



UNIVERSITÉ DE MONTPELLIER

MASTER 2 BIOSTATISTIQUES

HMMA414 STAGE

Préférences, barrières et facilitateurs à l'activité physique en prévention primaire et tertiaire d'un cancer du sein

Mémoire de :

Fanchon HERMAN

Tuteurs :

Lidia DELRIEU

Baptiste FOURNIER

Hwayoung NOH

Établissement :

Centre Léon Bérard

28 rue Laennec, 69008 Lyon

Jury :

Nicolas MOLINARI

Joseph SALMON

Gilles DUCHARME

Elodie BRUNEL

Année universitaire 2020 - 2021

Remerciements

Je souhaite adresser mes remerciements aux personnes qui m'ont permis de réaliser mon stage dans les meilleures conditions.

Tout d'abord, je remercie la professeure Béatrice Fervers, directrice du département Prévention Cancer Environnement du Centre Léon Bérard, pour m'avoir accueillie dans son équipe. De plus, je souhaite remercier Élodie Brunel, ma tutrice à l'Université de Montpellier, pour son investissement et son suivi tout au long du semestre. Le master de Biostatistiques m'a permis d'acquérir et de développer d'innombrables connaissances et compétences sur le domaine des mathématiques et des statistiques.

Je souhaiterais tout particulièrement remercier Lidia Delrieu et Baptiste Fournier chercheurs en activité physique et Hwayoung Noh épidémiologiste pour leur investissement et leurs estimables conseils tout au long de mon stage.

Pour finir, je souhaite adresser ma reconnaissance à l'équipe du département Prévention Cancer Environnement pour leur accueil et leur bienveillance. Leur bonne humeur et leur dynamisme ont contribué à créer un environnement de travail stimulant et agréable.

Résumé

En France, environ 58 459 nouveaux cas de cancer du sein ont été diagnostiqués en 2018 et il reste le cancer le plus meurtrier chez les femmes avec 12 146 décès en 2018. Selon le Centre International de Recherche sur le Cancer (CIRC), en France, 37% des cancers du sein seraient liés à notre mode de vie et à notre environnement. Parmi les principales causes de cancers évitables figure une pratique d'activité physique régulière. L'objectif de ce rapport est d'identifier les préférences, les barrières et les facteurs facilitants l'activité physique auprès de femmes incluses dans la communauté des Seintinelles en prévention primaire et tertiaire du cancer du sein. Ces résultats contribueront à mieux comprendre les comportements des femmes avec et sans historique de cancer du sein pour leur proposer des interventions adaptées et améliorer les recommandations actuelles.

*L'ensemble du code réalisé lors de ce stage est présent sur **GitHub** à l'adresse suivante : <https://github.com/fanchonherman/Stage-Centre-Leon-Berard>.*

Table des matières

Remerciements	i
Résumé	i
1 Introduction	1
1 Structure d'accueil	1
1.1 Centre Léon Bérard	1
1.2 Département Prévention Cancer Environnement	2
2 Plateforme de recherche collaborative des Seintinelles sur le cancer	3
3 Cancer du sein	4
3.1 Épidémiologie	4
3.2 Traitements contre le cancer	4
3.3 Facteurs de risques	4
4 Activité physique	5
5 Problématique	7
2 Données et méthodes	8
1 Contexte de l'étude	8
1.1 Population	8
1.2 Questionnaire	8
1.3 Gestion des données	10
2 Méthode de description de la population	12
3 Méthodes de partitionnement	12
3.1 Analyse des Correspondances Multiples	13
3.2 Classification Ascendante Hiérarchique	19
3.3 K-means	20
3 Résultats	21
1 Étude descriptive des participantes	21
1.1 Caractéristiques de la population	21
1.2 Préférences, barrières et facilitateurs de la population	24
1.3 Population des Seintinelles	28
2 Partitionnement des participantes	30
2.1 Caractéristiques des classes	33
2.2 Étude de l'activité physique	35
4 Discussion	40

Conclusion	42
Annexes	46
A Questionnaire	47
B Démonstrations des formules mathématiques	67
C Graphiques et tableaux annexes	68
D Indices de pénibilités de Capgemini	74

Liste des abréviations

ACM Analyse des Correspondances Multiples

AIC Critère d'information d'Akaike

AP Activité Physique

CAH Classification Ascendante Hiérarchique

CIRC Centre International de Recherche sur le Cancer

CLB Centre Léon Bérard

CLCC Centre de Lutte Contre le Cancer

GSLTPAQ Godin-Shephard Leisure-Time Physical Activity Questionnaire

HCPC Hierarchical Clustering on Principal Components

IMC Indice de Masse Corporelle

LSI Leisure Score Index

OMS Organisation Mondiale de la Santé

Table des figures

1.1	Organigramme du Département Prévention Cancer Environnement [9].	2
1.2	Délais de recrutements avec/sans la plateforme Seintinelles [25]. . .	3
1.3	Bénéfices de la pratique d'une activité physique pendant et après un cancer [10].	6
2.1	Flow chart de la population d'étude.	11
2.2	Tableau de données et notations.	14
2.3	Tableau de fréquences et notations.	14
2.4	Traduction du tableau logique disjonctif complet en espaces métriques.	16
2.5	Dendrogramme issu d'une Classification Ascendante Hiérarchique (CAH).	20
3.1	Âge des participantes en fonction du statut du cancer du sein. . . .	22
3.2	Niveau d'activité des participantes en fonction du statut du cancer.	22
3.3	Pourcentage de participantes en fonction des départements.	24
3.4	Type d'activité physique pratiqué par les participantes.	25
3.5	Barrières à l'activité physique en fonction du statut du cancer. . . .	26
3.6	Facilitateurs à l'activité physique en fonction du statut du cancer. .	27
3.7	Représentation des indices de pénibilités, 1ère partie (France vs population des Seintinelles).	29
3.8	Représentation des indices de pénibilités, 2ème partie (France vs population des Seintinelles).	29
3.9	Graphique des pourcentages d'inertie selon les axes.	31
3.10	Dendrogramme issus de la méthode Hierarchical Clustering on Principal Components (HCPC).	33
3.11	Violin plot de l'âge des participantes en fonction des classes obtenues.	35
3.12	Représentation graphique des OR obtenus pour la régression logistique multinomiale multivariée.	39
1	Questionnaire de l'étude.	47
2	Violin plot du nombre de maladie en fonction du statut du cancer. .	68
3	Niveau d'éducation des participantes en fonction du statut du cancer.	68
4	De qui les participantes souhaitent obtenir des informations sur la pratique d'une activité physique (en %).	69
5	Barrières à l'activité physique en fonctions des classes issues du partitionnement.	71
6	Facilitateurs à l'activité physique en fonctions des classes issues du partitionnement.	72
7	Descriptions des indices de pénibilités de Capgemini, 1ère partie. . .	74
8	Descriptions des indices de pénibilités de Capgemini, 2ème partie. .	75

Liste des tableaux

1.1	Diminution (en %) du risque de développer des cancers par le biais d'une activité physique régulière [10].	5
2.1	Pourcentage de valeurs manquantes.	12
3.1	Caractéristiques des participantes en fonction des sous-groupes.	23
3.2	Premières valeurs propres issues de l'Analyse des Correspondances Multiples (ACM) et pourcentage d'inerties cumulées.	30
3.3	Indice de Rand : paires d'éléments groupés et séparés dans deux partitions P_1 et P_2 [6].	32
3.4	Principales caractéristiques des classes de participantes.	34
3.5	Odds-Ratio des caractéristiques des participantes pour l'appartenance aux classes par rapport à la classe 2 (les résultats en bleu correspondent à des p -values inférieures à 5%).	38
3.6	Modèles de régression logistique multinomiales et Critère d'information d'Akaike (AIC) correspondant.	38
1	Préférences des participantes sur la pratique d'une activité physique. . .	70
2	Utilisation des technologies (% de participantes ayant répondu oui). . . .	73

Chapitre 1

Introduction

1 Structure d'accueil

Dans cette partie, nous allons tout d'abord présenter l'établissement qui m'a permis d'effectuer mon stage ainsi que la plateforme dans laquelle l'étude a été publiée afin d'inclure des participants. Ensuite, nous vous présentons le cadre du projet en définissant les termes cancer du sein et activité physique avant d'énoncer la problématique qui nous guidera tout au long de ce rapport.

1.1 Centre Léon Bérard

Le Centre Léon Bérard (CLB) à Lyon dans la région Auvergne-Rhône-Alpes est un Centre de Lutte Contre le Cancer (CLCC) membre de la fédération nationale des centres de lutte contre le cancer UNICANCER. Le premier centre anticancéreux Léon Bérard de Lyon a été inauguré en 1923 et le Centre Léon Bérard en 1958 [11].

Il dispose de trois missions principales au service de la cancérologie : les soins, la recherche et l'enseignement :

- **Mission de soins** : Offrir une prise en charge de qualité aux patients atteints de cancer. Le CLB possède une organisation de soin pluridisciplinaire en collaboration avec les médecins, les soignants et permet l'implication des patients dans leurs soins.
- **Mission de recherche** : Garantir une recherche d'excellence tournée vers l'innovation, à l'aide d'échanges nationaux et internationaux. Le CLB réunit sur un seul et même site trois grandes composantes de recherche : la recherche fondamentale, translationnelle et clinique.
- **Mission d'enseignement** : Transmettre les connaissances de l'établissement par l'enseignement à trois publics : les futurs professionnels de santé et de recherche, les professionnels de santé en exercice, mais également leurs salariés afin de maintenir et de réactualiser les méthodes et compétences.

1.2 Département Prévention Cancer Environnement

Le département Prévention Cancer Environnement du CLB a pour mission d'étudier l'impact de l'environnement sur les risques de cancer (expositions environnementales, nutrition, activité physique, etc.) et se divise en trois axes. Premièrement la prévention et la prise en charge des facteurs de risque de cancer – avant, pendant et après la maladie, par la participation à la prise en charge globale et personnalisée des patients atteints de cancer. Le second axe concerne le développement et l'implémentation d'interventions permettant d'améliorer les comportements de prévention avant, pendant et après un cancer. Enfin, le troisième axe s'intéresse à l'étude des facteurs environnementaux et des risques de cancer [9].

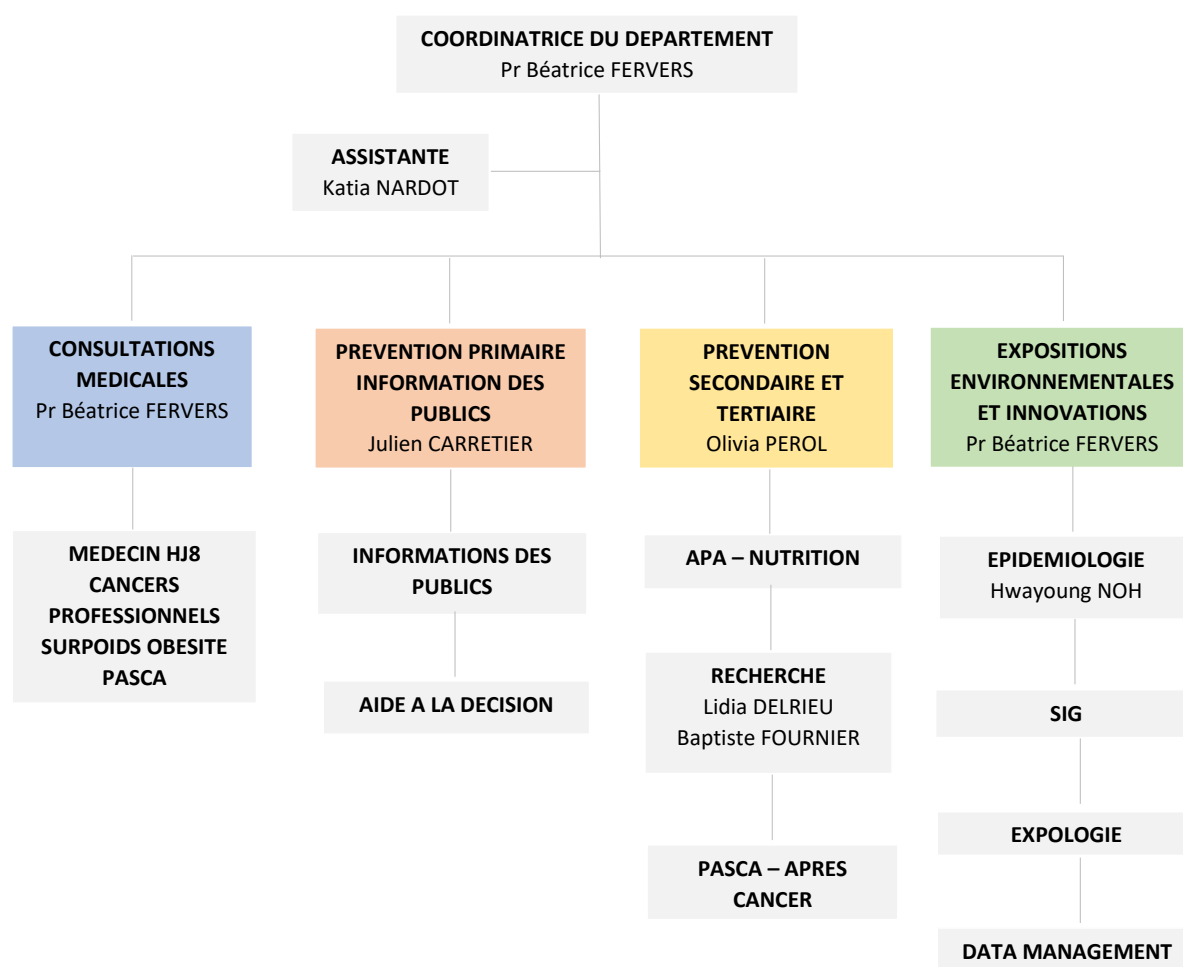


FIGURE 1.1 – Organigramme du Département Prévention Cancer Environnement [9].

2 Plateforme de recherche collaborative des Seintinelles sur le cancer

La plateforme des Seintinelles est la seule plateforme de recherche collaborative sur le cancer en France dont la mission est d'accélérer le processus de recherche [25]. Lancée en 2013 par Fabien Reyat (chirurgien de l'institut Curie) et Guillemette Jacob (ancienne patiente), la plateforme mobilise actuellement 35 139 volontaires, hommes et femmes, malades ou non. On comptabilise également 7199 patientes ayant des antécédents de cancer du sein [31]. Une enquête datant de 2018 a été proposée afin de mieux connaître le profil des participants à cette plateforme [31]. Au total, 10 054 participants ont répondu à l'enquête, 94% étaient des femmes d'un âge moyen de 46 ans, d'un niveau socio-économique élevé et dont 28% présentaient un historique de cancer. De plus, 96% d'entre elles étaient intéressées à répondre à des questionnaires. [25].

Grâce à cette plateforme, le recrutement des patients n'apparaît plus comme un obstacle à la recherche et permet de considérablement optimiser les délais de recrutement comme on peut le voir sur la Figure 1.2. Les Seintinelles ont participé à 31 études avec la collaboration de plus de 30 000 personnes [25].

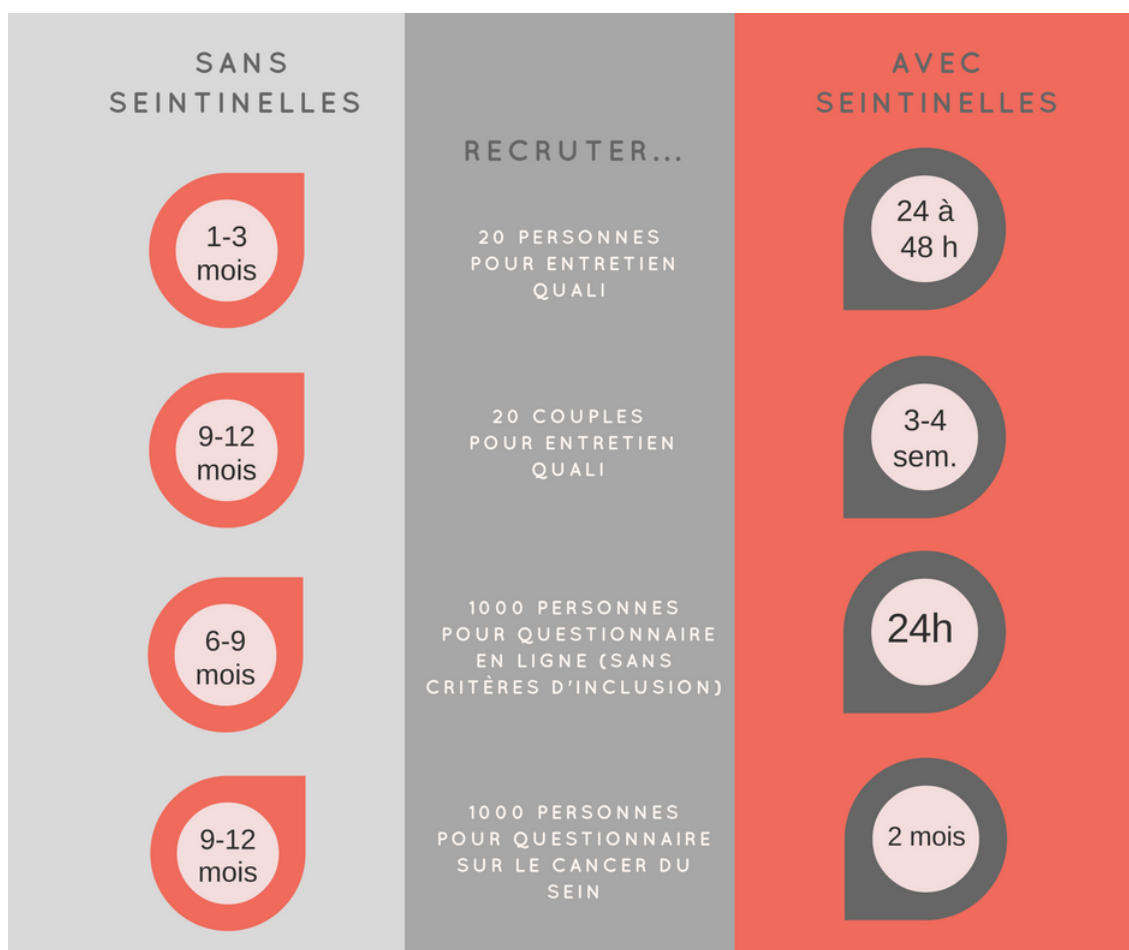


FIGURE 1.2 – Délais de recrutements avec/sans la plateforme Seintinelles [25].

3 Cancer du sein

3.1 Épidémiologie

Le cancer est le nom donné à tout type de tumeur maligne à évolution rapide qui se propage vers le reste de l'organisme. Ces tumeurs sont causées par la multiplication et la mutation non contrôlée de cellules du corps [8]. Le cancer du sein est une maladie hétérogène associée à des caractéristiques histologiques et biologiques. En France, environ 58 459 nouveaux cas de cancer du sein ont été diagnostiqués en 2018 et il reste le cancer le plus meurtrier chez les femmes avec 12 146 décès en 2018 [22]. L'âge médian des femmes au moment du constat du cancer du sein est de 63 ans tandis que celui au moment du décès est de 74 ans [15].

3.2 Traitements contre le cancer

Il existe différents types de traitement afin de soigner un cancer du sein [33] :

- Chirurgie : permet d'éliminer les tissus atteints par les cellules cancéreuses,
- Chimiothérapie : traitement médicamenteux agissant sur les mécanismes de la division cellulaire et permettant d'altérer les cellules cancéreuses quelles que soient leurs localisations,
- Radiothérapie : permet de détruire et d'empêcher le développement des cellules cancéreuses à l'aide de radiation à haute énergie,
- Immunothérapie : aide à renforcer le système immunitaire afin de combattre et de se protéger des cellules cancéreuses,
- Hormonothérapie : peut être utilisée lorsque la tumeur est « hormonodépendante », c'est-à-dire lorsque les hormones féminines stimulent la croissance de la tumeur. L'objectif est d'enrayer la stimulation des cellules cancéreuses par les hormones féminines,
- Thérapies ciblées : médicaments anticancéreux permettant d'agir sur les mécanismes des cellules cancéreuses en atténuant leurs croissances.

3.3 Facteurs de risques

Le cancer est une maladie multifactorielle puisque plusieurs facteurs de risque ont été identifiés comme augmentant le risque de développer un cancer, le plus souvent par des mécanismes indirects. Dans le cadre du cancer du sein, les principaux facteurs de risque connus sont essentiellement l'âge, les facteurs hormonaux et le mode de vie incluant l'activité physique [16]. Au niveau mondial, le cancer du sein apparaît comme la cinquième cause de mortalité avec 684 996 décès en 2020 [18]. De plus, on estime à 2 261 419 le nombre de femmes ayant été diagnostiquées d'un cancer du sein en 2020 tout âge confondu [18]. Parmi les principales causes de cancer, certaines sont inévitables telles que la génétique ou encore l'âge. Cependant, parmi celles évitables, plusieurs concernent des facteurs liés à la consommation d'alcool et au tabagisme, mais également au mode de vie, notamment le manque d'activité physique [16].

4 Activité physique

Selon l'Organisation Mondiale de la Santé (OMS), l'activité physique est définie comme tout mouvement corporel produit par les muscles squelettiques produisant une augmentation de la dépense énergétique [29]. Dans le cadre des loisirs, des tâches domestiques, sur le lieu de travail ou encore pour des déplacements, l'Activité Physique (AP) possède des bénéfices importants sur la santé notamment sur la réduction des risques de cancer [22]. En effet, les bénéfices ont largement été démontrés en prévention primaire puisqu'une pratique régulière permettrait de diminuer le risque de développer certains cancers tels que le cancer du sein, le cancer du colon, etc. comme on peut le voir dans le tableau 1.1.

