102	0.0272	0.0041	0.5410

Table 15 Prédictions marginales moyennes, intervalles de confiance à 95% et p-values :
Obésité selon l'activité physique et le revenu.

Variable	Modalité	Marge prédite	IC 95%	p-value
Charge de travail	Pas de tâche spécifique	14.7%	10.8%, 18.6%	<0.001
	Souvent assis ou debout	13.6%	11.6%, 15.5%	<0.001
	Souvent marche ou tâche modérée	15.9%	12.8%, 18.9%	<0.001
	Travail dur et demandant	22.1%	15.9%, 28.3%	<0.001
Catégories d'âges	18-24	12.5%	5.7%, 19.2%	<0.001
	25-34	9.9%	7.1%, 12.7%	<0.001
	35-44	15.7%	12.8%, 18.7%	<0.001
	45-54	17.5%	14.9%, 20.2%	<0.001
	55-64	17.2%	14.4%, 20.0%	<0.001
Sexe	Homme	16.3%	14.3%, 18.3%	<0.001
	Femme	14.2%	12.5%, 15.9%	<0.001
Milieu de résidence	Grandes villes	15.2%	13.0%, 17.4%	<0.001
	Banlieue	15.8%	12.3%, 19.3%	<0.001
	Villes urbanisées	15.4%	13.0%, 17.8%	<0.001
	Rural	13.7%	9.9%, 17.6%	<0.001
Niveau de diplôme	Autres	4.6%	0.6%, 8.6%	0.025
	Primaire	19.3%	13.2%, 25.4%	<0.001
	Secondaire	18.1%	14.2%, 21.9%	<0.001
	Supérieur court	16.8%	14.4%, 19.2%	<0.001
	Supérieur long	13.8%	11.5%, 16.1%	<0.001
Statut matrimonial	Célibataire	14.1%	11.3%, 16.9%	<0.001
	Marié ou cohabitant légal	15.7%	13.9%, 17.5%	<0.001
	Veuf	30.0%	20.2%, 39.8%	<0.001
	Divorcé	13.1%	9.6%, 16.6%	<0.001
Origine	Belgique	15.5%	13.9%, 17.1%	<0.001
	EU	15.4%	10.7%, 20.1%	<0.001
	Hors EU	13.4%	10.0%, 16.8%	<0.001
clusters * revenu	1 * 1	27.8%	20.7%, 34.8%	<0.001
	1 * 2	27.8%	19.2%, 36.4%	<0.001
	1 * 3	18.9%	12.9%, 25.0%	<0.001
	1 * 4	21.2%	15.0%, 27.4%	<0.001
	1 * 5	17.6%	11.4%, 23.9%	<0.001
	2 * 1	18.8%	12.3%, 25.3%	<0.001
	2 * 2	13.3%	8.2%, 18.4%	<0.001
	2 * 3	14.8%	9.7%, 19.8%	<0.001
	2 * 4	15.7%	10.2%, 21.2%	<0.001

Variable	Modalité	Marge prédite	IC 95%	p-value
	2 * 5	9.9%	5.5%, 14.3%	<0.001
	3 * 1	11.4%	4.3%, 18.6%	0.002
	3 * 2	15.8%	2.0%, 29.6%	0.025
	3 * 3	23.6%	11.2%, 36.0%	<0.001
	3 * 4	6.5%	2.5%, 10.5%	0.001
	3 * 5	7.5%	3.3%, 11.8%	<0.001
	4 * 1	11.0%	3.2%, 18.8%	0.006
	4 * 2	13.6%	6.3%, 20.8%	<0.001
	4 * 3	5.9%	1.5%, 10.3%	0.008
	4 * 4	10.9%	3.4%, 18.3%	0.004
	4 * 5	10.5%	4.8%, 16.3%	<0.001
	5 * 1	18.8%	-0.5%, 38.1%	0.056
	5 * 2	8.1%	-1.6%, 17.9%	0.10
	5 * 3	17.2%	2.5%, 31.9%	0.022
	5 * 4	5.7%	0.8%, 10.5%	0.021
	5 * 5	1.1%	-0.8%, 3.0%	0.3