Type de cancer	↘ (%)
Cancer du sein	10-27%
Cancer du côlon	25%
Cancer de l'endomètre	20%
Cancer du poumon	20%
Cancer de la vessie	15%
Cancer de l'œsophage	21%
Cancer du rein	12%

TABLE 1.1 – Diminution (en %) du risque de développer des cancers par le biais d'une activité physique régulière [10].

D'après l'étude de Aude-Marie Foucaut et son équipe en 2015, il a été constaté une dégradation significative de l'activité physique suite à la découverte d'un cancer du sein, ainsi que pendant et après le traitement de chimiothérapie [4]. Pourtant, la pratique physique améliore significativement la qualité musculaire [32], la composition corporelle [23, 13], la qualité de vie [24, 30, 27], la douleur et diminue la fatigue de 25 à 35% [20] (Figure 1.3). Il a également été montré dans des études observationnelles que la pratique d'une activité physique permettrait de diminuer le risque de décès du cancer du sein en comparaison avec des femmes plus sédentaires [26].

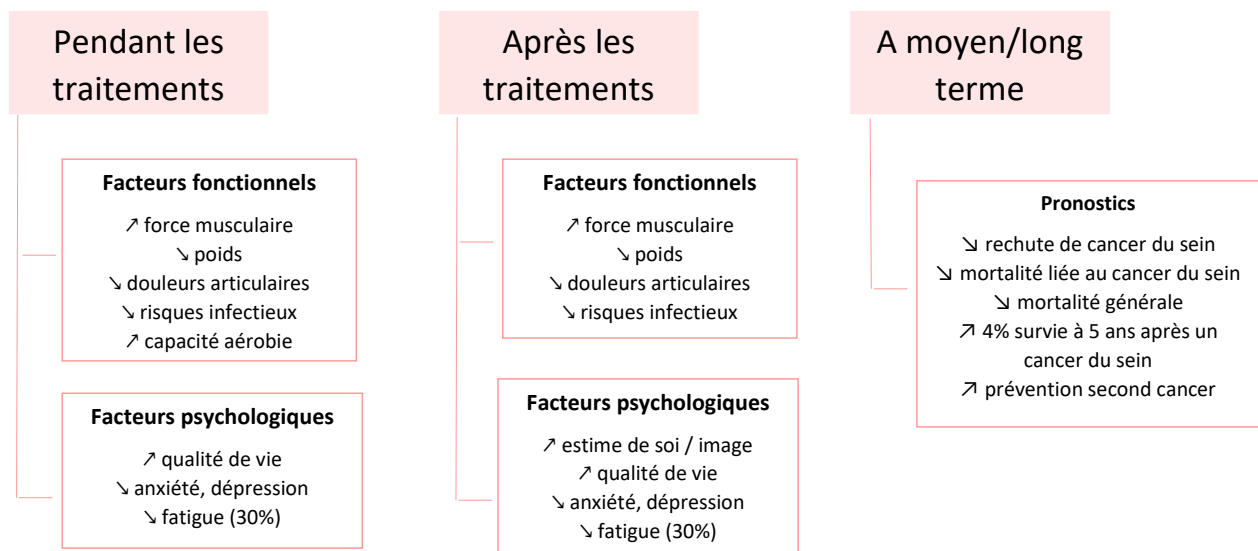


FIGURE 1.3 – Bénéfices de la pratique d’une activité physique pendant et après un cancer [10].

Les recommandations en matière d’AP pour les adultes de plus de 18 ans sont les suivantes [29] :

- pratiquer au moins 150 minutes par semaine d’activité physique d’intensité modérée,
- ou pratiquer au moins 75 minutes par semaine d’activité physique d’intensité soutenue,
- ou pratiquer une combinaison équivalente d’activité d’intensité modérée et soutenue la semaine.

En 2017, seulement 54% des femmes majeures atteignaient les recommandations de l’OMS en matière d’activité physique [2]. Ceci peut s’expliquer par le mode de vie, par une augmentation du temps passé assis aussi bien dans les activités professionnelles que les loisirs. Évoquons la notion de sédentarité : elle correspond au temps passé assis ou allongé durant la phase d’éveil entre le lever et le coucher. Elle est caractérisée par une dépense énergétique inférieure ou égale à la dépense de repos en position assise ou allongée [1]. Étant considérée comme la 4^{ème} cause de mortalité au niveau mondial [29], la sédentarité augmente significativement le risque de nombreuses maladies chroniques telles que le diabète, l’hypertension et plusieurs formes de cancer.

Il apparaît donc important d’augmenter le niveau d’activité physique des femmes avant l’apparition d’un cancer (prévention primaire) afin de réduire les risques de survenue de la maladie, mais également de limiter les effets secondaires des traitements et limiter le risque de récurrence (prévention tertiaire).

5 Problématique

Afin de promouvoir plus de femmes à la pratique d'une activité physique, il apparaît judicieux de connaître leurs préférences, leurs freins et les facteurs permettant une meilleure adhésion. Chez les femmes atteintes de cancer du sein, plusieurs barrières ont été identifiées, principalement liées aux effets secondaires des traitements (fatigue, douleur), aux facteurs environnementaux et temporels (distance, météo, installations, manque de temps), aux facteurs économiques (coût), psychologiques (manque de confiance en soi, intérêt pour l'activité physique, manque de motivation) et sociaux (manque d'intérêt pour l'activité physique, de motivation, de soutien ou encouragement des autres) [28, 14]. Ces barrières varient en plus en fonction du stade du cancer. Bien que moins souvent décrits, des facilitateurs à la pratique d'une activité physique ont également été identifiés chez les femmes atteintes ou non d'un cancer du sein, comme avoir un meilleur contrôle sur sa santé ou les contacts sociaux durant la pratique [28]. Les études disponibles ont essentiellement été réalisées en prévention tertiaire, auprès d'une population spécifique de femmes atteintes de cancer du sein et avec un effectif limité. Aucune étude n'a proposé à des femmes, à la fois en prévention primaire et en prévention tertiaire d'un cancer du sein, de remplir au même moment des questionnaires pour connaître leurs préférences, barrières et facilitateurs à l'activité physique.

L'objectif de cette étude est de décrire les préférences, les barrières et les facteurs influençant la pratique d'une activité physique des femmes en prévention primaire et tertiaire du cancer du sein.

Après avoir renseigné le contexte du projet et l'environnement professionnel dans lequel il a été mis en œuvre, nous poursuivrons dans la seconde partie par expliquer la méthode de l'étude afin de répondre à la problématique. Nous déterminerons les facteurs influant le plus la pratique d'activité physique et évaluerons leur impact. Finalement, nous terminerons ce projet par la présentation des résultats obtenus et des perspectives.

Chapitre 2

Données et méthodes

Dans cette partie, nous commencerons par définir le contexte de l'étude en abordant la population, le questionnaire mis en place et la gestion des données recueillies. Ensuite, nous exposons la méthodologie utilisée afin de répondre à la problématique.

1 Contexte de l'étude

1.1 Population

Afin d'étudier les préférences, les barrières et les facilitateurs de notre population cible, nous avons mis en place une étude transversale réalisée à l'aide d'un questionnaire en ligne codé sur LimeSurvey. Ce dernier est un logiciel d'enquêtes statistiques permettant aux usagers, aux enquêteurs et aux statisticiens de mettre en place et de publier des questionnaires. Le lien du questionnaire a été distribué en ligne via la plateforme collaborative des Seintinelles. Les données ont été recueillies du 4 février au 31 mars 2021. De plus, elles ont été stockées sur les serveurs sécurisés du Centre Léon Bérard.

Les personnes souhaitant participer à l'étude devaient respecter les critères suivants :

- être une femme,
- avoir au moins 18 ans.

Il pouvait s'agir de femmes n'ayant jamais eu de cancer du sein ou ayant des antécédents de cancer du sein qu'il soit en cours de traitement ou passé.

1.2 Questionnaire

Le questionnaire de l'étude est présenté en Annexe A. Une collaboration avec deux chercheurs internationaux spécialistes des questions liées aux préférences, barrières et facteurs favorisant la pratique de l'activité physique auprès des patients atteints de cancer a été mise en place afin de structurer le questionnaire. Ces derniers sont Jeff Vallance et Siobhan Phillips. Le questionnaire comporte 59 questions réparties en 6 domaines :

- les caractéristiques socio-démographiques, de lieu de vie et de santé (17 questions),
- le niveau d'activité physique (4 questions),
- les préférences en matière d'activité physique (19 questions),

- l'utilisation des technologies liées à l'activité physique (13 questions),
- les barrières à la pratique d'une activité physique (3 questions),
- les facilitateurs à la pratique d'une activité physique (3 questions).

Domaine : identification

Le premier domaine étudié est l'identification des participantes. Ce dernier comporte des informations socio-démographiques telles que l'âge, la situation familiale, le niveau d'éducation, la situation professionnelle, la catégorie socio-professionnelle, le département de résidence, la situation du logement, la proximité d'espaces verts mais encore le poids, la taille, le nombre d'enfants présents au domicile, les traitements reçus pour le cancer du sein et les problèmes de santé (par exemple maladies respiratoires, inflammatoires, etc.).

Nous avons également analysé l'Indice de Masse Corporelle (IMC) à l'aide de la formule suivante :

$$\text{IMC} = \frac{\text{poids (kg)}}{\text{taille}^2(\text{cm})}.$$

Puis, nous l'avons réparti en 5 catégories selon sa valeur : maigreur ($\text{IMC} < 18.5$), corpulence normale ($18.5 \leq \text{IMC} < 25$), surpoids ($25 \leq \text{IMC} < 30$) et obésité ($\text{IMC} \geq 30$) [12].

Domaine : niveau d'activité physique

Concernant les informations sur le niveau d'activité physique, nous avons utilisé le score du questionnaire Godin-Shephard Leisure-Time Physical Activity Questionnaire (GSLTPAQ) [19]. Ce dernier est utilisé dans la recherche en oncologie afin d'étudier le niveau d'activité physique relatif aux activités de loisirs. Les questions visent à obtenir des informations sur le nombre de fois par semaine qu'une personne effectue une activité physique de loisirs d'intensité légère, modérée et intense d'une durée d'au moins 15 minutes [19]. A l'aide du GSLTPAQ, on peut en déduire un score appelé Leisure Score Index (LSI) afin de classer les participantes dans les catégories active / inactive concernant leur activité physique. La formule du LSI est la suivante :

$$\text{LSI} = (\text{fréquence AP léger} \times 3) + (\text{fréquence AP modéré} \times 5) + (\text{fréquence AP intense} \times 9).$$

Les valeurs des seuils sont définies comme suit : une personne ayant un LSI supérieur ou égal à 24 est classée comme active tandis qu'une personne ayant un LSI inférieur ou égal à 23 est classée comme inactive [17].

Domaine : préférences

Les préférences des participantes en matière d'activité physique ont été mesurées à l'aide de diverses questions tirées d'études similaires. Il leur a été demandé les **moyens** (téléphone, mail, etc.) et **par quelle personne** (médecin de famille, infirmière, etc.) elles aimeraient recevoir des informations/conseils sur la pratique d'une AP. Mais également le **type d'AP** que les participantes souhaitent pratiquer, le **lieu** de pratique et **comment** pratiquer (seul, en groupe, etc.). De plus, des questions concernaient également la

manière d’être **supervisé** lors des séances, l’**intensité** de pratique, l’**organisation**, le **nombre de jours** par semaine, le **moment** de la journée et la **durée**. Finalement, des préférences concernant des stratégies en distanciel (SMS d’encouragement, utilisation de montre connectée, compétition de type gaming, etc.) ont été évaluées en utilisant une échelle de Likert (outil permettant de mesurer une attitude chez des individus) à 5 items.

Domaine : technologies

À propos du domaine des technologies, les participantes ont dû répondre à des questions concernant l’utilisation des réseaux sociaux, des applications liées à l’AP et des traqueurs d’AP. De plus, des interrogations sur le souhait de participer à un programme à distance via un traqueur d’AP, via une application téléphonique et via un site web ont également été posées. Des items de non-intérêt à cette adhésion tels que la peur de ne pas savoir s’en servir, la dépendance ou le manque de temps ont également été évalués.

Domaine : barrières

Les barrières des participantes sur la pratique d’activité physique ont été recensées à l’aide de questions sur une échelle de Likert. Les barrières étudiées et tirées d’études similaires sont la fatigue, la douleur, le regard des autres, les obligations familiales, les effets secondaires du traitement, etc.

Domaine : facilitateurs

Les facilitateurs à la pratique d’une AP concernent l’amélioration des capacités physiques, des symptômes de la maladie, de la qualité de vie, la diminution de la fatigue, le maintien/perte de poids, etc. Ces dernières également tirées d’études similaires sont mesurées à l’aide d’une échelle de Likert.

1.3 Gestion des données

Le jeu de données disponible comprend $n = 2766$ participantes ainsi que 263 variables codées dans un format international.

Avant toute analyse de la base de données, un travail de data management a été réalisé afin d’en vérifier la fiabilité. Toutes les analyses ont été effectuées à l’aide du logiciel R (version 4.0.4) et les codes sont disponibles dans **GitHub** dont l’adresse est précisé au début du document. Nous avons vérifié le type des variables et la cohérence des données. De plus, des variables ont été créées à partir de données disponibles telles que l’âge des participantes, leurs IMC, le nombre de maladies et le statut d’activité physique. Finalement, nous avons ajouté une variable permettant de faire des sous-groupes de participantes en fonction du stade du cancer du sein (localisé, métastatique) et du stade des traitements (en cours, fini). Nous avons ainsi mis en place les sous-groupes suivants :

- pas de cancer du sein,
- cancer du sein localisé en cours de traitement,
- cancer du sein localisé ancien,
- cancer du sein métastatique.

Sur les 2766 participations, certaines femmes n'ont pas poursuivi le questionnaire jusqu'à la fin (Figure 2.1), nous avons recensé 458 (16.6%) réponses incomplètes. De plus, parmi ces dernières, 53 (11.6%) femmes ont seulement ouvert le lien du questionnaire sans répondre à aucune question. Après suppression de ces 53 participantes, nous avons 10,3% de données manquantes. Étant donné que cela représente un pourcentage assez faible de la population d'étude, nous avons pris la décision de les exclure. Ainsi, nous avons $n = 2308$ participantes. Cependant, parmi celles-ci, nous avons exclu 2 femmes de l'étude : une diagnostiquée d'un cancer du sein et n'ayant pas renseigné le statut du cancer (métastatique ou localisé) et une ayant un IMC aberrant environ égal à 87. En effet, une personne ayant un IMC supérieur ou égal à 40 est considérée comme obèse de grade III. Finalement, nous avons retenu 2306 participantes.

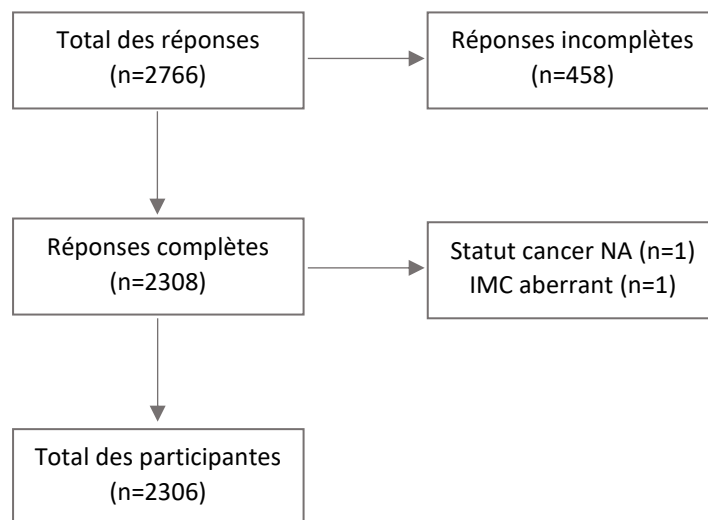


FIGURE 2.1 – Flow chart de la population d'étude.

Dans le jeu de données final (2306 participantes), nous constatons des données manquantes pour certaines variables tels que l'âge des enfants des participantes, le statut du cancer, le type de traitements ou encore les raisons pour lesquelles les participantes ne sont pas intéressées pour réaliser un programme d'activité physique à distance (via un traqueur d'AP, un site web et une application mobile). Cependant, ces données manquantes correspondent à des items pour lesquels les femmes n'étaient pas concernées. En effet, nous nous trouvons dans cette situation par exemple lorsqu'une participante indique qu'elle n'a pas d'enfants et qu'elle ne répond pas à la question concernant l'âge de ces enfants ou encore lorsqu'une répondante n'a pas de cancer du sein et qu'on lui demande quel est le stade de sa maladie. Les pourcentages de valeurs manquantes pour les variables concernées sont exposés dans le tableau 2.1 ci-dessous.

Valeurs manquantes	n (%)
Âge des enfants	1291 (55.98)
Statut cancer	1245 (53.99)
Traitements cancer	1245 (53.99)
Traqueur d'AP	1045 (45.32)
Programme AP via traqueur	761 (33.00)
Programme AP via app	872 (37.81)
Programme AP via web	785 (34.04)

TABLE 2.1 – Pourcentage de valeurs manquantes.

Concernant les valeurs manquantes pour l'âge des enfant, étant donné que ces dernières correspondent à des participantes n'ayant pas d'enfants, nous avons fixé la valeur à 0 dans chacune des variables correspondantes. Pour le statut du cancer du sein et les traitements anti-cancéreux, les valeurs manquantes correspondent à des participantes n'ayant pas de cancer du sein, ainsi, nous avons ajouté une catégorie "Pas de cancer".

2 Méthode de description de la population

Les variables du jeu de données se répartissent en deux types : variables qualitatives et quantitatives. Avant toute analyse statistique, nous avons réalisé une analyse descriptive des 4 sous-groupes précédemment énoncés. Les variables qualitatives ont été décrites au moyen de distributions, de fréquences et de pourcentages. Concernant les variables quantitatives, nous avons étudié les tendances centrales ainsi que leurs dispersions à l'aide d'indicateurs et de représentations graphiques. De plus, afin d'augmenter la lisibilité, certaines variables sont présentées également en classes telles que l'âge ou l'IMC.

3 Méthodes de partitionnement

Les méthodes de partitionnement permettent à partir d'un ensemble de données, de créer des groupes d'individus homogènes i.e. présentant des caractéristiques communes. Il existe actuellement diverses méthodes dont certaines sont plus ou moins adéquates selon le type de variables. Afin de décrire les préférences, les barrières et les facteurs facilitant la pratique d'activité physique des femmes en prévention primaire et tertiaire d'un cancer du sein, nous avons utilisé **la classification hiérarchique sur composantes principales**. L'objectif est d'extraire à partir du jeu de données des sous-ensembles de participantes non identifiables au premier abord et présentant des caractéristiques similaires afin de les décrire.

L'approche Hierarchical Clustering on Principal Components (HCPC) est une méthode complète permettant de combiner trois méthodes d'analyses de données multivariées. L'algorithme de la classification hiérarchique sur composantes principales repose sur le principe suivant :

1. Effectuer une analyse des correspondances multiples,
2. Effectuer une Classification Ascendante Hiérarchique (CAH) sur les résultats de l'étape précédente,

3. Effectuer une optimisation de la partition initiale à l'aide des k -means.

Les méthodes d'analyse factorielle et de classification amènent une complémentarité concernant l'analyse d'un jeu de données : l'analyse factorielle permet de transformer les variables qualitatives en quantitatives via les axes principaux et de supprimer le bruit du jeu de données afin de rendre la classification plus robuste.

3.1 Analyse des Correspondances Multiples

Choix de la méthode. Une première phase de discussion a porté sur le choix de la méthode de composantes principales. Étant donné que nous sommes en présence de données mixtes (quantitatives et qualitatives), de nombreuses méthodes ne sont pas applicables. Une pratique couramment utilisée consiste à transformer les variables quantitatives en qualitatives ordinales par découpage en tranches afin d'obtenir des données de même nature [3]. Néanmoins, cette démarche possède des limites :

- lorsque que le nombre d'individus est faible ("moins d'une centaine pour fixer les idées" [3]),
- lorsqu'il y a peu de variables qualitatives en comparaison aux variables quantitatives.

Au regard du jeu de données disponible, ces barrières ne sont pas un frein à notre analyse. Nous faisons donc le choix d'utiliser l'Analyse des Correspondances Multiples (ACM). Cette dernière est une méthode factorielle adaptée à l'exploration de données de type enquête décrivant des individus à l'aide de variables.

Tableau de données et notations. Considérons un tableau de données où n individus sont décrits à l'aide de p variables. L'élément h_{ik} de ce tableau représente la modalité de la variable k , $k \in \{1, \dots, p\}$ possédée par l'individu i , $i \in \{1, \dots, n\}$.

Une étape essentielle avant de mettre en place l'ACM est le recodage des variables quantitatives en variables qualitatives ordinales par découpage en tranches. Néanmoins, afin de prendre en compte l'interprétabilité des variables, il est important d'en respecter le sens.

Nous pouvons ainsi construire le tableau logique disjonctif complet où chaque individu possède une seule modalité pour une variable qualitative. Pour ce faire, les p variables qualitatives sont représentées par p tableaux logiques disjonctifs complets nommés H_1, \dots, H_p . Ces derniers possèdent autant de colonnes que de modalités possédées par chaque variable. Finalement les p tableaux sont concaténés en un seul tableau disjonctif complet H (Figure 2.2).

Le nombre de modalités d'une variable $k \in \{1, \dots, p\}$ est noté m_k et le nombre total des modalités est noté $m = \sum_{k=1}^p m_k$. Le tableau H contient ainsi n lignes et m colonnes. L'élément h_{is} vaut 1 si l'individu $i \in \{1, \dots, n\}$ possède la modalité s , de la variable k et 0 sinon. La somme des éléments de chaque ligne vaut p et la somme du tableau vaut np . De plus, notons N_s la somme de la colonne d'une modalité s , $N_s = \sum_{i=1}^n h_{is}$. Finalement, étant donné qu'un individu possède une et une seule modalité d'une variable qualitative quelconque, les vecteurs indicatrices d'une même variable sont orthogonaux deux à deux.

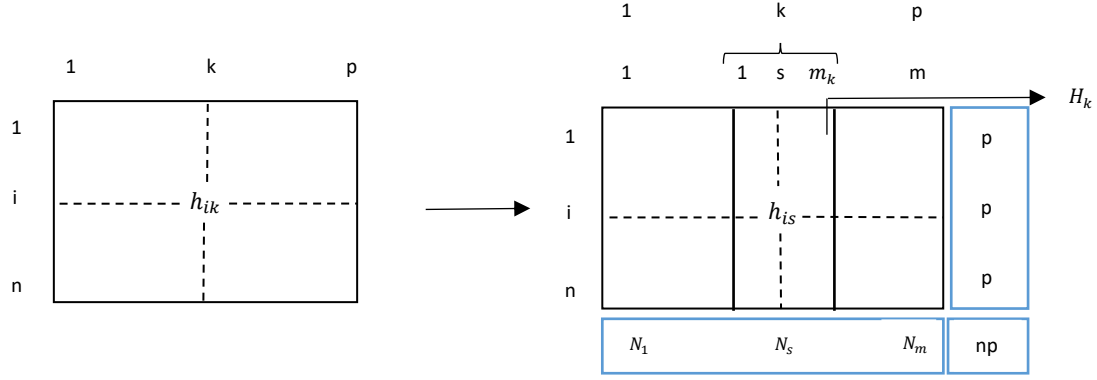


FIGURE 2.2 – Tableau de données et notations.

Ainsi, concernant le tableau logique disjonctif complet H , nous avons :

- $\sum_{s=1}^{m_k} h_{is} = 1 \quad \forall (i, k),$
- $\sum_{s=1}^m h_{is} = p \quad \forall i,$
- $\sum_{i=1}^n h_{is} = N_s \quad \forall s,$
- $\sum_{s=1}^{m_k} N_s = n \quad \forall k.$

Notons $F \in \mathbb{R}^{n \times m}$ la matrice des fréquences obtenue en divisant le tableau disjonctif complet par sa somme np . Cette dernière est représentée matriciellement dans la Figure 2.3. L'élément f_{is} de la matrice vaut $\frac{h_{is}}{np}$ où $i \in \{1, \dots, n\}$ et $s \in \{1, \dots, m\}$. Ainsi, la somme de chaque ligne de la matrice est égale à $\frac{1}{n}$ et le poids de chaque colonne vaut $\frac{N_s}{np}$ étant donné que la somme d'une colonne s du tableau disjonctif complet correspond au nombre d'individu possédant la modalité (i.e. N_s). De plus, on a $\sum_{s=1}^m \frac{h_{is}}{np} = \frac{1}{n}$ étant donné que $\sum_{s=1}^m h_{is} = p$.

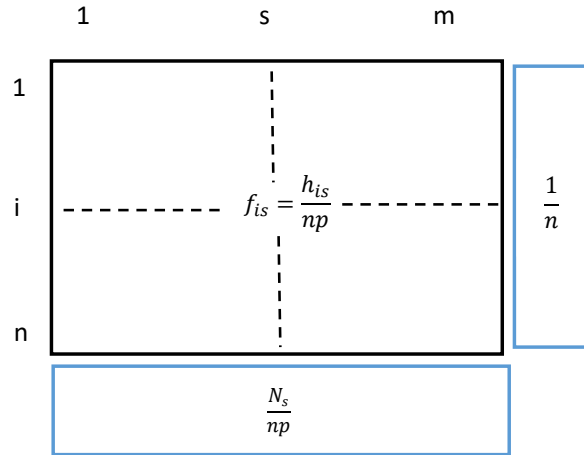


FIGURE 2.3 – Tableau de fréquences et notations.

Nuages direct et dual. Définissons plusieurs notations qui nous seront utiles par la suite :

- $w = \left(\dots, \frac{1}{n}, \dots \right)' \in \mathbb{R}^n$ le vecteur des poids des individus,
- $z = \left(\dots, \frac{N_s}{np}, \dots \right)' \in \mathbb{R}^m$ le vecteur du poids des modalités,
- $W = \text{diag}(w) \in \mathbb{R}^{n \times n}$ la matrice des poids des lignes,
- $Z = \text{diag}(z) \in \mathbb{R}^{m \times m}$ la matrice des poids des colonnes.

On définit la matrice des profils-lignes L où l'élément $l_{is} = \frac{f_{is}}{\sum_{s=1}^m f_{is}} = n f_{is} = \frac{h_{is}}{p}$. Les profils lignes sont donc obtenus en divisant les lignes du tableau disjonctif complet par le nombre de variables qualitatives. De plus, la matrice des profils-colonnes est obtenue en divisant les éléments de la matrice de fréquences par sa marge colonne. Nous avons les notations suivantes :

$$L = W^{-1}F = \frac{1}{p}H \in \mathbb{R}^{n \times m},$$

$$C = FZ^{-1} = \left(\frac{H}{N_s} \right)_{s=1, \dots, m} \in \mathbb{R}^{n \times m}.$$

De plus, le tableau des profils-lignes L est analysé à l'aide de la métrique $M_L = Z^{-1}$ correspondant à la diagonale des inverses des fréquences des colonnes et le tableau des profils-colonnes C avec la métrique $M_C = W^{-1}$.

Géométriquement, le tableau H se traduit en deux nuages de points pondérés : le nuage direct des individus de \mathbb{R}^m avec un poids de $1/n$ et la métrique M_L et le nuage dual de \mathbb{R}^n avec un poids de N_s/np et la métrique M_C (Figure 2.4).

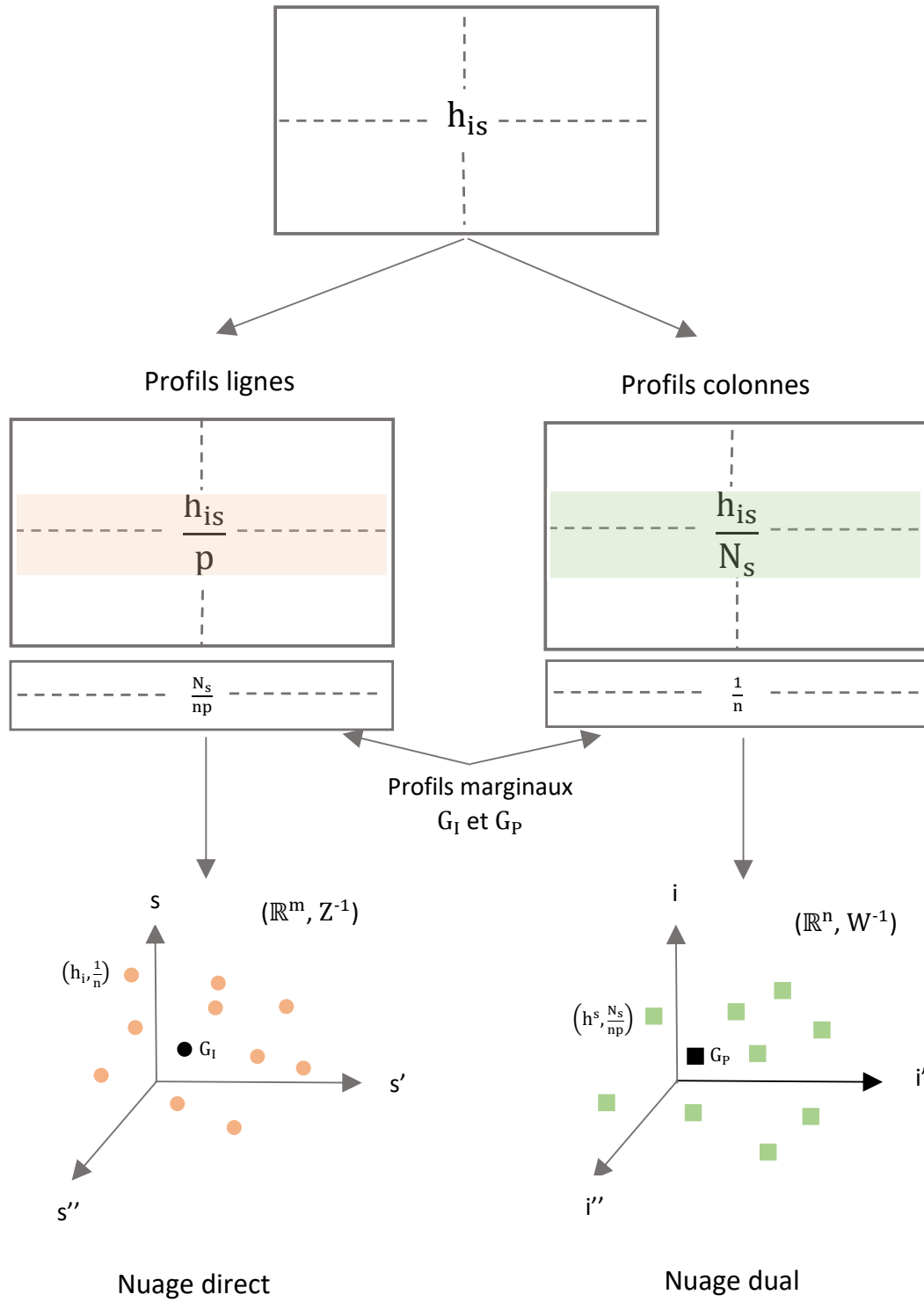


FIGURE 2.4 – Traduction du tableau logique disjointif complet en espaces métriques.

- Comparaison de deux individus i et l , $i, l \in \{1, \dots, n\}$ à l'aide de la distance du χ^2 :

$$d_{\chi^2}^2(i, l) = \sum_{s=1}^m \frac{\left(\frac{h_{is}}{p} - \frac{h_{ls}}{p}\right)^2}{f_{.s}} = \sum_{s=1}^m \frac{np}{N_s} \left(\frac{h_{is}}{p} - \frac{h_{ls}}{p}\right)^2 = \frac{1}{p} \sum_{s=1}^m \frac{n}{N_s} (h_{is} - h_{ls})^2,$$

où : N_s est la somme de la colonne-modalité s du tableau H ,

$f_{.s} = \frac{N_s}{np}$ est la marge colonne de F .

Ainsi, plus les individus i et l auront des modalités différentes, plus leur distance sera grande. De plus, lorsqu'un individu possède une modalité rare, celui-ci sera éloigné des autres individus étant donné que dans cette distance le poids d'une modalité s est l'inverse de sa fréquence i.e. $\frac{np}{N_s}$.

- Distance entre deux modalités s et d , $s, d \in \{1, \dots, M\}$:

$$d_{\chi^2}^2(m, d) = \sum_{i=1}^n \frac{\left(\frac{h_{is}}{N_s} - \frac{h_{id}}{N_d}\right)^2}{f_{i.}} = n \sum_{i=1}^n \left(\frac{h_{is}}{N_s} - \frac{h_{id}}{N_d}\right)^2,$$

où : N_s est la somme de la colonne-modalité s du tableau H ,

$f_{i.} = \frac{1}{n}$ est la marge ligne de F .

- Le centre de gravité du nuage des individus $G_I = z$ correspond au profil ligne uniforme. De façon similaire, le centre de gravité du nuage dual $G_P = w$ est un vecteur contenant le poids des individus. On a donc les notations suivantes :

$$G_I = \left(\frac{N_1}{np}, \dots, \frac{N_m}{np}\right)' \in \mathbb{R}^m$$

$$G_P = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)' \in \mathbb{R}^m.$$

- Distance entre une modalité s , $s \in \{1, \dots, m\}$ et G_P : (démonstration en Annexe B) :

$$d^2(s, G_P) = \frac{n}{N_s} - 1.$$

Inertie des nuages de points pondérés. Définissons l'inertie d'un nuage de points pondéré dans un espace muni de la métrique M . L'inertie est une mesure de la dispersion du nuage. Considérons n points de \mathbb{R}^m muni d'un poids w_i , $i \in \{1, \dots, n\}$: $\{h_i, w_i\}_{i=1, \dots, n}$ et $y \in \mathbb{R}^m$ un point de référence. L'inertie du nuage par rapport au point de référence y est défini par [7] :

$$\text{In}_y\{h_i, w_i\}_i = \sum_{i=1}^n w_i \|h_i - y\|_M^2.$$

A présent, nous allons évoquer la notion d'inertie le long d'un sous espace noté F .

$$\text{In}^F(\{h_i, w_i\}) = \text{In}(\{x_i, w_i\}) = \sum_{i=1}^n w_i \|\pi_F h_i\|_M^2,$$

où x_i est la projection orthogonale du point h_i sur le sous espace F . Finalement, nous pouvons exposer la notion d'inertie le long d'un axe $\langle v \rangle$ où v est un vecteur unitaire. En projetant orthogonalement les points $(h_i)_i$ sur l'axe $\langle v \rangle$, on obtient n coordonnées sur cet axe. Ces dernières forment un vecteur de \mathbb{R}^n appelé composante. La coordonnée du point h_i sur l'axe $\langle v \rangle$ est déterminée par $f_i = \langle h_i, v \rangle_M = h_i' M u$. Par conséquent, le vecteur de \mathbb{R}^n des composantes est $f = H M u$. Ainsi, l'inertie du nuage le long d'un axe $\langle v \rangle$ avec v un vecteur unitaire est :

$$\begin{aligned} \text{In}_{\langle v \rangle}(\{h_i, w_i\}_i) &= \sum_{i=1}^n \|\pi_{\langle v \rangle} h_i\|_M^2 \\ &= \sum_{i=1}^n w_i f_i^2 \\ &= f' W f \\ &= u' M H' W H M u \\ &= \langle u, \underbrace{H' W H M}_{\text{matrice d'inertie}} u \rangle_M. \end{aligned}$$

L'inertie d'une modalité s , $s \in \{1, \dots, m\}$ est définie par $I(s) = \frac{1}{p} \left(1 - \frac{N_s}{n}\right)$ et l'inertie d'une variable k , $k \in \{1, \dots, p\}$ par : $I(k) = \frac{1}{p} (m_k - 1)$, où m_p est le nombre de modalité de la variable k . L'inertie totale est définie comme la somme pondérée des distances au carré entre les individus et le centre de gravité. Ainsi, on peut définir l'inertie totale comme étant la somme des inerties des variables :

$$I_T = \sum_{k=1}^p \frac{1}{p} (m_k - 1) = \frac{1}{p} \left(\sum_{k=1}^p m_k - p \right) = \frac{m - p}{p} = \frac{m}{p} - 1.$$

Analyses directe et duale. Les vecteurs principaux directs peuvent être caractérisés comme vecteurs propres de la matrice d'inertie via l'équation suivante [5] :

$$\begin{aligned} L' W L M_L u &= \lambda u \\ \Leftrightarrow \frac{1}{p^2} H' W H Z^{-1} u &= \lambda u \\ \Leftrightarrow \frac{1}{p} H' W H D u &= \lambda u \text{ en posant } D = \frac{1}{p} Z^{-1} \\ \Leftrightarrow H' W H D u &= \lambda p u \\ \Leftrightarrow H' W H D u &= \gamma u \text{ en posant } \gamma = \lambda p. \end{aligned}$$

On retrouve ainsi l'équation caractéristique de l'ACP du tableau H muni des poids W et de la métrique D correspondant à l'inverse des fréquences des modalités car $D = \frac{1}{p} Z^{-1} =$

$$\frac{1}{p} \left(\text{diag} \left(\frac{N_s}{np} \right)_s \right)^{-1} = \text{diag} \left(\frac{n}{N_s} \right)_s.$$

Relations barycentriques Lors de la pratique de l'analyse des correspondances multiples, il est possible d'effectuer une représentation simultanée des individus et des modalités sur un même graphique. Pour cela, nous avons besoin des relations barycentriques suivantes [5] :

$$f_i^q = \frac{1}{\sqrt{\lambda_q}} \sum_{s=1}^m \frac{1}{p} h_{is} \phi_q^s = \frac{1}{\sqrt{\lambda_q}} \sum_{s/i \text{ possède } s} \frac{1}{p} \phi_q^s, \quad (2.1)$$

$$\phi_q^s = \frac{1}{\sqrt{\lambda_q}} \sum_{i=1}^n \frac{1}{N_s} h_{is} f_i^q = \frac{1}{\sqrt{\lambda_q}} \sum_{i/i \text{ possède } s} \frac{1}{N_s} f_i^q, \quad (2.2)$$

où $i \in \{1, \dots, n\}$, $s \in \{1, \dots, m\}$, q représente un axe, f_i^q est la projection de la ligne-individu i sur l'axe q , ϕ_q^s est la projection de la colonne-modalité s sur l'axe q et λ_q la $q^{\text{ème}}$ valeur propre.

Ainsi, par (2.1), les individus se positionnent de manière claire par rapport aux modalités. En effet, un individu se positionne sur l'axe q , au coefficient de dilatation près, à l'isobarycentre de ces modalités. De manière symétrique, par (2.2), une modalité se trouve sur l'axe q , au coefficient de dilatation près, au barycentre équilibré des individus possédant cette modalité.

3.2 Classification Ascendante Hiérarchique

Son objectif est le partitionnement d'une population en groupes homogènes. L'idée principale est la maximisation de l'homogénéité intra-classe et de l'hétérogénéité inter-classe [6].

Une méthode ascendante repose sur le principe suivant : partir d'une partition où les classes sont des singletons i.e. chaque individu représente une unique classe. Puis, ces individus sont agrégés en classe de plus en plus grande jusqu'à obtenir une unique classe contenant tous les individus.

Afin d'agréger deux classes, nous utilisons un indice d'inertie appelé l'indice de Ward. L'indice de Ward consiste à regrouper les classes de façon à ce que l'augmentation de l'inertie intra-classe soit minimale. Pour deux classes A et B , il est défini par :

$$\mu(A, B) = \text{In}(\{(\bar{x}_A, w_A), (\bar{x}_B, w_B)\}) = \frac{w_A w_B}{w_A + w_B} \|\bar{x}_A - \bar{x}_B\|^2,$$

où \bar{x}_A et \bar{x}_B sont respectivement les centres de gravité des classes A et B et w_A , w_B les poids des classes A et B .

Algorithme 1 : Algorithme CAH

Résultat : Partition des individus

Entrées : H (n individus, p variables)

1 Partir de n classes initiales qui sont les singletons;

2 **répéter**

3 | agrégations des deux classes les plus proches selon le critère de Ward;

4 **jusqu'à** les n individus forment un unique groupe;

Les résultats de la CAH sont représentés via un arbre appelé dendrogramme. Dans la partie la plus haute du dendrogramme tous les individus sont agrégés dans une même classe tandis que dans la partie la plus basse, les individus forment des singletons dans leurs groupes comme on peut le constater sur la Figure 2.5. Les individus se présentent le long de l'axe horizontal tandis que le long de l'axe vertical nous avons la valeur de l'indice d'agrégation des individus. Le dendrogramme nous permet de déterminer des partitions prometteuses en coupant l'arbre au plus grand décrochement de niveau.

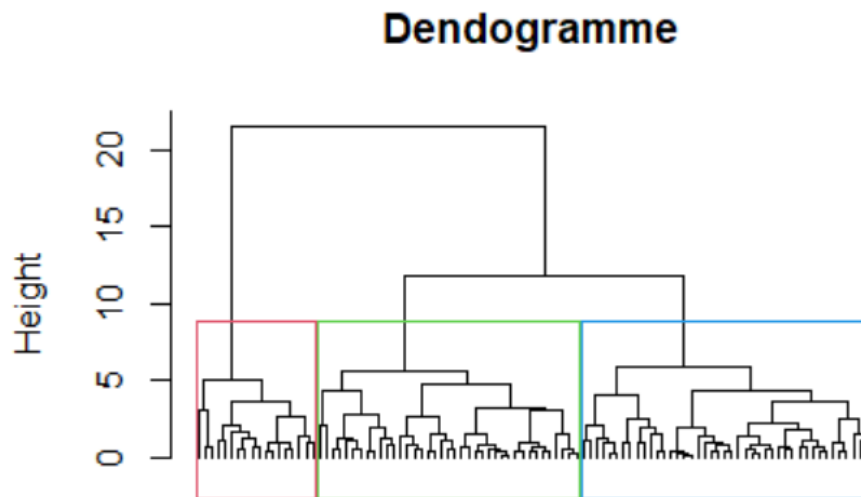


FIGURE 2.5 – Dendrogramme issu d'une CAH.

3.3 K-means

Enfin, la dernière étape de la classification hiérarchique sur composantes principales est la méthode des k -means. Nous utilisons les barycentres des classes issues de la CAH comme initialisation des centres des classes afin de stabiliser la partition. L'algorithme de la méthode des k -means est décrit ci-dessous :

Algorithme 2 : Algorithme des k -means

Résultat : Partition des individus

Entrées : Nombre de classes k , matrice de données

- 1 Initialisation des centres des k classes;
 - 2 **répéter**
 - 3 | Affecter chaque point individu à la classe du centre le plus proche ;
 - 4 | Recalculer le centre de chaque classe ;
 - 5 **jusqu'à** *Stabilisation des classes*;
-

Chapitre 3

Résultats

Dans ce chapitre, nous allons tout d’abord énoncer les résultats obtenus lors de l’étude descriptive des participantes, premièrement en mentionnant leurs caractéristiques puis leurs préférences, barrières et facilitateurs à l’activité physique. Dans une seconde partie, nous verrons les résultats obtenus en utilisant la méthode de partitionnement énoncé dans le chapitre précédent.