Remarque : Pour la variable « clusters * revenu », le symbole « * » représente l'interaction entre les variables « clusters » et « revenu ». Pour la variable clusters : 1 = Sédentaires, 2 = Actifs modérés, 3 = Sportifs irréguliers, 4 = Sportifs complets, 5 = Sportifs intensifs. Pour la variable revenus : 1 = revenu1, 2 = revenu2, 3 = revenu3, 4 = revenu4, 5 = revenu5.

Table 16 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0016	0.8459	-7.6039	0.0000	0.0003	0.0084
Clusters						
Sédentaires (réf.)	—	—	—	—	—	—
Actifs modérés	0.8441	0.2131	-0.7951	0.4266	0.5558	1.2820
Sportifs irréguliers	0.6749	0.4073	-0.9654	0.3344	0.3037	1.4998
Sportifs complets	0.3472	0.3409	-3.1025	0.0019	0.1780	0.6775
Sportifs intensifs	0.2142	0.5741	-2.6842	0.0073	0.0695	0.6601
Catégories d'âges						
18–24 (réf.)	—	—	—	—	—	—
25–34	11.1165	0.7009	3.4363	0.0006	2.8130	43.9298
35–44	13.0521	0.7549	3.4029	0.0007	2.9708	57.3446
45–54	20.7738	0.7246	4.1869	0.0000	5.0183	85.9955
55–64	65.3754	0.7168	5.8319	0.0000	16.0359	266.5237
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	0.9011	0.2019	-0.5156	0.6062	0.6065	1.3388
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	1.1076	0.3000	0.3406	0.7334	0.6151	1.9943
Ville urbanisée	1.3915	0.2568	1.2862	0.1984	0.8410	2.3023
Rural	2.2118	0.2984	2.6604	0.0078	1.2322	3.9703
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	1.2973	0.7189	0.3621	0.7173	0.3169	5.3109
Secondaire	1.4403	0.6547	0.5573	0.5774	0.3990	5.1990
Supérieur court	0.9521	0.6289	-0.0780	0.9378	0.2774	3.2676
Supérieur long	0.5989	0.6478	-0.7915	0.4287	0.1682	2.1325
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	0.9210	0.2499	-0.3291	0.7421	0.5642	1.5034
Veuf	0.8321	0.4535	-0.4052	0.6853	0.3420	2.0246
Divorcé	1.0515	0.3021	0.1662	0.8680	0.5815	1.9013
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	1.2187	0.3691	0.5358	0.5921	0.5910	2.5131
Hors UE	2.9826	0.2505	4.3620	0.0000	1.8251	4.8743

Table 17 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique et la charge de travail