1 Étude descriptive des participantes

1.1 Caractéristiques de la population

Une première approche s’est portée sur l’étude des participantes du point de vue de la maladie. La population étudiée ($n = 2306$) est constituée de femmes dont l’âge moyen est de 48.7 ± 13.2 ans. Nous avons 54.0% des femmes qui ont déclarées ne pas avoir d’historique de cancer du sein tandis que 46.0% en possède un, actuel ou passé. Parmi ces dernières, 87.3% ont un cancer localisé et 12.7% un cancer métastatique. De plus, celles ayant été diagnostiquées d’un cancer du sein localisé sont soit en cours de traitements (57.2%), soit en rémission (42.8%). Rappelons que nous avons mis en place les 4 sous-groupes suivants afin d’en étudier les différences :

1. pas de cancer du sein : 54.0%,
2. cancer du sein localisé actuel (en cours de traitements) : 23.0%,
3. cancer du sein passé (rémission) : 17.2%,
4. cancer du sein métastatique : 5.9%.

On remarque sur la Figure 3.1 que les participantes sans cancer du sein sont assez jeunes en comparaison avec les participantes ayant un cancer du sein passé. Dans le tableau 3.1, nous observons que 14.1% de participantes ayant un cancer métastatique sont classées comme obèse selon l’IMC. Une tendance globale s’observe dans les 4 groupes : principalement des personnes de corpulence normale. La population de l’étude possède un niveau d’éducation élevé avec au minima un baccalauréat et 47.5% de femmes sans cancer ayant un niveau BAC +5 et plus. On peut noter que les participantes sont principalement issues de catégorie socio-professionnelle élevées avec au total 40.4% de répondantes ayant une profession intermédiaire supérieure. Enfin, on remarque par la Figure 3.2 que les

participantes sont plutôt des femmes ayant une activité physique classée comme inactive par le score issu du GSLTPAQ : en effet, nous avons respectivement pour les 4 sous-groupes 64.1%, 60.8%, 66.2% et 65.9% de femmes inactives.

Concernant les caractéristiques des traitements du cancer du sein, 87.9% des participantes ayant un cancer en cours de traitements sont en cours d'hormonothérapie. Les femmes ayant un cancer passé ont principalement eu recours à la chirurgie (89.1%).

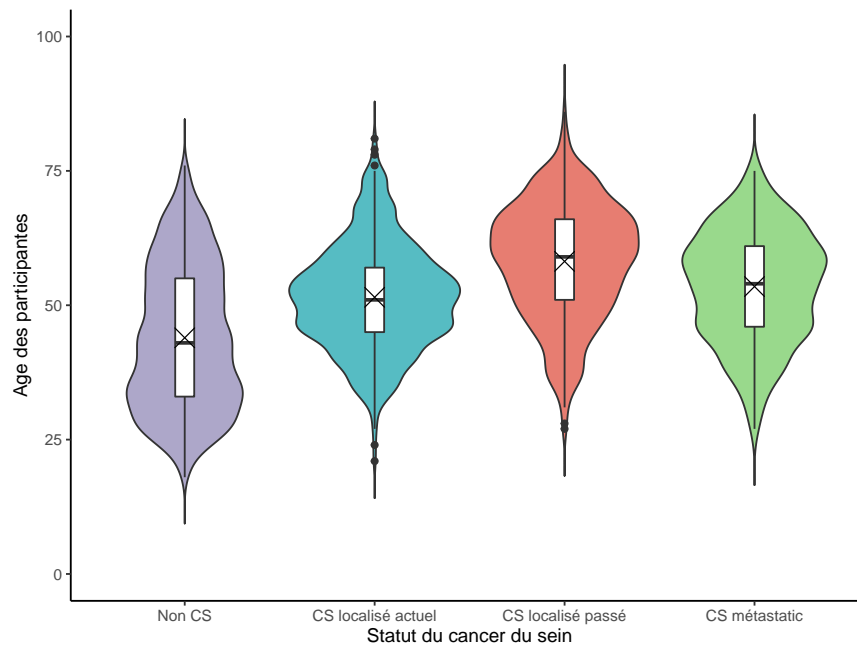


FIGURE 3.1 – Âge des participantes en fonction du statut du cancer du sein.

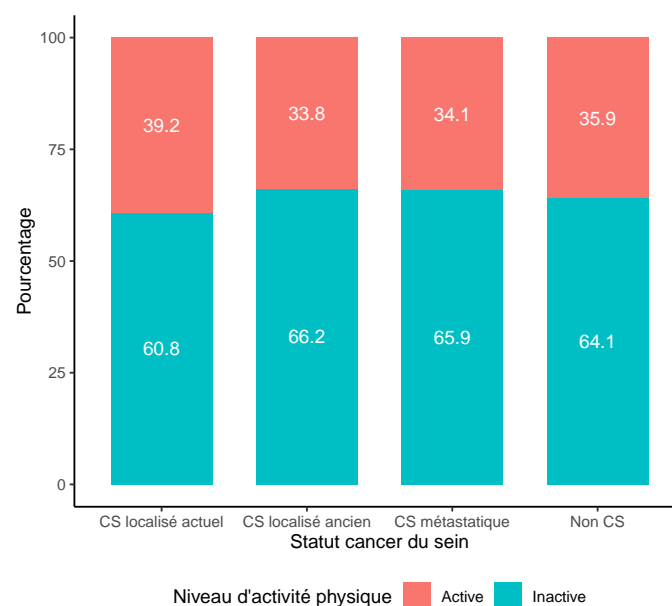


FIGURE 3.2 – Niveau d'activité des participantes en fonction du statut du cancer.

Caractéristiques	CS non	CS localisé actuel	CS localisé ancien	CS métastatique
Âge	43.9 ± 13.3	51.5 ± 9.82	58.2 ± 10.7	53.5 ± 10.3
IMC	23.5 ± 4.36	24.3 ± 4.67	23.5 ± 4.14	24.8 ± 4.94
Classification de l'IMC				
Insuffisance pondérale	84 (6.7%)	17 (3.2%)	21 (5.3%)	5 (3.7%)
Corpulence normale	820 (65.9%)	329 (62.1%)	274 (69.2%)	79 (58.5%)
Surpoids	239 (19.2%)	124 (23.4%)	70 (17.7%)	32 (23.7%)
Obésité	102 (8.2%)	60 (11.3%)	31 (7.8%)	19 (14.1%)
Statut marital				
Mariée ou pacsée	679 (54.5%)	336 (63.4%)	258 (65.2%)	79 (58.5%)
Concubinage	206 (16.5%)	56 (10.6%)	34 (8.6%)	16 (11.9%)
Divorcée	78 (6.3%)	49 (9.2%)	53 (13.4%)	19 (14.1%)
Veuve	22 (1.8%)	13 (2.5%)	12 (3.0%)	3 (2.2%)
Célibataire	260 (20.9%)	76 (14.3%)	39 (9.8%)	18 (13.3%)
Niveau d'éducation				
Pas de diplôme	7 (0.6%)	1 (0.2%)	1 (0.3%)	2 (1.5%)
CAP-BEP	40 (3.2%)	34 (6.4%)	32 (8.1%)	10 (7.4%)
BAC	89 (7.1%)	49 (9.2%)	49 (12.4%)	12 (8.9%)
BAC + 1-2	191 (15.3%)	103 (19.4%)	81 (20.5%)	30 (22.2%)
BAC + 3-4	327 (26.3%)	170 (32.1%)	109 (27.5%)	44 (32.6%)
BAC + 5 et plus	591 (47.5%)	173 (32.6%)	124 (31.3%)	37 (27.4%)
Catégorie socio-professionnelle				
Agriculteur-exploitant	9 (0.7%)	2 (0.4%)	2 (0.5%)	0 (0%)
Artisan-commerçant-chef	48 (3.9%)	17 (3.2%)	15 (3.8%)	4 (3.0%)
Cadre-profession intellectuelle	533 (42.8%)	223 (42.1%)	132 (33.3%)	44 (32.6%)
Profession intermédiaire	191 (15.3%)	85 (16.0%)	41 (10.4%)	21 (15.6%)
Employé	200 (16.1%)	89 (16.8%)	44 (11.1%)	19 (14.1%)
Ouvrier	12 (1.0%)	5 (0.9%)	2 (0.5%)	1 (0.7%)
Retraite	108 (8.7%)	71 (13.4%)	129 (32.6%)	25 (18.5%)
Sans activité	118 (9.5%)	29 (5.5%)	22 (5.6%)	17 (12.6%)
Autre	26 (2.1%)	9 (1.7%)	9 (2.3%)	4 (3.0%)
Localisation de l'habitat				
Grande ville et sa banlieue	488 (39.2%)	164 (30.9%)	114 (28.8%)	46 (34.1%)
Ville/petite ville et subdivision	324 (26.0%)	137 (25.8%)	114 (28.8%)	28 (20.7%)
Village et lotissement	258 (20.7%)	132 (24.9%)	98 (24.7%)	38 (28.1%)
Isolé/hameau	175 (14.1%)	97 (18.3%)	70 (17.7%)	23 (17.0%)
Nombre de maladies	0.67 ± 0.99	0.86 ± 1.05	1.14 ± 1.25	1.24 ± 1.31
Niveau d'AP				
Active	446 (35.9%)	207 (39.2%)	134 (33.8%)	46 (34.1%)
Inactive	800 (64.1%)	323 (60.8%)	262 (66.2%)	89 (65.9%)

TABLE 3.1 – Caractéristiques des participantes en fonction des sous-groupes.

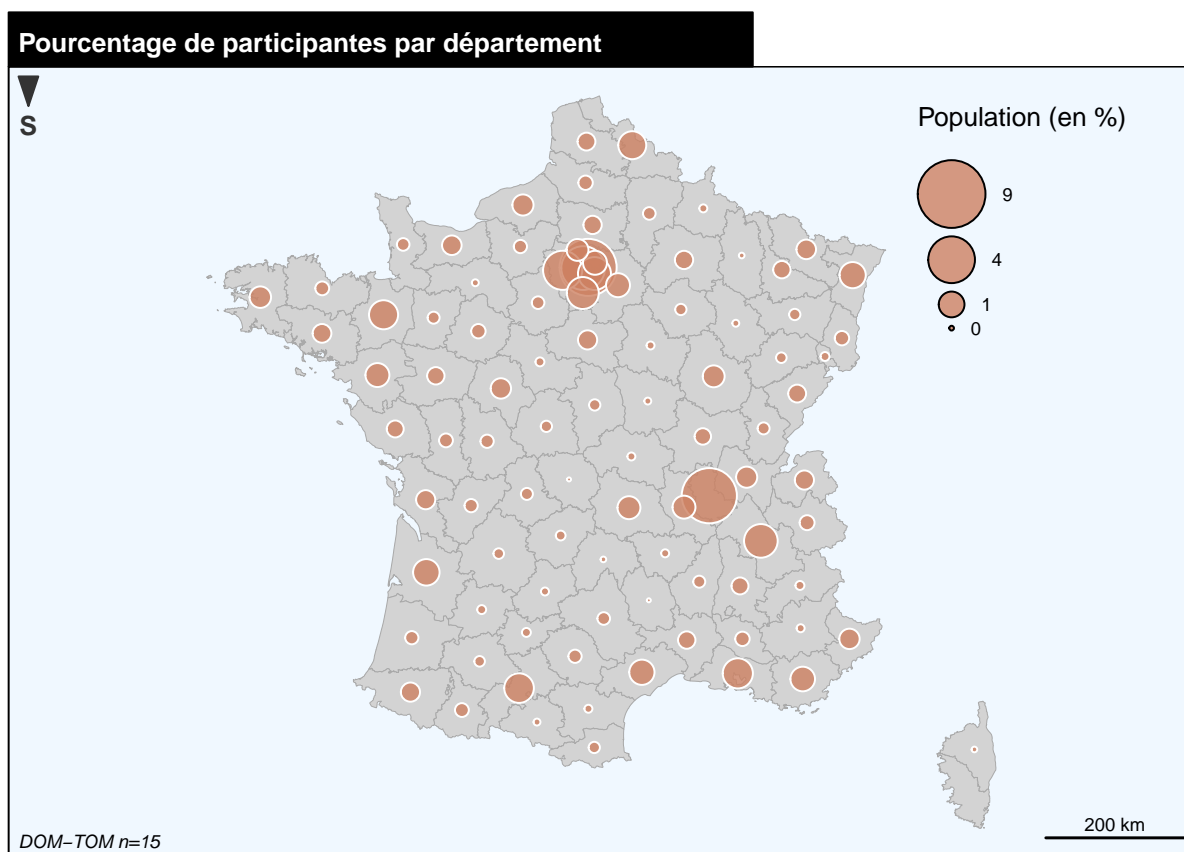


FIGURE 3.3 – Pourcentage de participantes en fonction des départements.

Intéressons-nous à la dimension géographique de la population d'étude. On remarque sur la Figure 3.3 que les participantes vivent dans les grandes villes, principalement des régions Île-de-France et Auvergne-Rhône-Alpes avec respectivement 28.4% et 18.7%. On peut noter que ces dernières régions apparaissent sur-représentées dans notre étude même si l'ensemble des participantes se dispersent dans toute la France.

1.2 Préférences, barrières et facilitateurs de la population

Conseils sur l'activité physique.

De manière générale, les répondantes de l'étude souhaitent recevoir des informations et des conseils sur la pratique d'une activité physique avec respectivement 53.4%, 70.0%, 61.4% et 78.5% pour les 4 sous-groupes. On constate sur la Figure 4 présente en Annexe C que les participantes ont la même préférence : recevoir des conseils et informations sur l'activité physique par le biais d'un spécialiste de l'activité physique affilié à un établissement de santé avec 45.2% pour les femmes sans cancer, 65.7% pour celles en cours de traitement, 50.8% pour les répondantes ayant eu un cancer et 63.7% pour celles diagnostiquées d'un cancer métastatique. Afin de recevoir ces informations, le moyen de communication mis en avant par chaque groupe de participantes est les vidéos en ligne avec respectivement 51.2%, 52.5%, 44.2% et 54.1% pour chaque sous-groupe.

Type d'activité physique privilégié.

Nous avons effectué une analyse de texte sur la variable énonçant le type d'AP que les participantes pratiquent. Sur la Figure 3.4, on peut noter que l'AP la plus pratiquée par les répondantes au questionnaire est la marche. L'étude des sous-groupes a également donné la même conclusion : quel que soit le stade de la maladie, les participantes souhaitent principalement effectuer un programme d'activité physique en lien avec la marche. La deuxième activité préférée était le yoga pour les femmes sans cancer (48,4%), le pilates pour les femmes en cours de traitement (53,6%) et les femmes en rémission (49,2%) et le renforcement musculaire pour les femmes ayant un cancer métastatique (51,1%).



FIGURE 3.4 – Type d'activité physique pratiqué par les participants.

D'autres informations concernant le lieu de pratique, l'intensité des séances ou encore l'organisation des séances sont disponibles dans le tableau 1 dans l'Annexe C.

Barrières à la pratique d'une activité physique.

La Figure 3.5 présente une visualisation des barrières à l'activité physique rencontrées par les répondantes en fonction du statut de la maladie. On peut noter que le manque de motivation est la barrière la plus souvent énoncée par les femmes sans cancer (63%), les femmes en cours de traitement (53%), celles en rémission (49%) et celles ayant un cancer métastatique (52%). Cette différence est significative avec une p -value inférieure à 5%. Exceptée la motivation, les barrières les plus fréquentes pour lesquelles les répondantes étaient "plutôt d'accord" et "tout à fait d'accord" sont la fatigue ou encore le mauvais temps. Cependant, on peut remarquer que les femmes avec un cancer métastatique ont également rapporté que la douleur due à la maladie est un frein à la pratique d'une activité (41%), différence significative (p -value= 2.2e - 16).

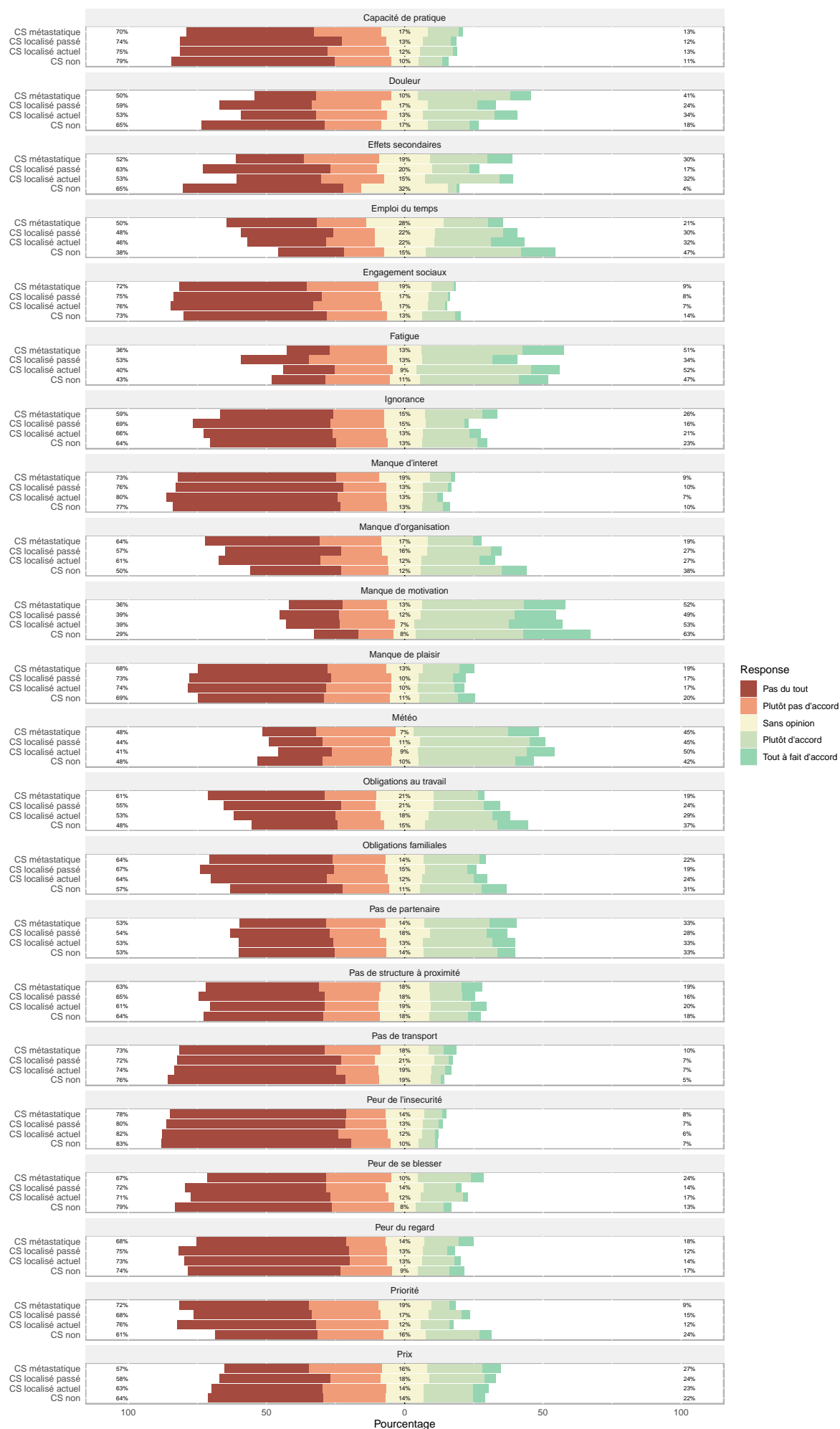


FIGURE 3.5 – Barrières à l'activité physique en fonction du statut du cancer.

Facteurs facilitant la pratique d'une activité physique.

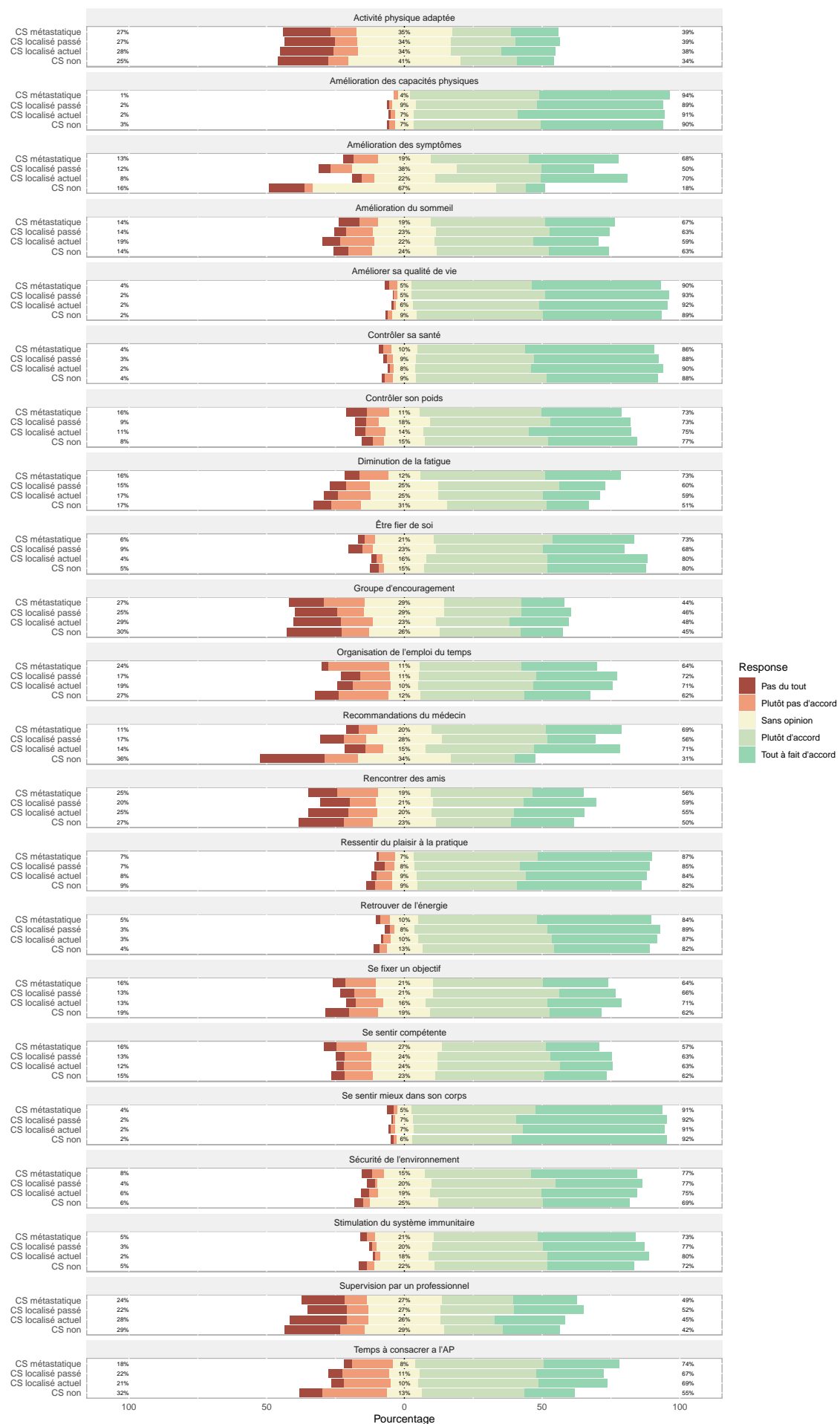


FIGURE 3.6 – Facilitateurs à l'activité physique en fonction du statut du cancer.

Intéressons-nous maintenant aux facteurs permettant de faciliter l'activité physique. Par la Figure 3.6, on constate que se sentir mieux dans son corps est le facilitateur le plus courant pour les femmes sans cancer (92%), améliorer sa qualité de vie pour les femmes en cours de traitements (92%) et en rémission (93%). Pour les répondantes avec un cancer métastatique, améliorer ses capacités semble être un facteur influant (94%). Les cinq facilitateurs les plus souvent cités par les femmes interrogées comme étant "plutôt d'accord" ou "tout à fait d'accord" pour faciliter la pratique d'une activité physique étaient "améliorer ses capacités", "être mieux dans son corps", "améliorer sa qualité de vie", "ressentir du plaisir" et "retrouver de l'énergie" pour les femmes sans cancer et pour les femmes avec un cancer métastatique. Ces facilitateurs étaient les mêmes pour les femmes en cours de traitements et en rémission sauf "retrouver de l'énergie" qui est dans les cinq premiers facilitateurs les plus courants au lieu de "ressentir du plaisir".

1.3 Population des Seintinelles

Une collaboration avec l'entreprise Capgemini spécialisée dans le service du numérique a été mise en place lors de cette étude. Le développement de 7 indices a permis de déterminer des scores de pénibilité à diverses échelles (communale, départementale, IRIS). La notion d'IRIS signifiant Îlots regroupés pour l'information statistique a été initiée par l'Institut national de la statistique et des études économiques (INSEE) et forme un découpage infra-communal de la France. Les communes d'au moins 10 000 habitants et une forte proportion des communes de 5 000 à 10 000 habitants sont découpées en IRIS. L'IRIS est la maille administrative la plus précise sur la base de laquelle sont effectuées les statistiques courantes en France [21]. Les indices sont les suivants et une description de ces derniers est présente en Annexe D :

- indice économique,
- indice d'accessibilité aux soins,
- indice socio-démographique,
- indice d'addiction,
- indice de sécurité,
- indice sur l'état de santé,
- indice sur les conditions matérielles,

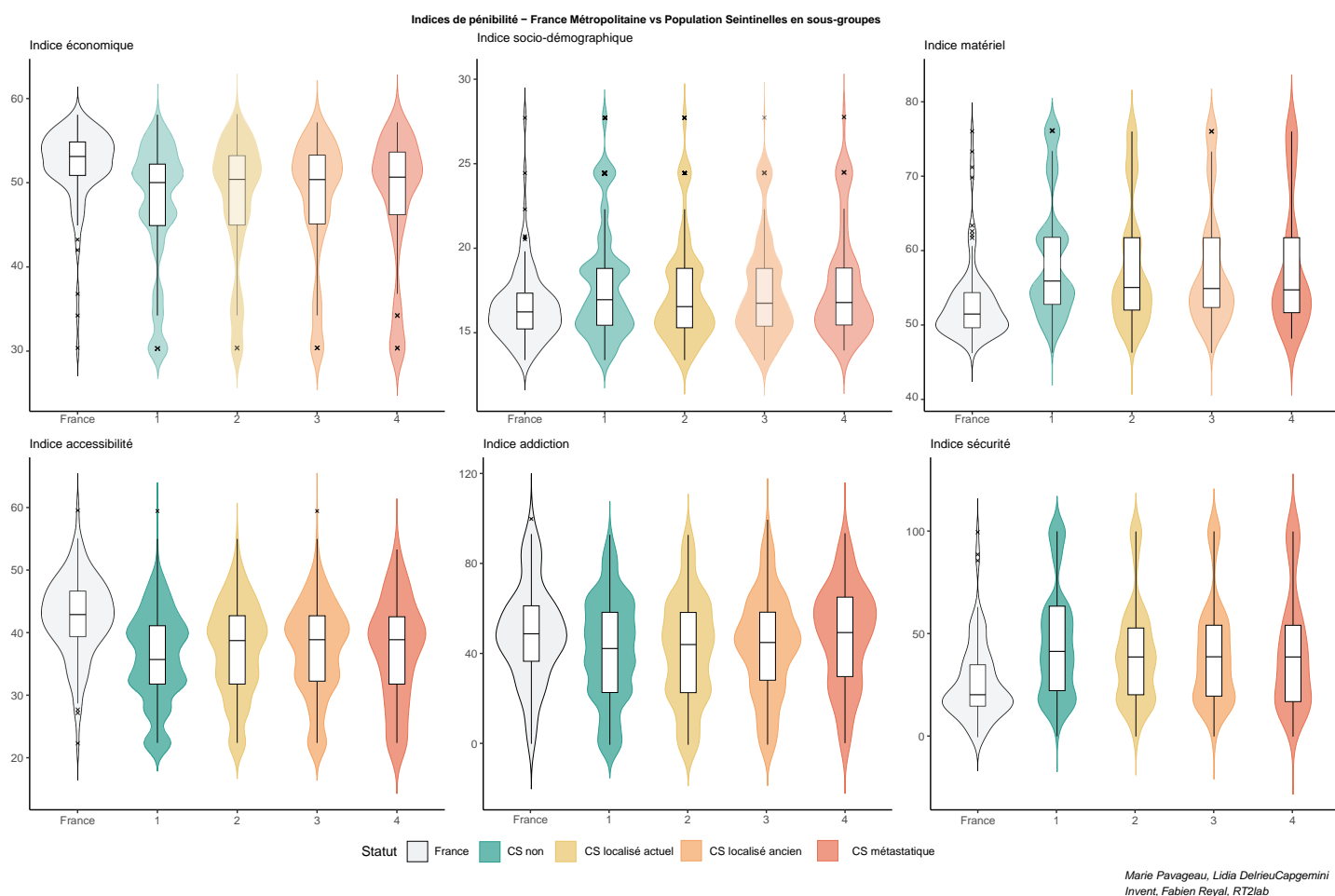


FIGURE 3.7 – Représentation des indices de pénibilités, 1ère partie (France vs population des Sentinelles).

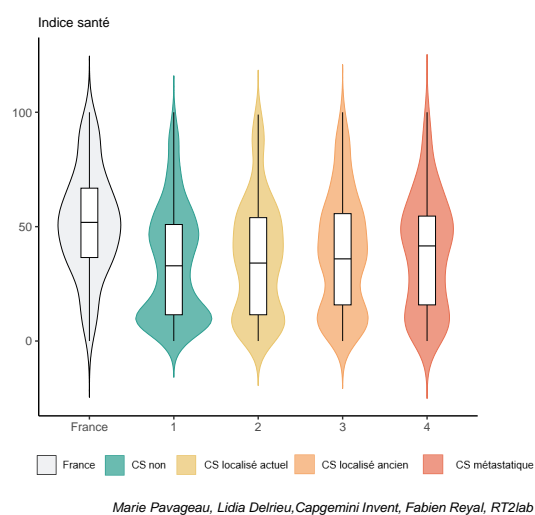


FIGURE 3.8 – Représentation des indices de pénibilités, 2ème partie (France vs population des Sentinelles).

On remarque sur les Figures 3.7 et 3.8 que les Seintinelles ayant participé à l'étude vivent dans un environnement dont la situation économique est y plus favorable qu'à l'échelle de la France avec une médiane plus basse en comparaison avec la population nationale. Étant donné que les participantes vivent principalement dans des grandes villes avec 47.1% des femmes qui habitent dans les régions Île-de-France et Auvergne-Rhône-Alpes, on observe qu'elles habitent dans des environnements plus accessibles en terme de santé, mais avec un sentiment d'insécurité plus important. Intéressons-nous à l'indice de la situation matérielle : les Seintinelles vivent dans un environnement plus précaire que la population nationale. Ceci peut s'expliquer d'une part, par le mode de transport utilisée pour se déplacer, on peut penser que dans les grandes villes beaucoup de personnes ont recourt aux transports en communs. D'autre part, étant donné le prix accrus des logement dans les grandes agglomérations, ceci peut favoriser la précarité de la situation matérielle. Enfin, les participantes vivent dans un environnement où l'état de santé moyen y est plus haut que la population française.

2 Partitionnement des participantes

Dans la section précédente, nous avons étudié les participantes en utilisant une approche médicale tandis que dans cette partie, nous effectuons une approche holistique afin de prendre en compte les participantes dans leur globalité. Nous avons appliqué la méthode de partitionnement vue précédemment afin de créer des classes de participantes présentant des caractéristiques similaires du point de vue de leurs réponses au questionnaire. Tout d'abord, afin de respecter l'homogénéité des données, nous avons pris soin de transformer les variables quantitatives en qualitatives par découpage en tranches de leurs domaines. De plus, les variables ne concernant pas les barrières, les facilitateurs et les préférences à la pratique d'une activité physique sont considérées en variables supplémentaires afin de chercher à les modéliser. Nous avons ainsi mis en supplémentaires les variables caractérisant les participantes telles que l'âge, l'IMC, le statut marital, etc.

Nous utilisons l'analyse en composantes multiples afin de supprimer le bruit du jeu de données dans le but d'effectuer une classification des individus. Nous obtenons de faibles taux d'inertie au vu des données telles que la redondance d'informations, les variables binaires. Nous retenons les composantes qui ne sont pas du bruit : on sélectionne les axes ayant un pourcentage d'inertie supérieur à $\frac{1}{m-p}$ où m est le nombre total de modalités des variables actives et p le nombre de variables qualitatives actives. Ainsi, tout axe au pourcentage d'inertie inférieur à $\frac{1}{626-201} \simeq 0.002$ est éliminé. Dans le tableau 3.2 et la Figure 3.9, nous présentons les valeurs propres ainsi que les pourcentages de variances expliqués par les 5 premières dimensions issues de l'ACM.

valeur propre	pourcentage de variance	pourcentage cumulé de la variance
0.167	7.221	7.221
0.063	2.735	9.956
0.049	2.120	12.076
0.045	1.933	14.009
0.041	1.784	15.793

TABLE 3.2 – Premières valeurs propres issues de l’ACM et pourcentage d’inerties cumulées.

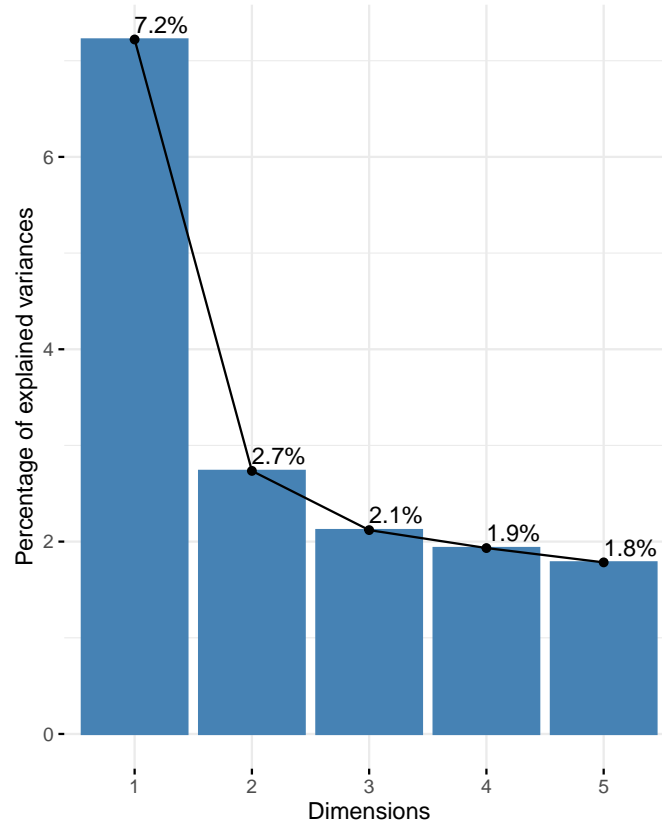


FIGURE 3.9 – Graphique des pourcentages d’inertie selon les axes.

Une valeur propre issue de l’ACM vaut en moyenne $\frac{1}{p}$ où p est le nombre de variables. En effet,

$$\frac{m}{p} - 1 = \sum_{k=1}^r \lambda_k \Leftrightarrow \frac{1}{r} \sum_{k=1}^r \lambda_k = \frac{1}{p},$$

où m est le nombre de modalités totales et $r = m - p$ le nombre de valeurs propres non nulles. Ainsi, les valeurs propres supérieures à 0.005 contribuent plus que la moyenne car nous avons 201 variables actives. Nous sélectionnons finalement les 300 premières composantes : ce choix permet de capturer environ 91% de l’information.

Nous effectuons maintenant une procédure CAH sur les résultats précédent de l’ACM. Tout d’abord, nous avons utilisé différents indices d’agrégation afin de comparer les résultats obtenus et de mettre en évidence une partition robuste. Pour ce faire, nous avons examiné les trois indices d’agrégation suivants :

- indice de Ward : $\mu(A, B) = \frac{w_A w_B}{w_A + w_B} ||\bar{x}_A - \bar{x}_B||,$
- indice de saut maximum : $\mu(A, B) = \max_{a \in A, b \in B} d(a, b),$

— indice de saut moyen : $\mu(A, B) = \frac{1}{w_A w_B} \sum_{a \in A, b \in B} w_a w_b d(a, b)$,

où A et B sont deux classes, w_A et w_B le poids des classes A et B et \bar{x}_A , \bar{x}_B les centres de gravité des classes.

Nous comparons les partitions obtenues à l'aide de l'indice de Rand. Ce dernier permet de mesurer le degré de coïncidence de deux partitions P_1 et P_2 . Soient a le nombre de paires d'éléments dans la même classe pour P_1 et pour P_2 , b le nombre de paires d'éléments dans la même classe pour P_1 , mais séparés pour P_2 , c le nombre de paires d'éléments dans la même classe pour P_2 , mais séparés pour P_1 et d le nombre de paires d'éléments séparés pour P_1 et P_2 (tableau 3.3) [6].

	groupés dans P_1	séparés dans P_1
groupés dans P_2	a	c
séparés dans P_2	b	d

TABLE 3.3 – Indice de Rand : paires d'éléments groupés et séparés dans deux partitions P_1 et P_2 [6].

L'indice de Rand est déterminé par la formule suivante : $Rand = \frac{a + d}{a + b + c + d}$.

En effectuant une CAH sur les résultats de l'ACM et en utilisant les trois indices décrits ci-dessus, nous obtenons les indices de Rand suivants : 1 entre l'indice du saut maximum et du saut moyen et 0.76 entre l'indice de Ward et l'indice de saut maximum. De ce fait, on peut considérer que la classification est assez robuste. Par soucis d'interprétabilité des classes, nous choisissons par la suite d'utiliser le critère de Ward puisque les deux autres indices partitionnent les individus en classes peu interprétables. On obtient ainsi le dendrogramme représenté dans la Figure 3.10 ci-dessous. Nous choisissons une partition en 3 classes. De plus, nous optimisons la partition obtenue à l'aide de l'algorithme des k -means afin d'avoir des classes plus robustes.

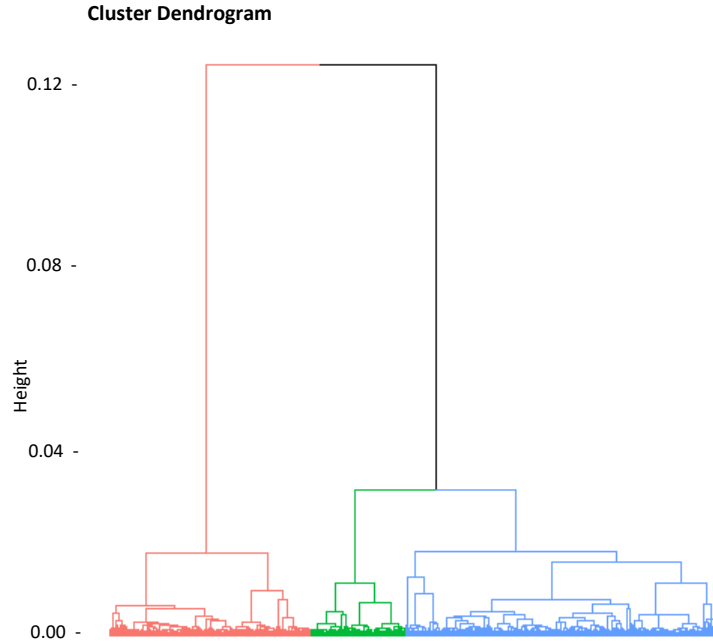


FIGURE 3.10 – Dendrogramme issu de la méthode HCPC.

On obtient ainsi 3 classes : la première contient 1124 participantes, la deuxième 431 et finalement la troisième en possède 751. Nous souhaitons à présent décrire les classes obtenues. Rappelons que les variables caractéristiques ont été mises en supplémentaires lors de l'ACM afin d'effectuer des modélisations entre celles-ci et la variable de classe.

2.1 Caractéristiques des classes

Dans le tableau 3.4, on constate que les personnes du premier groupe sont plus âgées (51.0 ± 13.5) par rapport à la classe 2 (47.3 ± 13.7) et la classe 3 (46.0 ± 11.7). Cependant, même si les participantes du troisième groupe ont une moyenne d'âge plus basse que les autres groupes, la majorité de la population a entre 30 et 50 ans (50.2%). De plus, les répondantes semblent avoir un IMC semblable dans les trois classes.

Concernant le stade du cancer du sein, les participantes de la première classe sont plutôt des femmes ayant terminé les traitements pour un cancer du sein localisé (21.0%) par rapport au groupe 2 dont la plupart des participantes n'ont pas d'historique de cancer du sein (58.5%). Pour la troisième classe, elle est principalement constituée de patientes en cours de traitements par chimiothérapie et hormonothérapie pour un cancer du sein localisé ou métastatique (35.3%). De manière générale, les trois classes sont principalement constituées de femmes inactives avec respectivement 65.6%, 64.5% et 61.0%. Enfin, le nombre moyen de comorbidités ainsi que le type de comorbidités semblent consistants entre les trois classes.

Caractéristiques	Classe 1	Classe 2	Classe 3
Âge	51.0 ± 13.5	47.3 ± 13.7	46.0 ± 11.7
Classification de l'IMC			
Insuffisance pondérale	65 (5.8%)	18 (4.2%)	44 (5.9%)
Corpulence normale	751 (66.8%)	278 (64.5%)	473 (63.0%)
Surpoids	208 (18.5%)	93 (21.6%)	164 (21.8%)
Obésité	100 (8.9%)	42 (9.7%)	70 (9.3%)
Catégorie socio-professionnelle			
Agriculteur-exploitant	7 (0.6%)	2 (0.5%)	4 (0.5%)
Artisan-commerçant-chef	41 (3.6%)	13 (3.0%)	30 (4.0%)
Cadre-profession intellectuelle	398 (35.4%)	191 (44.3%)	343 (45.7%)
Profession intermédiaire	159 (14.1%)	64 (14.8%)	115 (15.3%)
Employé	175 (15.6%)	58 (13.5%)	119 (15.8%)
Ouvrier	10 (0.9%)	0 (0.0%)	10 (1.3%)
Retraite	226 (20.1%)	53 (12.3%)	54 (7.2%)
Sans activité	79 (7.0%)	41 (9.5%)	66 (8.8%)
Autre	29 (2.6%)	9 (2.1%)	10 (1.3%)
Statut CS			
Pas de CS	603 (53.6%)	252 (58.5%)	390 (51.9%)
CS localisé actuel	231 (20.6%)	93 (21.6%)	206 (27.4%)
CS localisé ancien	236 (21.0%)	64 (14.8%)	96 (12.8%)
CS métastatique	54 (4.8%)	22 (5.1%)	59 (7.9%)
Niveau d'AP			
Active	387 (34.4%)	153 (35.5%)	293 (39.0%)
Inactive	737 (65.6%)	278 (64.5%)	458 (61.0%)
Conseils sur l'AP	484 (43.1%)	301 (69.8%)	600 (79.9%)

TABLE 3.4 – Principales caractéristiques des classes de participantes.