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0022	0.8636	-7.0611	0.0000	0.0004	0.0122
Clusters						
Sédentaires (réf.)	—	—	—	—	—	—
Actifs modérés	0.9019	0.4108	-0.2514	0.8015	0.4030	2.0181
Sportifs irréguliers	0.7293	0.4390	-0.7192	0.4721	0.3084	1.7245
Sportifs complets	0.3736	0.4523	-2.1769	0.0296	0.1539	0.9068
Sportifs intensifs	0.2266	0.6381	-2.3266	0.0200	0.0648	0.7917
Charge de travail						
Pas de tâche spécifique (réf.)	—	—	—	—	—	—
Souvent assis ou debout	0.6225	0.2521	-1.8807	0.0601	0.3798	1.0204
Marche ou tâche d'activité physique modéré	0.7967	0.4290	-0.5298	0.5963	0.3436	1.8475
Travail dur et demandant	0.3527	0.5096	-2.0450	0.0409	0.1299	0.9579
Catégories d'âges						
18-24 (réf.)	—	—	—	—	—	—
25-34	11.6655	0.7073	3.4735	0.0005	2.9153	46.6790
35-44	13.2635	0.7624	3.3907	0.0007	2.9750	59.1317
45-54	21.3241	0.7340	4.1687	0.0000	5.0568	89.9226
55-64	61.2652	0.7240	5.6839	0.0000	14.8159	253.3384
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	0.8658	0.1999	-0.7207	0.4712	0.5850	1.2814
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	1.0958	0.3057	0.2992	0.7648	0.6017	1.9955
Ville urbanisée	1.3839	0.2531	1.2835	0.1994	0.8425	2.2732
Rural	2.1869	0.2999	2.6095	0.0091	1.2148	3.9368
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	1.2812	0.7299	0.3396	0.7342	0.3063	5.3591
Secondaire	1.4280	0.6735	0.5289	0.5969	0.3813	5.3482
Supérieur court	0.9562	0.6388	-0.0701	0.9441	0.2733	3.3460
Supérieur long	0.6117	0.6599	-0.7450	0.4563	0.1677	2.2304
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	0.9493	0.2434	-0.2138	0.8307	0.5890	1.5298
Veuf	0.8291	0.4607	-0.4068	0.6842	0.3360	2.0460
Divorcé	1.0715	0.3000	0.2301	0.8181	0.5950	1.9293
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	1.2401	0.3685	0.5840	0.5593	0.6021	2.5540
Hors UE	2.9958	0.2542	4.3162	0.0000	1.8199	4.9312

Table 18 Odds ratios (OR), erreurs standards, z-statistiques, p-values et intervalles de confiance à 95% : Diabète selon le revenu

Variable	OR	SE	z	p-value	2.5%	97.5%
(Intercept)	0.0009	1.0075	-6.9165	0.0000	0.0001	0.0068
Catégorie de revenu						
Revenu bas et moyen (réf.)	—	—	—	—	—	—
Haut revenu	0.7217	0.2417	-1.3492	0.1774	0.4493	1.1593
Catégories d'âges						
18–24 (réf.)	—	—	—	—	—	—
25–34	7.9227	0.5528	3.7440	0.0002	2.6800	23.4207
35–44	9.1987	0.5868	3.7818	0.0002	2.9113	29.0648
45–54	13.2465	0.5619	4.5983	0.0000	4.4019	39.8622
55–64	41.0496	0.5451	6.8146	0.0000	14.0972	119.5319
Sexe						
Homme (réf.)	—	—	—	—	—	—
Femme	1.0040	0.2122	0.0186	0.9851	0.6623	1.5219
Milieu de résidence						
Grande ville (réf.)	—	—	—	—	—	—
Banlieue	0.9814	0.3302	-0.0570	0.9546	0.5136	1.8750
Ville urbanisée	1.4181	0.2652	1.3171	0.1879	0.8431	2.3852
Rural	2.3516	0.3134	2.7285	0.0064	1.2721	4.3473
Niveau de diplôme						
Autres (réf.)	—	—	—	—	—	—
Primaire	2.9657	0.6628	1.6403	0.1010	0.8087	10.8761
Secondaire	3.0455	0.6021	1.8496	0.0645	0.9353	9.9162
Supérieur court	1.9052	0.5691	1.1327	0.2574	0.6242	5.8150
Supérieur long	1.2945	0.5922	0.4359	0.6629	0.4054	4.1336
Statut matrimonial						
Célibataire (réf.)	—	—	—	—	—	—
Marié ou cohabitant légal	0.9279	0.2601	-0.2878	0.7735	0.5571	1.5453
Veuf	0.8918	0.4557	-0.2513	0.8016	0.3650	2.1791
Divorcé	1.1928	0.3146	0.5604	0.5752	0.6437	2.2105
Origine						
Belgique (réf.)	—	—	—	—	—	—
UE	1.2025	0.4060	0.4542	0.6497	0.5425	2.6656
Hors UE	2.6465	0.2709	3.5923	0.0003	1.5559	4.5017

Table 19 Odds ratios, erreurs standards, z, p-values et intervalles de confiance à 95% : Diabète selon l'activité physique et le revenu.

Variable	OR	Erreur Std.	z	p-value	IC 2.5%	IC 97.5%
(Intercept)	0.0027	0.9995	-5.9281	0.0000	0.0004	0.0190
Cluster						
Sédentaires (réf.)	-	-	-	-	-	-
Actifs modérés	0.4876	0.4636	-1.5493	0.1214	0.1965	1.2101
Sportifs irréguliers	0.2901	0.5656	-2.1878	0.0288	0.0957	0.8795
Sportifs complets	0.2986	0.5747	-2.1031	0.0355	0.0968	0.9214
Sportifs intensifs	0.1802	0.7707	-2.2234	0.0263	0.0398	0.8167
Quintile de revenu						
Bas et moyen revenu (réf.)	-	-	-	-	-	-
Haut revenu	0.4250	0.3720	-2.3001	0.0215	0.2049	0.8814
Charge de travail						
Pas de tâche spécifique (réf.)	-	-	-	-	-	-
Souvent assis ou debout	0.6713	0.2838	-1.4045	0.1603	0.3848	1.1710
Souvent marche ou tâche modérée	1.0441	0.4618	0.0935	0.9255	0.4222	2.5820
Travail dur et demandant	0.4854	0.5739	-1.2593	0.2080	0.1576	1.4955
Catégorie d'âge						
18-24 ans (réf.)	-	-	-	-	-	-
25-34	8.0833	0.5337	3.9154	0.0001	2.8386	23.0183
35-44	8.5234	0.5778	3.7087	0.0002	2.7455	26.4602
45-54	12.6270	0.5478	4.6294	0.0000	4.3139	36.9599
55-64	36.2187	0.5319	6.7490	0.0000	12.7656	102.7604
Sexe						
Homme (réf.)	-	-	-	-	-	-
Femme	0.9114	0.2095	-0.4430	0.6578	0.6044	1.3743
Milieu de résidence						
Grande ville (réf.)	-	-	-	-	-	-
Banlieue	0.9817	0.3391	-0.0545	0.9565	0.5050	1.9085
Ville urbanisée	1.4457	0.2694	1.3684	0.1713	0.8525	2.4516
Rural	2.1919	0.3187	2.4623	0.0139	1.1734	4.0947
Niveau de diplôme						
Aucun diplôme (réf.)	-	-	-	-	-	-
Primaire	2.0127	0.6729	1.0396	0.2986	0.5381	7.5291
Secondaire	2.3161	0.6039	1.3907	0.1644	0.7087	7.5687
Supérieur court	1.5401	0.5604	0.7707	0.4409	0.5133	4.6210
Supérieur long	1.1207	0.5722	0.1991	0.8422	0.3650	3.4410
Statut matrimonial						
Célibataire (réf.)	-	-	-	-	-	-

Table 19 – suite

Variable	OR	Erreur Std.	z	p-value	IC 2.5%	IC 97.5%
Marié ou cohabitant légal	1.0019	0.2552	0.0075	0.9940	0.6075	1.6525
Veuf	0.8172	0.4822	-0.4186	0.6755	0.3175	2.1036
Divorcé	1.2436	0.3118	0.6992	0.4845	0.6748	2.2918
Origine						
Belgique (réf.)	-	-	-	-	-	-
EU	1.2289	0.4154	0.4962	0.6198	0.5442	2.7751
Hors EU	2.5768	0.2735	3.4613	0.0005	1.5074	4.4048
Interaction clusters : revenu						
Sédentaires : Bas et moyen revenu (réf.)	-	-	-	-	-	-
Actifs modérés : Haut revenu	2.4289	0.4666	1.9019	0.0573	0.9729	6.0635
Sportifs irréguliers : Haut revenu	5.7613	0.7500	2.3348	0.0196	1.3239	25.0707
Sportifs complets : Haut revenu	1.0925	0.7892	0.1121	0.9107	0.2325	5.1341
Sportifs intensifs : Haut revenu	1.8269	1.1614	0.5189	0.6039	0.1874	17.8083