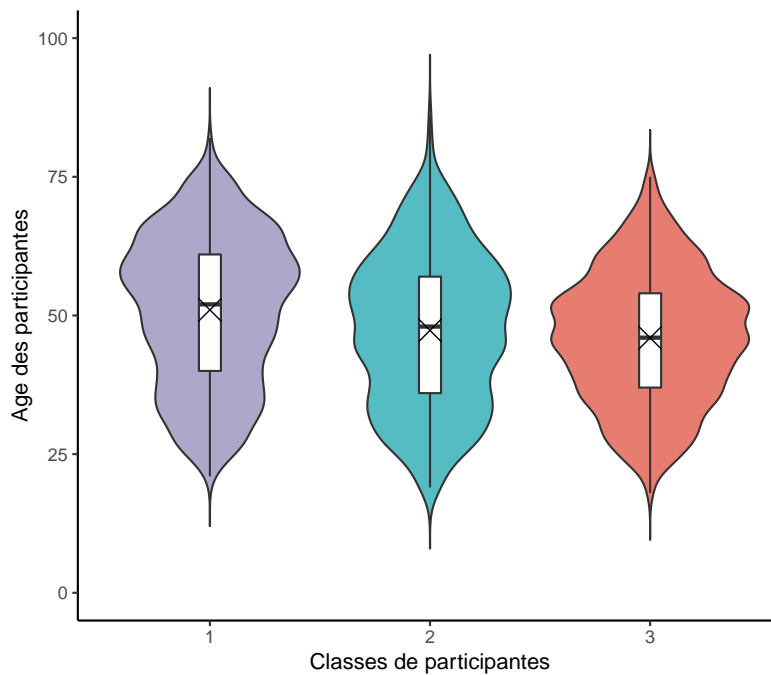


FIGURE 3.11 – Violin plot de l'âge des participantes en fonction des classes obtenues.

2.2 Étude de l'activité physique

Préférences des participantes.

On peut noter que les femmes présentes dans la troisième classe souhaitent majoritairement recevoir des conseils et des informations sur la pratique d'une activité physique (79.9%). On constate cette même tendance dans le second groupe alors que le premier semble moins intéressé avec un pourcentage de 43.1% .

On remarque que peu importe la classe, les femmes souhaitent peu recevoir des conseils de leur médecin, de leur oncologue et de leur kinésithérapeute. En revanche, les participantes de la classe 3, déjà fortement intéressées pour obtenir des conseils, aimeraient en recevoir principalement de la part d'un expert en activité physique qu'il soit enseignant en activité physique adaptée (62.8%) ou coach sportif (41.4%). Afin de transmettre des informations sur l'activité physique, les participantes du troisième groupe préféreraient utiliser une application mobile (52.9%) alors que les participantes du premier groupe souhaiteraient plutôt des conseils par mails (44.4%). De manière générale, les femmes se sentent intéressées et capables de pratiquer une activité physique adaptée à leurs capacités.

Intéressons-nous au type d'activité physique que les répondantes souhaitent pratiquer : dans les trois groupes, la marche reste l'activité physique privilégiée. On retrouve le même phénomène lors de l'étude des participantes en fonction du statut du cancer. En revanche, les sports de raquettes ainsi que les sports collectifs ne semblent pas intéresser les femmes de tous les groupes confondus.

Nouvelles technologies.

On souhaiterait par la suite pouvoir mettre en place des programmes d'activité physique tout en liant l'utilisation de technologies comme par exemple les montres connectées ou encore la pratique par le biais de lives sur les réseaux sociaux. Pour se faire, nous nous sommes intéressés à leur utilisation afin d'en dégager une première tendance. On peut relever la faible utilisation des réseaux sociaux par les femmes du premier groupe (31.0%) par rapport aux deux autres clusters (20.9% et 10.3%, différence significative). Les femmes du groupe 3 trouvent que l'utilisation de montres connectées serait assez utile pour augmenter leur niveau d'activité physique : en effet 100% des femmes issues de ce groupe souhaitent effectuer un programme d'AP via un traqueur (p -value < 5%). Ces femmes sont plus souvent utilisatrices de ces dispositifs par rapport aux deux autres clusters comme on peut le voir sur le tableau 2 présents en Annexe C. D'ailleurs, 45.9% des participantes de la troisième classe portent tous les jours une montre connectée contre respectivement 21.4% et 21.8% pour deux autres groupes. Concernant les principales raisons de ce désintérêt d'effectuer des programmes d'AP via un traqueur, une application et un site web pour les participantes du premier groupe, résident la peur que l'on utilise à tort leurs données (23.5%), le manque d'opportunité (18.1%) ou le manque d'information (14.9%). Finalement, les montres connectées sont recherchées par les participantes afin de mesurer la distance parcourue (74.8%, 71.7% et 89.3% pour les trois groupes respectivement) ou encore la fréquence cardiaque (52.8%, 58.0% et 72.0%).

Barrières et facilitateurs à l'activité physique.

Les principales barrières à l'activité physique perçues par les participantes sont principalement les mêmes énoncées lors de l'étude des sous-groupes. En effet, par la Figure 5 présente en Annexe C, le manque de motivation reste le premier obstacle à la pratique avec respectivement 52.0%, 62.0%, 64.0% pour les trois classes (p -value < 5%). Concernant les facteurs facilitant la pratique d'une AP présents sur la Figure 6 en Annexe C, se sentir mieux dans son corps ainsi que l'amélioration des capacités physiques semblent être les principaux facilitateurs.

Étude des caractéristiques avec la variable de classe.

Dans un premier temps, des modèles de régression logistique univariés ont été réalisés afin d'étudier les associations entre l'appartenance aux groupes de la variable de classe et les caractéristiques des participantes. Les variables comprenaient l'âge, l'IMC, le statut du cancer du sein, le niveau d'activité physique et le statut matrimonial. Dans un second temps, une analyse multivariée a été réalisée en effectuant une approche pas à pas, les modèles sont exposés dans le tableau 3.6.

La régression logistique permet de relier une variable qualitative Y à une série de variables explicatives. Rappelons le modèle de la régression logistique binaire : soient $Y \in \{0, 1\}$ la variable à prédire et $X = (X_1, \dots, X_N)$ les variables explicatives. On définit le logit en posant $p(1) = \mathbb{P}(Y = 1|X)$ par la formule suivante :

$$\text{logit}(p(1)) = \ln \left(\frac{p(1)}{1 - p(1)} \right) = \ln \left(\frac{p(1)}{p(0)} \right) = a_0 + a_1 x_1 + \dots + a_N x_N,$$

où x_1, \dots, x_N sont les valeurs prises par les variables explicatives X_1, \dots, X_N et a_0, \dots, a_N sont les paramètres.

Nous pouvons à présent définir la régression logistique multinomiale. Soient $Y \in \{0, \dots, L\}$ la variable qualitative à L modalités à prédire et $X = (X_1, \dots, X_N)$ les variables explicatives. On définit $p(l) = \mathbb{P}(Y = l|X)$. Etant donné la classe de référence, pour $l \in \{1, \dots, L-1\}$, on a :

$$\mathbb{P}(Y = l|X) = \frac{\exp(a_{0,l} + a_{1,l}x_1 + \dots + a_{N,l}x_N)}{1 + \sum_{l=1}^{L-1} \exp(a_{0,l} + a_{1,l}x_1 + \dots + a_{N,l}x_N)},$$

et pour $l = L$,

$$\mathbb{P}(Y = L|X) = 1 - \sum_{l=1}^{L-1} \mathbb{P}(Y = l|X),$$

de sorte que $\sum_{l=1}^L \mathbb{P}(Y = l|X) = 1$ et où $a_{0,l}, \dots, a_{N,l}$ sont les coefficients.

Dans les modèles de régression logistique, la classe 2 a été choisie comme classe de référence car elle correspond à la classe ayant le plus de participantes sans cancer du sein (58.5%). Les OR univariés pour l'appartenance aux classes 1 et 3 par rapport à la classe de référence sont présentés dans le tableau 3.5 ci-dessous.

Variables	Classe 1	Classe 3
Âge (ref : 18 – 30 ans)		
18 – 30 ans	0.69 (0.46 – 1.02)	0.68 (0.45 – 1.01)
50 – 65 ans	1.25 (0.97 – 1.60)	0.82 (0.63 – 1.07)
Plus de 65 ans	1.93 (1.34 – 2.77)	0.47 (0.30 – 0.74)
IMC (ref : normal)		
Poids insuffisant	1.34 (0.78 – 2.29)	1.44 (0.81 – 2.54)
Surpoids	0.83 (0.63 – 1.10)	1.04 (0.77 – 1.39)
Obésité	0.88 (0.60 – 1.30)	0.98 (0.65 – 1.48)
Statut CS (ref : CS non)		
CS localisé actuel	1.04 (0.78 – 1.38)	1.43 (1.07 – 1.92)
CS localisé ancien	1.54 (1.13 – 2.11)	0.97 (0.68 – 1.38)
CS métastatique	1.03 (0.61 – 1.72)	1.73 (1.04 – 2.90)
Niveau d'AP (ref : Active)		
Inactive	1.05 (0.83 – 1.32)	0.86 (0.67 – 1.10)
Statut matrimonial (ref : Célibataire)		
Mariée, pacsée	1.05 (0.77 – 1.42)	0.98 (0.71 – 1.35)
Concubinage	0.89 (0.60 – 1.33)	0.94 (0.62 – 1.43)
Divorcée	1.49 (0.92 – 2.41)	1.11 (0.66 – 1.87)
Veuve	1.62 (0.71 – 3.68)	0.70 (0.27 – 1.86)

TABLE 3.5 – Odds-Ratio des caractéristiques des participantes pour l'appartenance aux classes par rapport à la classe 2 (les résultats en bleu correspondent à des p -values inférieures à 5%).

Les résultats présents dans le tableau 3.5 montrent que le fait d'avoir 65 ans ou plus ($OR = 0.47$, 95% IC = 0.30–0.74) est significativement associé à une diminution du risque d'appartenir à la troisième classe et à une augmentation significative d'appartenir à la première classe ($OR = 1.93$, 95% IC = 1.34–2.77). De plus, avoir un cancer du sein en cours de traitement (actuel) ($OR = 1.43$, 95% IC = 1.07–1.92) et un cancer métastatique ($OR = 1.73$, 95% IC = 1.04–2.90) est significativement associé à un risque accru d'appartenir à la classe 3. Enfin, le fait d'être en rémission augmente le risque d'appartenir à la première classe ($OR = 1.54$, 95% IC = 1.13–2.11). Finalement, les résultats des analyses multivariées présents sur la Figure 3.12 sont similaires aux modèles univariés effectués plus haut.

Modèles	Variables explicatives	AIC
Modèle 1	Âge, IMC, statut matrimonial, statut du cancer, niveau d'AP	4676.3
Modèle 2	Âge, IMC, statut du cancer, niveau d'AP	4663.07
Modèle 3	Âge, statut du cancer, niveau d'AP	4659.78
Modèle 4	Âge, statut du cancer	4657.95

TABLE 3.6 – Modèles de régression logistique multinomiales et Critère d'information d'Akaike (AIC) correspondant.

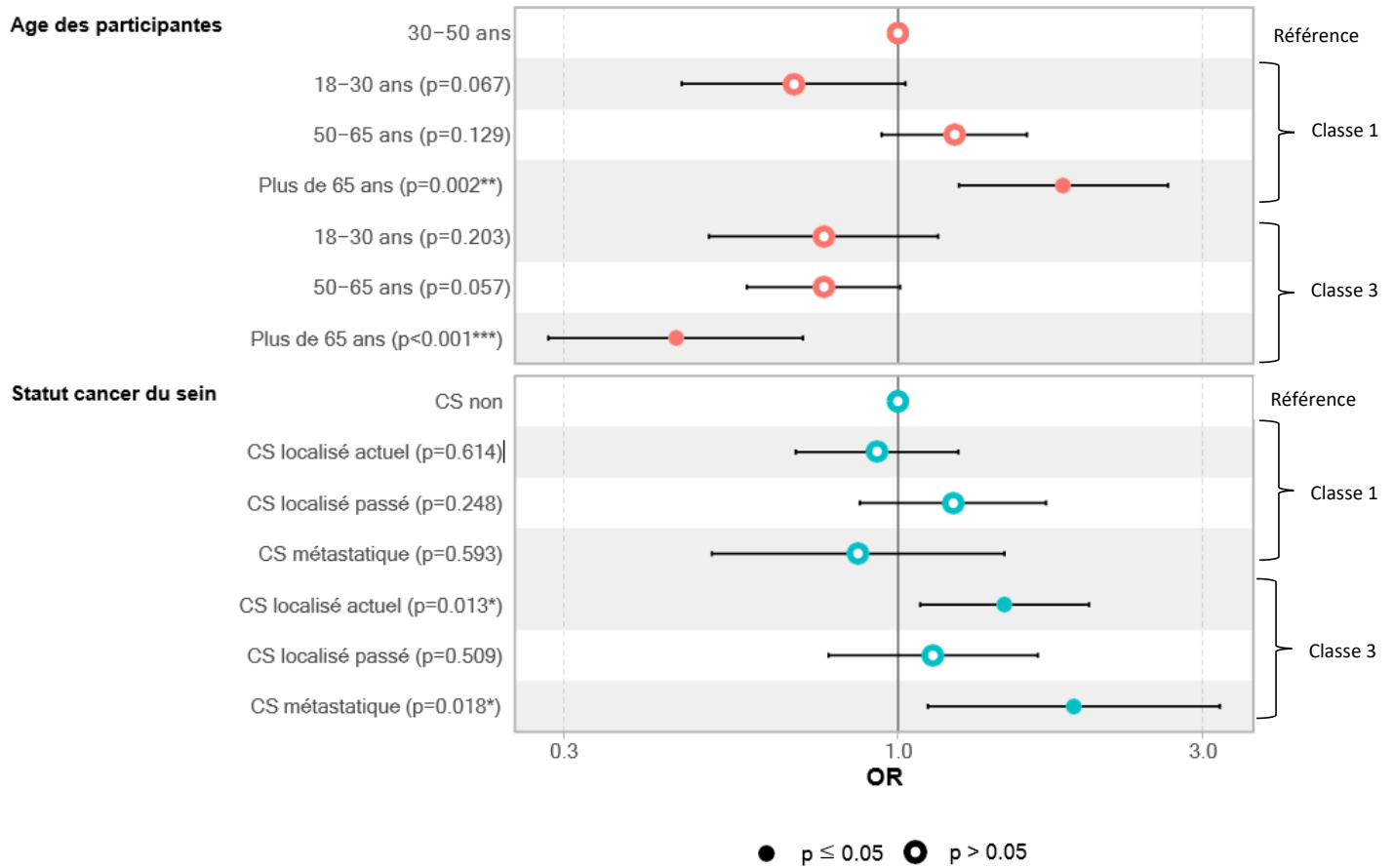


FIGURE 3.12 – Représentation graphique des OR obtenus pour la régression logistique multinomiale multivariée.

Chapitre 4

Discussion

Nouvelle approche pour la recherche. La mise en place de l'étude via la plateforme des Seintinelles apparaît comme un autre moyen de conduire une recherche : en effet, le délai de recrutement des participantes a considérablement été réduit puisque nous avons eu 2766 participations en moins de deux mois. Nous avons obtenu plus de 1500 participations la première semaine de l'étude. Ces modalités de recrutement ont permis des économies matérielles, humaines et financières.

Population d'étude. Concernant la population d'étude, on constate sur la Figure 3.3 que les régions Île-de-France et Auvergne-Rhône-Alpes sont sur-représentées. Cependant, on remarque également que les personnes ayant participé à l'étude proviennent de toute la France. De plus, toutes les catégories socioprofessionnelles ne sont pas également réparties : en effet, nous avons environ 40% de femmes cadres dans la population de l'étude. Il faut ainsi faire attention à ne pas faire des conclusions trop hâtives sur la représentativité des résultats obtenus. Enfin, l'étude présente un possible biais de sélection puisque les répondants ne sont que des personnes inscrites sur la plateforme des Seintinelles et donc sensibilisées à la recherche en cancérologie et ont généralement participé à d'autres études similaires, ce qui n'est pas forcément représentatif de la population générale.

Préférences, barrières et facilitateurs. Dans l'ensemble, les résultats ont permis de mettre en évidence que le fait d'avoir un cancer ou non, d'être en cours de traitement ou non et d'avoir un cancer localisé ou métastatique, ne semble pas influencer la plupart des préférences en termes d'activité physique. En effet, dans les 4 groupes, les participantes préféraient avoir des conseils d'un expert en AP à l'hôpital par le biais de vidéos. Elles préféraient aussi avoir un programme supervisé, à l'extérieur, d'intensité modérée, 3 jours par semaine et de 30 à 60 minutes par séance, avec la marche comme activité à privilégier. Cependant, on note que les femmes sans cancer du sein avaient certaines préférences différentes par rapport aux femmes ayant un antécédent de cancer, notamment sur le fait de moins vouloir recevoir de conseils en AP et de préférer un programme flexible avec des séances le soir. On a également pu observer que lorsque des aspects plus subjectifs pour faciliter la pratique d'activité physique étaient mis en avant, les participantes étaient plus réceptives. En effet, 90.5% des participantes ont indiqué que l'amélioration de la qualité de vie se présentait comme un facteur facilitant l'activité physique. A l'inverse, les aspects

liées à la maladie et aux traitements étaient moins facilitants. D'ailleurs, les participantes n'étaient pas forcément intéressées pour recevoir des conseils de leur médecin avec au total 85% d'entre elles ayant répondu non. Les femmes semblent avoir conscience de l'intérêt de pratiquer une activité physique régulière, mais la motivation reste un levier déterminant pour l'engagement. En effet, nous avons respectivement dans chaque sous-groupe (pas de cancer, cancer localisé actuel, cancer localisé ancien, cancer métastatique) 63%, 53%, 49%, 52% de participantes ayant rapporté que le manque de motivation était la principale source de non pratique. Il serait ainsi judicieux de mettre en place des dispositifs afin que les femmes intègrent par elles-mêmes l'activité physique dans leur quotidien.

Méthodes statistiques. Le choix de la méthode de classification a longuement été discuté. Tout d'abord, étant donné le nombre important de variables, nous nous sommes dans un premier temps tournés vers la classification de variables avant celles des individus afin de diminuer la dimension du jeu de données. Cependant, cette méthode n'a pas donné de résultats cliniquement interprétables. Ensuite, comme nous étions en présence de variables principalement qualitatives, de nombreuses méthodes n'étaient pas applicables. Nous nous sommes posé la question d'effectuer de l'encodage one-hot, mais cette méthode allait considérablement augmenter le nombre de variables du jeu de données ce à quoi nous ne voulions pas aboutir. Nous avons ainsi décidé de garder les variables originales et d'effectuer une analyse des correspondances multiples. Néanmoins, en effectuant uniquement cette méthode et en choisissant un nombre d'axes raisonnable, nous avons récupéré seulement 10% de l'information totale ce qui est très faible. Notre choix s'est finalement tourné vers la classification hiérarchique sur composantes principales qui nous a permis de supprimer le bruit du jeu de données pour ensuite effectuer un partitionnement des participantes.

Perspectives. Un de nos objectifs secondaires à la suite de ce stage est la prédiction de l'utilisation des objets connectés. En effet, on souhaiterait à partir des caractéristiques des participantes, pouvoir prédire si ces dernières seront des utilisatrices d'objets connectés, mais également si elles souhaitent participer à un programme d'activité physique. Cette démarche permettra de cibler les personnes pouvant y être intéressées et ainsi mettre en place des dispositifs facilitant la pratique d'une activité physique.

Conclusion

Afin de poursuivre ce travail sur les préférences, barrières et facilitateurs en activité physique en prévention primaire et tertiaire du cancer, il pourrait être intéressant de mettre en place de nouvelles études auprès de populations atteintes d'autres types de cancer et de les comparer pour voir les différences induites par la pathologie. On peut aussi se demander si le fait d'étudier principalement des femmes apparaît comme un facteur de confusion sur la pratique d'activité physique. De plus, il pourrait être judicieux de continuer cette étude en étudiant la pratique d'une activité physique en fonction des traitements reçus. L'analyse des résultats obtenus va ensuite faire l'objet de publications scientifiques à comité de lecture internationale par l'équipe d'encadrement. Ainsi, la suite de mon stage se concentrera sur la participation à l'écriture de ces articles. Finalement, effectuer mon stage au sein du Centre Léon Bérard m'a permis de développer ma curiosité sur le domaine de la cancérologie. Cela m'a également permis d'entrer dans une équipe pluridisciplinaire et d'avoir une idée pour concise sur le métier des biostatistiques.

Bibliographie

- [1] Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. <https://www.anses.fr/fr/system/files/NUT2012SA0155Ra.pdf>.
- [2] Agence nationale de sécurité sanitaire de l'alimentation de l'environnement et du travail (France). Étude individuelle nationale des consommations alimentaires : avis de l'ANSES, rapport d'expertise collective.
- [3] Analyse factorielle de données mixtes, J. Pagès. http://www.numdam.org/article/RSA_2004__52_4_93_0.pdf.
- [4] Foucaut AM et BERTHOUBE S ET TOUILLAUD M ET MORELLE M ET BOURNEBRANCHU V ET KEMPF-LÉPINE AS ET CARRETIER J ET PÉROL D ET TRÉDAN O ET BACHMANN P ET FERVERS B. « Deterioration of Physical Activity Level and Metabolic Risk Factors After Early-Stage Breast Cancer Diagnosis. » In : *Cancer Nurs* (2015). DOI : 10.1097/NCC.0000000000000187.
- [5] Xavier BRY. *Analyse des Correspondances Multiples*. Cours d'Analyse des Données Multidimensionnelles (HMMA127).
- [6] Xavier BRY. *Classification automatique*. Cours d'Analyse des Données Multidimensionnelles (HMMA127).
- [7] Xavier BRY. *Inertie d'un nuage*. Cours d'Analyse des Données Multidimensionnelles (HMMA127).
- [8] *Cancer : symptômes, types, en France, quelle survie ?* <https://sante.journaldesfemmes.fr/maladies/2510959-cancer-symptomes-types-definition-chiffre-france-evolution-pronostic-survie-depistage/>.
- [9] *Cancer Environnement*. <https://www.centreleonberard.fr/patient-proche/prevention-et-depistage/cancer-environnement>.
- [10] *Cancer environnement, activité physique et cancer*. <https://www.cancer-environnement.fr/178-Activite-physique-et-cancer.ce.aspx>.
- [11] *Centre Léon Bérard*. <https://www.centreleonberard.fr/>.
- [12] *Classification de l'Indice de Masse Corporelle*. <https://www.who.int/features/factfiles/obesity/facts/fr/>.
- [13] Rebecca M Speck et CYNTHIA R GROSS ET JULIA M HORMES ET REHANA L AHMED ET LESLIE A LYTLE ET WEI-TING HWANG ET KATHRYN H SCHMITZ. « Changes in the Body Image and Relationship Scale following a one-year strength training trial for breast cancer survivors with or at risk for lymphedema. » In : *Breast Cancer Res Treat* (2009). DOI : 10.1007/s10549-009-0550-7..

- [14] Jackson C et DOWD AJ ET CAPOZZI LC ET BRIDEL W ET LAU HY ET CULOS-REED SN. « A turning point : Head and neck cancer patients' exercise preferences and barriers before and after participation in an exercise intervention ». In : *Eur J Cancer Care* (2018). DOI : 10.1111/ecc.12826.
- [15] *Épidémiologie du cancer du sein*. <https://www.e-cancer.fr/Professionnels-de-sante/Les-chiffres-du-cancer-en-France/Epidemiologie-des-cancers/Les-cancers-les-plus-frequents/Cancer-du-sein>.
- [16] *Facteurs de risque du cancer du sein*. <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein/Facteurs-de-risque>.
- [17] S Amireault et G GODIN ET J LACOMBE ET C M. SABISTON. « The use of the Godin-Shephard Leisure-Time Physical Activity Questionnaire in oncology research : a systematic review ». In : (2015). DOI : 10.1186/s12874-015-0045-7.
- [18] *Global Cancer Observatory*. <https://gco.iarc.fr/>.
- [19] *Godin-Shephard Leisure-Time Physical Activity Questionnaire*. https://www.ons.org/sites/default/files/Godin%20Leisure-Time%20Exercise%20Questionnaire_070815.pdf.
- [20] L K Juvet et I THUNE ET I K Ø ELVSAAS ET E A FORS ET S LUNDGREN ET G BERTHEUSSEN ET G LEIVSETH ET L M OLDERVOLL. « The effect of exercise on fatigue and physical functioning in breast cancer patients during and after treatment and at 6 months follow-up : A meta-analysis ». In : *Breast* (2017). DOI : 10.1016/j.breast.2017.04.003.
- [21] *Îlots regroupés pour l'information statistique*. https://fr.wikipedia.org/wiki/%C3%8Elots_regroup%C3%A9s_pour_l%27information_statistique.
- [22] *Institut National du Cancer*. <https://www.e-cancer.fr/>.
- [23] Daniel Y T Fong et JUDY W C HO ET BRYANT P H HUI ET ANTOINETTE M LEE ET DUNCAN J MACFARLANE ET AL. « Physical activity for cancer survivors : meta-analysis of randomised controlled trials. » In : *BMJ* (2012). DOI : 10.1136/bmj.e70.
- [24] Rebecca M Speck et KERRY S COURNEYA ET LOUISE C MÂSSE ET SUE DUVAL ET KATHRYN H SCHMITZ. « An update of controlled physical activity trials in cancer survivors : a systematic review and meta-analysis. » In : *J Cancer Surviv* (2020). DOI : 10.1007/s11764-009-0110-5..
- [25] *Les Seintinelles*. <https://www.seintinelles.com/>.
- [26] MD et Wendy Y. Chen et Diane Feskanich et al MICHELLE D. HOLMES. « Physical Activity and Survival After Breast Cancer Diagnosis ». In : (2005). DOI : 10.1001/jama.293.20.2479.
- [27] Saskia F A Duijts et MIZJA M FABER ET HESTER S A OLDENBURG ET MARC VAN BEURDEN ET NEIL K AARONSON. « Effectiveness of behavioral techniques and physical exercise on psychosocial functioning and health-related quality of life in breast cancer patients and survivors—a meta-analysis ». In : *Psychooncology* (2011). DOI : 10.1002/pon.1728.

- [28] Clifford BK et MIZRAHI D ET SANDLER CX ET BARRY BK ET SIMAR D ET WAKEFIELD CE ET AL. « Barriers and facilitators of exercise experienced by cancer survivors : a mixed methods systematic review ». In : *Support Care Cancer Off J Multinatl Assoc Support Care Cancer* (2018). DOI : 10.1007/s00520-017-3964-5.
- [29] *Organisme Mondial de la Santé, activité physique*. <https://www.who.int/fr/news-room/fact-sheets/detail/physical-activity>.
- [30] M Carayol et P BERNARD ET J BOICHÉ ET F RIOU ET B MERCIER ET F COUSSON-GÉLIE ET AJ ROMAIN ET C DELPIERRE ET G NINOT. « Psychological effect of exercise in women with breast cancer receiving adjuvant therapy : what is the optimal dose needed ? » In : *Ann Oncol* (2013). DOI : 10.1093/annonc/mds342.
- [31] Sablone L et Jacob G et Reyat F et al PANNARD M ET BAUQUIER C ET BASSOLEIL L. « Ces citoyen(nes) qui se portent volontaires pour participer aux recherches dans le champ du cancer, résultats du Baromètre Seintinelles 2018 ». In : *Bulletin du Cancer* (2020). DOI : 10.1016/j.bulcan.2019.11.012.
- [32] Adams SC et SEGAL RJ ET MCKENZIE DC ET VALLERAND JR ET MORIELLI AR ET MACKEY JR ET AL. « Impact of resistance and aerobic exercise on sarcopenia and dynapenia in breast cancer patients receiving adjuvant chemotherapy : a multicenter randomized controlled trial. » In : *Breast Cancer Res Treat* (2016). DOI : 10.1007/s10549-016-3900-2.
- [33] *Traitements anticancéreux*. <https://www.e-cancer.fr/Patients-et-proches/Se-faire-soigner/Traitements>.

A Questionnaire



Préférences, barrières et facilitateurs à l'activité physique des femmes

Prévention primaire et
tertiaire d'un cancer
du sein

Questionnaires

Dans ce projet, il vous sera demandé de remplir des questionnaires portant sur 6 domaines :

- vos caractéristiques sociodémographiques, de lieu de vie et de santé,
- votre niveau d'activité physique,
- vos préférences en matière d'activité physique,
- votre utilisation des technologies liées à l'activité physique,
- vos barrières à la pratique d'activité physique,
- vos facilitateurs à la pratique d'activité physique.

DOMAINE 1 : QUESTIONNAIRE SOCIODÉMOGRAPHIQUE, DE LIEU DE VIE ET DE SANTÉ

Nous allons tout d'abord vous poser quelques questions pour mieux connaître votre situation de vie et votre santé.

1. Quelle est votre année de naissance ? : __ __
2. Quelle est votre situation familiale actuelle ?
 - a) Mariée
 - b) En concubinage
 - c) Divorcée
 - d) Veuve
 - e) Célibataire
3. Quel est votre diplôme d'études le plus élevé ?
 - a) Pas de diplôme
 - b) CAP ou BEP
 - c) Bac général/professionnel/technologique
 - d) Bac+1, bac+2
 - e) Bac +3, bac +4
 - f) Bac +5 et plus
4. Quelle est votre situation professionnelle actuelle principale ?
 - a) En activité à temps plein
 - b) En activité à temps partiel
 - c) En congé maladie
 - d) En invalidité
 - e) Retraitée
 - f) Etudiante
 - g) A la recherche d'un emploi
 - h) Ne travaille pas
5. Quelle est votre catégorie socio-professionnelle actuelle ?
 - a) Agriculteur exploitant
 - b) Artisan, commerçant ou chef d'entreprise (professionnel indépendant)

- c) Cadre et profession intellectuelle supérieure (cadre A de la fonction publique, cadre administratif et commercial d'entreprise, profession libérale, professeurs, ingénieur...)
- d) Profession intermédiaire (instituteur, profession intermédiaire des entreprises, cadre B de la fonction publique, technicien...)
- e) Employé (cadre C et D de la fonction publique, employé administratif d'entreprise, employé de commerce, personnel des services aux particuliers, militaire, policier...)
- f) Ouvrier qualifié ou non qualifié
- g) Retraité
- h) Sans activité professionnelle (étudiant, chômeur, autres...)
 - Veuillez préciser _____
- i) Autre catégorie socio-professionnelle
 - Veuillez préciser _____

6. Quelle est votre département de résidence ? : _ _ _

7. Quelle est la proposition qui correspond **le mieux** à votre logement principal ?

- a) Dans une maison
- b) Dans un petit collectif (jusqu'à 3 étages)
- c) Dans un immeuble
- d) Dans une grande tour

8. Quelle est la proposition qui correspond **le mieux** à l'endroit où est situé votre logement ?

- a) Au centre d'une grande ville
- b) Dans un quartier situé entre le centre d'une grande ville et sa banlieue
- c) Dans la banlieue d'une grande ville
- d) Au centre d'un bourg ou d'une petite ville
- e) Dans un lotissement d'un bourg ou d'une petite ville
- f) Au cœur d'un village
- g) Dans un lotissement d'un village
- h) Mon logement est isolé, ou situé dans un hameau

9. Quelle est votre taille, en cm ? : _ _ _

10. Quel est votre poids à jeun le matin, en kg ? : _ _ _

11. Êtes-vous enceinte ?

- a) Oui
- b) Non

12. Si vous avez des enfants, combien d'entre eux vivent encore à votre domicile ? _ _ _

13. (Si q12 différent de 0) Quel âge ont-ils ? _ _ _

14. Êtes-vous... ? :

- a) Une femme qui n'a jamais eu de cancer du sein (Poursuivre avec q17)
- b) Une femme actuellement atteinte d'un cancer du sein

c) Une femme anciennement atteinte d'un cancer du sein

15. (Si b ou c à q14) Quel est le stade de votre maladie ?

- a) Localisé
- b) Métastatique

16. (Si b ou c à q14) Quels traitements avez-vous reçus depuis le diagnostic ?

- a) Chirurgie
 - Oui
 - Non
- b) Chimiothérapie
 - Finie
 - En cours
 - Pas de traitement par chimiothérapie
- c) Radiothérapie
 - Finie
 - En cours
 - Pas de traitement par radiothérapie
- d) Hormonothérapie
 - Finie
 - En cours
 - Pas de traitement par hormonothérapie
- e) Immunothérapie
 - Finie
 - En cours
 - Pas de traitement par immunothérapie
 - Je ne sais pas
- f) Thérapie ciblée
 - Finie
 - En cours
 - Pas de traitement par thérapie ciblée
 - Je ne sais pas

17. Rencontrez-vous actuellement les problèmes de santé suivants ?

	Oui	Non
Maladies cardiovasculaires (par exemple, hypertension artérielle, troubles du rythme cardiaque)		
Maladies neurologiques (par exemple, maladie de Parkinson) et neurovasculaires (par exemple, accidents vasculaires cérébraux)		
Pathologie endocrinienne (par exemple, diabète, hypothyroïdie)		
Maladie des voies respiratoires (par exemple, asthme)		
Maladies inflammatoires (par exemple, maladie de Crohn)		
Maladies auto-immunes (par exemple, lupus)		

Maladies musculosquelettiques (par exemple, ostéoporose)		
Troubles de l'humeur (par exemple, dépression, troubles psychiatriques)		
Chirurgie musculosquelettique (colonne vertébrale, jambes)		
Autre cancer que le cancer du sein ➤ Veuillez préciser _____		
Autres problèmes de santé importants ➤ Veuillez préciser _____		

DOMAINE 2 : QUESTIONNAIRE DU NIVEAU D'ACTIVITÉ PHYSIQUE

Nous allons à présent vous interroger sur votre pratique d'activité physique.

18. Considérez une période typique d'une semaine. Combien de fois, en moyenne, pratiquez-vous les types d'activités physiques suivantes pendant plus de 15 minutes durant vos temps libres ?
- a) ACTIVITÉ PHYSIQUE D'INTENSITÉ ÉLEVÉE (le cœur bat rapidement, j'ai du mal à parler) (exemple : course à pied ou jogging, natation vigoureuse, cyclisme de longue distance vigoureux, football, squash, basket-ball, ski de fond, judo) : _ _ _ (fois par semaine)
 - b) ACTIVITÉ PHYSIQUE D'INTENSITÉ MODÉRÉE (je suis essouffée et je suis très limitée pour parler) (exemple : marche rapide, vélo de promenade, danses populaires, volley-ball, natation facile, ski alpin, tennis, badminton): _ _ _ (fois par semaine)
 - c) ACTIVITÉ PHYSIQUE D'INTENSITÉ FAIBLE (je peux avoir une discussion normale et je ne ressens pas particulièrement d'essoufflement) (exemple : marche lente, yoga, tir à l'arc, bowling, promenade à cheval, golf) : _ _ _ (fois par semaine)
19. Considérez une période typique d'une semaine. En dehors de votre travail ou de vos tâches ménagères, combien de jours consacrez-vous à des activités physiques de loisirs spécifiquement conçues pour renforcer vos muscles, tels que la musculation avec des poids ou le circuit training (ne pas inclure les exercices cardio tels que la marche, le vélo ou la natation) ?
- a) 0
 - b) 1 fois par semaine
 - c) 2 fois par semaine
 - d) 3 fois par semaine
 - e) 4 fois par semaine
 - f) 5 fois par semaine
 - g) 6 fois par semaine
 - h) 7 fois par semaine
20. Considérez une période typique d'une semaine. Quelles sont les activités physiques que vous pratiquez pendant plus de 15 minutes durant vos temps libres : _ _ _ _ _
_ _ _ _ _
21. Est-ce que la crise sanitaire du COVID-19 a eu une influence sur votre niveau d'activité physique actuel ?
- a) Oui
 - Veuillez préciser
 - _____
 - _____
 - _____
 - b) Non

DOMAINE 3 : QUESTIONNAIRE DES PRÉFÉRENCES EN MATIÈRE D'ACTIVITÉ PHYSIQUE

Maintenant que nous vous connaissons un peu mieux, nous aimerions connaître vos préférences en matière d'activité physique afin de nous aider à construire des programmes d'activité physique qui prennent en compte vos suggestions. Nous allons tout d'abord vous questionner sur la façon dont vous aimeriez recevoir des conseils et des informations sur l'activité physique.

22. Cela vous intéresserait-il de recevoir des **INFORMATIONS** et des **CONSEILS** sur la manière de pratiquer une activité physique ? (une seule possibilité)
- a) Oui
 - b) Non
 - c) Peut-être/ je ne suis pas sûre
23. Parmi ces propositions, de qui aimeriez-vous recevoir des **INFORMATIONS** et des **CONSEILS** sur la manière de pratiquer ? (cochez toutes les réponses applicables)
- a) Médecin de famille
 - b) Oncologue
 - c) Infirmière
 - d) Spécialiste de l'activité physique affilié à un établissement de santé (hôpitaux, maison médicale de santé, centre de rééducation...)
 - e) Spécialiste de l'activité physique dans un centre de loisirs ou une salle de sport
 - f) Patient ou groupe de patient atteint(s) de cancer
 - g) Entraîneur personnel
 - h) Kinésithérapeute
 - i) Je n'ai pas de préférence
 - j) Autres
- Veuillez préciser _____
24. Parmi ces propositions, par quel moyen souhaiteriez-vous recevoir des **INFORMATIONS** et des **CONSEILS** sur la manière de pratiquer ? (cochez toutes les réponses applicables)
- a) Téléphone
 - b) Sms
 - c) Application mobile
 - d) Email
 - e) Page web
 - f) DVD ou Blu-Ray
 - g) Vidéos en ligne
 - h) Matériel écrit (par exemple, brochure, dépliant, livre, cahier d'exercices)
 - i) Vidéoconférence en ligne
 - j) Podcast
 - k) Discussion en tête-à-tête
 - l) Ateliers de groupe

Pour les questions ci-dessous, merci d'indiquer les activités/les modalités qu'il vous plairait de pratiquer, même si elles sont différentes de ce que vous faites actuellement. Par exemple, si vous faites du vélo mais que vous préféreriez pratiquer du yoga, alors indiquez yoga.

25. Cela vous intéresserait-il de pratiquer une activité physique adaptée à vos capacités ? (une seule possibilité)
- d) Oui
 - e) Non
 - f) Peut-être/ je ne suis pas sûre
26. Seriez-vous capable de pratiquer une activité physique adaptée à vos capacités ? (une seule possibilité)
- g) Oui
 - h) Non
 - i) Peut-être/ je ne suis pas sûre
27. Parmi ces propositions, quels types d'activité physique voudriez-vous **PRATIQUER** ? (cochez toutes les réponses applicables)
- a) Marche
 - b) Natation, nage
 - c) Gymnastique
 - d) Cyclisme, vélo
 - e) Danse
 - f) Jogging ou course à pied
 - g) Renforcement musculaire/musculation (poids libres, machines, bandes élastiques, etc.)
 - h) Yoga
 - i) Pilates
 - j) Cours collectif type aérobic
 - k) Sports de raquette
 - Veuillez préciser _____
 - l) Sports collectifs
 - Veuillez préciser _____
 - m) Autres
 - Veuillez préciser _____
28. Parmi ces propositions, où aimeriez-vous **PRATIQUER** votre activité physique ? (cochez toutes les réponses applicables)
- a) En plein air (par exemple, un sentier de randonnée ou un parc)
 - b) A la maison
 - c) Au travail
 - d) Dans un club de sport ou un centre de loisirs
 - e) Dans une salle de sport au sein d'un établissement de santé
 - f) Je n'ai pas de préférence
 - g) Autres
 - Veuillez préciser _____
29. Parmi ces propositions, comment préféreriez-vous le plus **PRATIQUER** votre activité physique ? (une seule réponse possible)

- a) Seule
 - b) En binôme ou avec un proche
 - c) Avec d'autres patients dans un groupe
 - d) Avec d'autres personnes non malades dans un groupe
 - e) Combinaison de séances individuelles et de séances avec d'autres patients en groupe
 - f) Combinaison de séances individuelles et de séances avec d'autres personnes non malades en groupe
 - g) Combinaison de séances individuelles et de séances en binôme ou avec un proche
 - h) Je n'ai pas de préférence
 - i) Autres
- Veuillez préciser _____

30. Parmi ces propositions, de quelle manière souhaiteriez-vous le plus être supervisée pour **PRATIQUER** votre activité physique ? (**une seule réponse possible**)

- a) Supervisée en présentiel par un professionnel de l'activité physique
- b) Supervisée à distance par un professionnel de l'activité physique
- c) Non supervisée, en autonomie
- d) Supervisée en présentiel pour les premières sessions puis supervisée à distance (b)
- e) Supervisée en présentiel (a) pour les premières sessions puis non supervisée, en autonomie (c)
- f) Supervisée à distance (b) pour les premières sessions puis non supervisée, en autonomie (c)
- g) Je n'ai pas de préférence

31. Parmi ces propositions, à quelle intensité aimeriez-vous le plus **PRATIQUER** votre activité physique ? (**une seule réponse possible**)

- a) Intensité faible (EFFORT MINIMAL -> pas d'essoufflement, pas de transpiration, effort ressenti de 3 à 4 sur une échelle de 0 à 10)
- b) Intensité modérée (PAS EXTÉNUANTE -> essoufflement modéré, transpiration modérée, conversation possible, effort ressenti de 5 à 6 sur une échelle de 0 à 10, peut être maintenu 30 à 60 minutes)
- c) Intensité élevée (LE CŒUR BAT RAPIDEMENT -> essoufflement marqué, conversation difficile, transpiration abondante, effort ressenti de 7 à 8 sur une échelle de 0 à 10, ne peut être maintenue plus de 30 minutes)
- d) Je n'ai pas de préférence

32. Parmi ces propositions, quelle serait l'organisation la plus adaptée pour **PRATIQUER** votre activité physique ? (**une seule réponse possible**)

- a) Flexible (des jours et horaires qui varient)
- b) Planifiée (par exemple tous les jours à la même heure)
- c) Je n'ai pas de préférence

33. Parmi ces propositions, quels types d'activités voudriez-vous **PRATIQUER** dans votre programme ? (**une seule réponse possible**)

- a) Les mêmes activités physiques à chaque fois
- b) Des activités physiques différentes à chaque fois
- c) Je n'ai pas de préférence

34. Parmi ces propositions, quels types d'activités voudriez-vous **PRATIQUER** dans votre programme ? (une seule réponse possible)
- a) Des activités physiques de compétition
 - b) Des activités physiques de loisirs
 - c) A la fois des activités physiques de compétition et des activités physiques de loisirs
 - d) Je n'ai pas de préférence
35. Parmi ces propositions, quand préféreriez-vous le plus **PRATIQUER** votre activité physique ? (une seule réponse possible)
- a) Matin
 - b) Pendant l'heure du déjeuner
 - c) Après-midi
 - d) Soirée
 - e) Je n'ai pas de préférence
36. Parmi ces propositions, quand préféreriez-vous le plus **PRATIQUER** votre activité physique ? (une seule réponse possible)
- a) Du samedi au dimanche (week-end)
 - b) Du lundi au vendredi
 - c) Tout au long de la semaine
 - d) Je n'ai pas de préférence
37. Parmi ces propositions, quel est le nombre maximum de jours durant lesquels vous seriez prêt à **PRATIQUER** votre activité physique par semaine ? (une seule réponse possible)
- a) 1
 - b) 2
 - c) 3
 - d) 4
 - e) 5
 - f) 6
 - g) 7
38. Parmi ces propositions, pendant combien de temps seriez-vous prêt à **PRATIQUER** votre activité physique par séance ? (une seule réponse possible)
- a) Moins de 15 minutes
 - b) De 15 à 30 minutes
 - c) De 30 à 60 minutes
 - d) Plus de 60 minutes

39. Parmi ces propositions, dans quelle mesure pensez-vous que chacune des stratégies suivantes vous serait utile pour **PRATIQUER** votre activité physique ?