**Table 20 Prédictions marginales moyennes, intervalles de confiance à 95% et p-values :
Diabète selon l'activité physique et le revenu.**

Variable	Modalité	Marge prédite	IC 95%	p-value
Charge de travail	Pas de tâche spécifique	4.1%	2.1%, 6.0%	<0.001
	Souvent assis ou debout	2.8%	1.9%, 3.8%	<0.001
	Souvent marche ou tâche modérée	4.2%	2.2%, 6.3%	<0.001
	Travail dur et demandant	2.1%	0.4%, 3.8%	0.018
Catégories d'âges	18-24	0.2%	0.0%, 0.5%	0.052
	25-34	1.8%	0.8%, 2.8%	<0.001
	35-44	1.9%	0.9%, 3.0%	<0.001
	45-54	2.8%	1.8%, 3.8%	<0.001
	55-64	7.4%	5.3%, 9.4%	<0.001
Sexe	Homme	3.6%	2.6%, 4.5%	<0.001
	Femme	3.3%	2.5%, 4.1%	<0.001
Milieu de résidence	Grandes villes	2.7%	1.8%, 3.7%	<0.001
	Banlieue	2.7%	1.4%, 4.0%	<0.001
	Ville urbanisée	3.8%	2.6%, 5.1%	<0.001
	Rural	5.5%	3.3%, 7.8%	<0.001
Niveau de diplôme	Autres	2.3%	-0.1%, 4.7%	0.058
	Primaire	4.4%	1.5%, 7.2%	0.003
	Secondaire	4.9%	3.1%, 6.8%	<0.001
	Supérieur court	3.4%	2.5%, 4.4%	<0.001
	Supérieur long	2.5%	1.6%, 3.5%	<0.001
Statut matrimonial	Célibataire	3.3%	2.0%, 4.6%	<0.001
	Marié ou cohabitant légal	3.3%	2.5%, 4.1%	<0.001
	Veuf	2.8%	0.6%, 4.9%	0.011
	Divorcé	4.0%	2.5%, 5.6%	<0.001
Origine	Belgique	3.0%	2.3%, 3.6%	<0.001
	EU	3.6%	1.1%, 6.1%	0.005
	Hors EU	6.9%	4.2%, 9.6%	<0.001

Variable	Modalité	Marge prédite	IC 95%	p-value
clusters * revenu	1 * Bas et moyen revenu	6.2%	2.7%, 9.8%	<0.001
	1 * Haut revenu	2.9%	1.2%, 4.5%	<0.001
	2 * Bas et moyen revenu	3.2%	1.9%, 4.6%	<0.001
	2 * Haut revenu	3.3%	1.4%, 5.3%	<0.001
	3 * Bas et moyen revenu	2.0%	0.2%, 3.7%	0.026
	3 * Haut revenu	4.6%	0.7%, 8.4%	0.019
	4 * Bas et moyen revenu	2.0%	0.5%, 3.5%	0.007
	4 * Haut revenu	1.0%	-0.2%, 2.1%	0.094
	5 * Bas et moyen revenu	1.3%	-0.3%, 2.8%	0.12
	5 * Haut revenu	1.0%	-0.7%, 2.7%	0.2

Remarque : Pour la variable « clusters * revenu », le symbole « * » représente l'interaction entre les variables « clusters » et « revenu ». Pour la variable clusters : 1 = Sédentaires, 2 = Actifs modérés, 3 = Sportifs irréguliers, 4 = Sportifs complets, 5 = Sportifs intensifs.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des sciences économiques, sociales, politiques et de communication

École des sciences politiques et sociales (PSAD)

Place Montesquieu, 1 bte L2.08.05, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/psad