	Pas utile	Pas très utile	Sans opinion	Plutôt utile	Très utile
Cours d'activité physique en groupe, en ligne et en direct par un professionnel de l'activité physique					
Cours d'activité physique individuel, en ligne et en direct par un professionnel de l'activité physique					
Cours d'activité physique mis en ligne par un professionnel de l'activité physique afin d'être réalisé en différé en autonomie					
Compétitions avec d'autres personnes dans une intervention de type « Gaming »* (*Pokémon GO, défis, énigmes, jeu de l'oie...)					
Appels téléphoniques réguliers par un professionnel de l'activité physique					
SMS d'encouragement					
E-mails d'encouragement					
Brochure sur les recommandations et les bénéfices de l'activité physique sur la fatigue, la qualité de vie, la condition physique etc.					
Vidéos de témoignages de personnes vivant la même situation que vous					
Motivation à pratiquer via les réseaux sociaux (Facebook, Twitter, Instagram, etc.)					
Utilisation de montres connectées ou podomètres (par ex. FitBit, podomètre, Jawbone)					
Cahier vous permettant de noter et suivre vos activités physiques réalisées					
Application vous permettant de noter et suivre vos activités physiques réalisées					
Page web vous permettant de noter et suivre vos activités physiques réalisées					
Possibilité de partager les informations sur vos activités physiques avec votre entourage					
Possibilité de partager les informations sur vos activités physiques avec votre médecin ou un autre professionnel					
Document téléchargeable en ligne avec des exercices à réaliser en autonomie					
Clé USB avec des exercices à réaliser en autonomie					
DVD avec des exercices à réaliser en autonomie					

40. Est-ce que la crise sanitaire du COVID-19 a eu un impact sur vos préférences actuelles en activité physique ?

a) Oui



➤ Veuillez préciser

b) Non

DOMAINE 4 : QUESTIONNAIRE SUR L'UTILISATION DES TECHNOLOGIES LIÉES À L'ACTIVITÉ PHYSIQUE

A présent, nous allons vous poser des questions concernant l'utilisation des nouvelles technologies. Vous trouverez le terme "traqueur d'activité physique" aussi appelé bracelet connecté ou moniteur d'activité. Il s'agit d'un appareil électronique, souvent relié à une application téléphonique, qui permet de mesurer l'intensité et la quantité d'activité physique effectuée par un individu tels que la distance marchée ou courue, la consommation et la dépense de calories, et dans certains cas le rythme cardiaque et la qualité du sommeil.

41. Parmi les réseaux sociaux suivants, lesquels utilisez-vous ? (Cochez toutes les réponses applicables)
- a) Facebook
 - b) Twitter
 - c) Instagram
 - d) LinkedIn
 - e) Autres
 - Veuillez préciser _____
 - f) Je n'utilise pas les réseaux sociaux
42. Parmi les applications suivantes liées à l'activité physique, lesquelles utilisez-vous ? (Cochez toutes les réponses applicables)
- a) FizzUp
 - b) Nike Training Club
 - c) Adidas running
 - d) Strava
 - e) Fitbit
 - f) Garmin
 - g) Jawbone
 - h) Withings
 - i) Xiaomi
 - j) Application santé du téléphone portable
 - k) Runtastic
 - l) Autres
 - Veuillez préciser _____
 - m) Je n'utilise aucune application liée à l'activité physique
43. Parmi les traqueurs d'activité physique suivants, lesquels possédez-vous ? (Cochez toutes les réponses applicables)
- a) Podomètre du téléphone portable
 - b) Podomètre classique (compte les pas)
 - c) FitBit
 - d) Jawbone
 - e) Apple Watch
 - f) Galaxy Watch
 - g) Garmin
 - h) Xiaomi
 - i) Withings

- j) Autres
➤ Veuillez préciser _____
- k) Je ne possède pas de traqueur d'activité
44. (si a à j à q43), À quelle fréquence portez-vous normalement ce traqueur d'activité ? (une seule réponse possible)
- a) Jamais
 - b) Moins d'un jour par semaine
 - c) 1 à 2 jours par semaine
 - d) 3-4 jours par semaine
 - e) 5-6 jours par semaine
 - f) Tous les jours
45. Souhaiteriez-vous participer à un programme d'activité physique réalisé à distance via un traqueur d'activité physique ? (une seule réponse possible).
- a) Oui
 - b) Non
 - c) Peut-être/ je ne suis pas sûre
46. (si b ou c à q45), Pourquoi pourriez-vous ne pas être intéressée par ce type de programme réalisé à distance via un **traqueur d'activité physique** ? (cochez toutes les réponses applicables).
- a) J'ai peur de ne pas savoir m'en servir
 - b) J'ai eu peu d'information à ce sujet
 - c) Le prix est trop élevé
 - d) Je crains que mes données soient utilisées à mauvais escient
 - e) Je crains la dépendance à l'objet
 - f) Ces objets ne me semblent pas fiables
 - g) Je n'aime pas porter quelque chose au poignet
 - h) Je suis souvent allergique à certains produits portés au poignet
 - i) Je n'ai pas eu l'occasion d'en avoir un
 - j) Ces objets ne sont pas très conviviaux
 - k) Cela me prendra trop de temps à utiliser ces objets
 - l) Ces objets ne fonctionnent pas encore très bien
 - m) Je ne pourrais pas suivre un programme où je dois faire de l'activité physique par moi-même
 - n) Je fais déjà suffisamment d'activité physique par moi-même
 - o) Je ne suis pas intéressé de pratiquer plus d'activité physique
 - p) Je suis physiquement incapable de faire de l'activité physique
 - q) Autres
➤ Veuillez préciser _____
47. Si vous deviez utiliser un traqueur d'activité qui pourrait être spécialement conçu pour répondre à vos besoins, quel type d'informations souhaiteriez-vous connaître ? (Cochez toutes les réponses applicables)
- a) Nombre de pas

- b) Distance parcourue
 - c) Estimation des calories brûlées
 - d) Nombre d'étages montés
 - e) Votre niveau d'activité par rapport aux autres
 - f) Fréquence cardiaque
 - g) Temps passé assis
 - h) Temps passé debout
 - i) Temps passé dans différentes intensités d'activité (légère, modérée, vigoureuse)
 - j) Temps passé dans différents types d'activités (par exemple, entraînement, marche, yoga, etc.)
 - k) Habitudes de sommeil
 - l) Autres
 - Veuillez préciser _____
48. Si vous deviez utiliser un traqueur d'activité, seriez-vous disposée à ce que vos données soient partagées avec votre équipe soignante (par exemple, médecin, infirmières, assistant(e) médical(e)) ? **(une seule réponse possible)**
- a) Oui
 - b) Non
 - c) Peut-être/ je ne suis pas sûre
49. Souhaiteriez-vous participer à un programme d'activité physique réalisé à distance via une application téléphonique ? **(une seule réponse possible)**.
- a) Oui
 - b) Non
 - c) Peut-être/ je ne suis pas sûre
50. **(si b ou c à q49)** Pourquoi pourriez-vous ne pas être intéressée par ce type de programme réalisé à distance via une **application téléphonique** ? **(Cochez toutes les réponses applicables)**
- a) J'ai peur de ne pas savoir m'en servir
 - b) J'ai eu peu d'information à ce sujet
 - c) Je crains que mes données soient utilisées à mauvais escient
 - d) Je crains la dépendance à ces outils
 - e) Je n'ai pas eu l'occasion de m'en servir
 - f) Ces outils ne sont pas très conviviaux
 - g) Cela me prendra trop de temps à utiliser ces outils
 - h) Ces outils ne fonctionnent pas encore très bien
 - i) Je ne pourrais pas suivre un programme où je dois faire de l'activité physique par moi-même
 - j) Je fais déjà suffisamment d'activité physique par moi-même
 - k) Je ne suis pas intéressé de pratiquer plus d'activité physique
 - l) Je suis physiquement incapable de faire de l'activité physique
 - m) Autres
 - Veuillez préciser _____
51. Souhaiteriez-vous participer à un programme d'activité physique réalisé à distance via un **site web** ? **(une seule réponse possible)**.

- a) Oui
- b) Non
- c) Peut-être/ je ne suis pas sûre

52. (si b ou c à q51) Pourquoi pourriez-vous ne pas être intéressée par ce type de programme réalisé à distance via un site web ? (Cochez toutes les réponses applicables)

- a) J'ai peur de ne pas savoir m'en servir
- b) J'ai eu peu d'information à ce sujet
- c) Je crains que mes données soient utilisées à mauvais escient
- d) Je crains la dépendance à ces outils
- e) Je n'ai pas eu l'occasion de m'en servir
- f) Ces outils ne sont pas très conviviaux
- g) Cela me prendra trop de temps à utiliser ces outils
- h) Ces outils ne fonctionnent pas encore très bien
- i) Je ne pourrais pas suivre un programme où je dois faire de l'activité physique par moi-même
- j) Je fais déjà suffisamment d'activité physique par moi-même
- k) Je ne suis pas intéressé de pratiquer plus d'activité physique
- l) Je suis physiquement incapable de faire de l'activité physique
- m) Autres

➤ Veuillez préciser _____

53. Est-ce que la crise sanitaire du COVID-19 a eu un impact sur votre utilisation actuelle des technologies liées à l'activité physique ?

- a) Oui

➤ Veuillez préciser

- b) Non

DOMAINE 5 : QUESTIONNAIRE DES BARRIÈRES À LA PRATIQUE D'ACTIVITÉ PHYSIQUE

Vous êtes bientôt à la fin du questionnaire.

54. Vous trouverez ci-dessous une liste de raisons que les gens donnent pour expliquer pourquoi ils ne font pas autant d'activité physique qu'ils le voudraient. Veuillez lire chaque raison et indiquer à quelle fréquence cela peut s'appliquer à votre situation actuelle.

Les réponses vont de 1 à 5 et correspondent à : 1 = jamais, 2 = rarement, 3 = parfois, 4 = souvent, 5 = très souvent. Merci de répondre à toutes les questions, en ne cochant qu'une seule case par ligne.

Barrières : « Je ne pratique pas autant d'activité physique que je le voudrais car... »	Jamais	Rarement	Parfois	Souvent	Très souvent
Je suis trop fatiguée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
La mauvaise météo m'empêche de pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai des douleurs trop importantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je n'ai pas de partenaire avec qui pratiquer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Les horaires ne correspondent pas à mon emploi du temps	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je ne ressens pas de plaisir à pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Il n'existe pas de structure autour de chez moi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai peur du regard des autres	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cela ne m'intéresse pas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mes obligations familiales m'empêchent de pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je ne sais pas quoi faire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pratiquer une activité physique coûte trop cher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai peur de me blesser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mes engagements sociaux m'empêchent de pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je n'arrive pas à me motiver	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai des effets secondaires trop importants dus aux traitements d'une maladie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Il n'y a pas de transport accessible pour me rendre sur les lieux de pratique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je n'arrive pas à m' organiser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mes obligations au travail m'empêchent de pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je ne me sens pas capable de pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai peur de pratiquer à cause de l'insécurité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ce n'est pas ma priorité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

55. Existe-t-il d'autres éléments qui vous freinent **souvent ou très souvent** dans votre pratique d'activité physique ?

56. Est-ce que la crise sanitaire du COVID-19 a eu un impact sur vos barrières actuelles à la pratique d'activité physique ?

a) Oui

➤ Veuillez préciser

b) Non

DOMAINE 6 : QUESTIONNAIRE DES FACILITATEURS À LA PRATIQUE D'ACTIVITÉ PHYSIQUE

Voici la dernière partie du questionnaire qui concerne les potentiels facteurs facilitant la pratique de l'activité physique.

57. Vous trouverez ci-dessous une liste de raisons que les gens donnent pour expliquer ce qui facilite leur pratique d'activité physique. Veuillez lire chaque raison et indiquer à quelle fréquence cela peut s'appliquer à votre situation actuelle.

Les réponses vont de 1 à 5 et correspondent à : 1 = jamais, 2 = rarement, 3 = parfois, 4 = souvent, 5 = très souvent. Merci de répondre à toutes les questions, en ne cochant qu'une seule case par ligne.

Facilitateurs : « Ma pratique d'activité physique est facilitée par le fait que... »	Jamais	Rarement	Parfois	Souvent	Très souvent
Elle me permet d'avoir un meilleur <u>contrôle sur ma santé</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C'est une <u>recommandation de mon médecin</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet d'améliorer mes <u>capacités physiques</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai <u>du temps à consacrer</u> à ma pratique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet d'améliorer les <u>symptômes</u> de ma maladie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je me fixe un <u>but et un objectif</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet d'améliorer ma <u>qualité de vie</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cela me permet de <u>retrouver un groupe, un(e) ami(e), un proche</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet d'améliorer mon <u>sommeil</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je pratique dans une <u>structure qui adapte</u> l'activité physique à mes capacités	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de diminuer ma <u>fatigue</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis <u>encouragée par un groupe, un(e) ami(e), un proche</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de <u>maintenir ou perdre du poids</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je pratique dans un <u>environnement</u> où je me sens en sécurité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de me <u>sentir mieux dans mon corps</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis <u>encadrée par un professionnel de l'activité physique</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de me sentir <u>fière</u> de moi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de stimuler mon système <u>immunitaire</u> et d'éviter les maladies	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je ressens du <u>plaisir</u> à pratiquer une activité physique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je me sens <u>compétente</u> dans l'activité physique que je pratique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'arrive à organiser mon planning pour que l'activité physique devienne une <u>habitude</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elle me permet de retrouver de l' <u>énergie</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

58. Existe-t-il d'autres éléments qui facilitent **souvent ou très souvent** votre pratique d'activité physique :

59. Est-ce que la crise sanitaire du COVID-19 a eu un impact sur les facteurs actuels qui facilitent votre pratique d'activité physique ?

a) Oui

➤ Veuillez préciser

b) Non

B Démonstrations des formules mathématiques

Inertie entre modalité et barycentre

$$\begin{aligned} d^2(m, G_P) &= n \sum_{i=1}^n \left(\frac{h_{im}}{N_m} - \frac{1}{n} \right)^2 \\ &= n \sum_{i=1}^n \left(\frac{h_{im}}{N_m} - \frac{1}{n} \right)^2 \\ &= f_{.m} n \sum_{i=1}^n \left(\frac{h_{im}}{N_m^2} - \frac{2h_{im}}{nN_m} + \frac{1}{n^2} \right) \quad \text{car } \sum_{i=1}^n h_{im} = N_m \\ &= \frac{n}{N_m} - 1. \end{aligned}$$

Inertie d'une modalité

$$\begin{aligned} I(m) &= f_{.m} d^2(m, G_P) \\ &= f_{.m} \frac{n}{N_m} - 1 \\ &= \frac{N_m}{nP} \left(\frac{n}{N_m} - 1 \right) \\ &= \frac{1}{P} - \frac{N_m}{nP} \\ &= \frac{1}{P} \left(1 - \frac{N_m}{n} \right). \end{aligned}$$

Inertie d'une variable

$$\begin{aligned} I(p) &= \sum_{m=1}^{M_p} f_{.m} d^2(m, G_P) \quad \text{où } M_p \text{ est le nombre de modalités de la variable } p \\ &= \sum_{m=1}^{M_p} \frac{1}{P} \left(1 - \frac{N_m}{n} \right) \\ &= \frac{1}{P} \left(M_p - \frac{1}{n} \sum_{m=1}^{M_p} N_m \right) \\ &= \frac{1}{P} (M_p - 1) \quad \text{car } \sum_{m=1}^{M_p} N_m = n. \end{aligned}$$

C Graphiques et tableaux annexes

Caractéristiques des participants

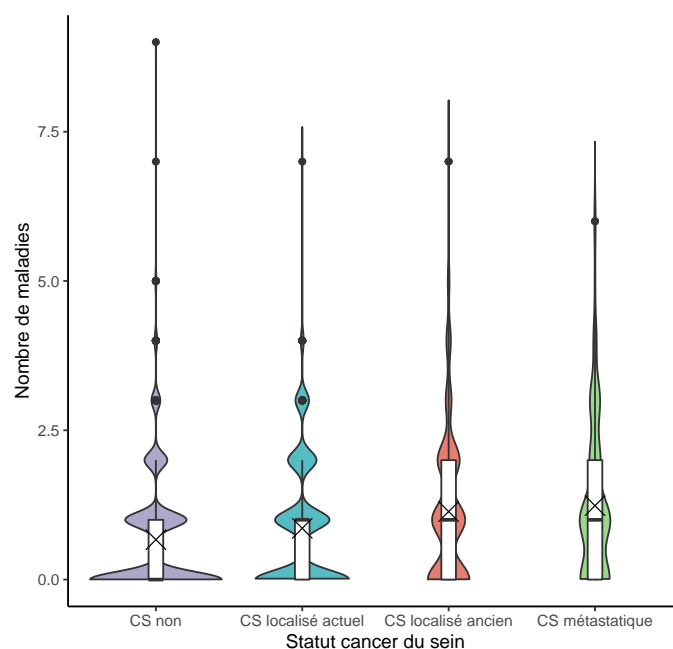


FIGURE 2 – Violin plot du nombre de maladie en fonction du statut du cancer.

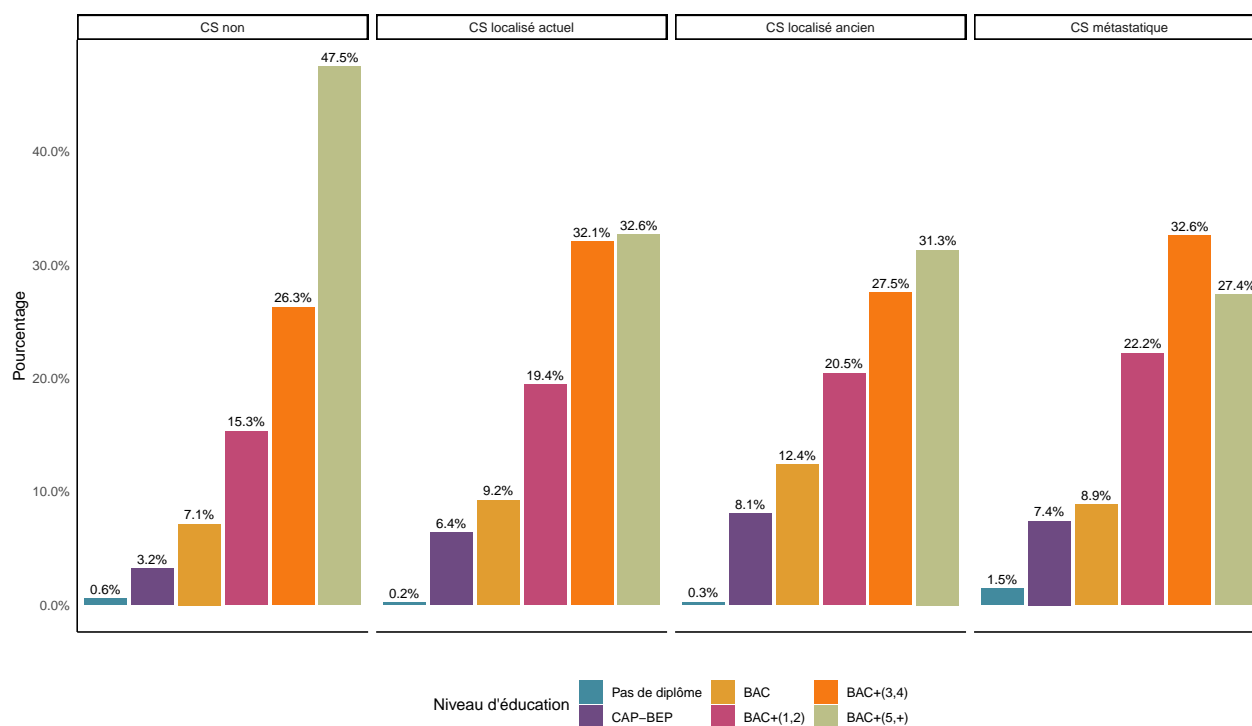


FIGURE 3 – Niveau d'éducation des participantes en fonction du statut du cancer.

Préférences en activité physique

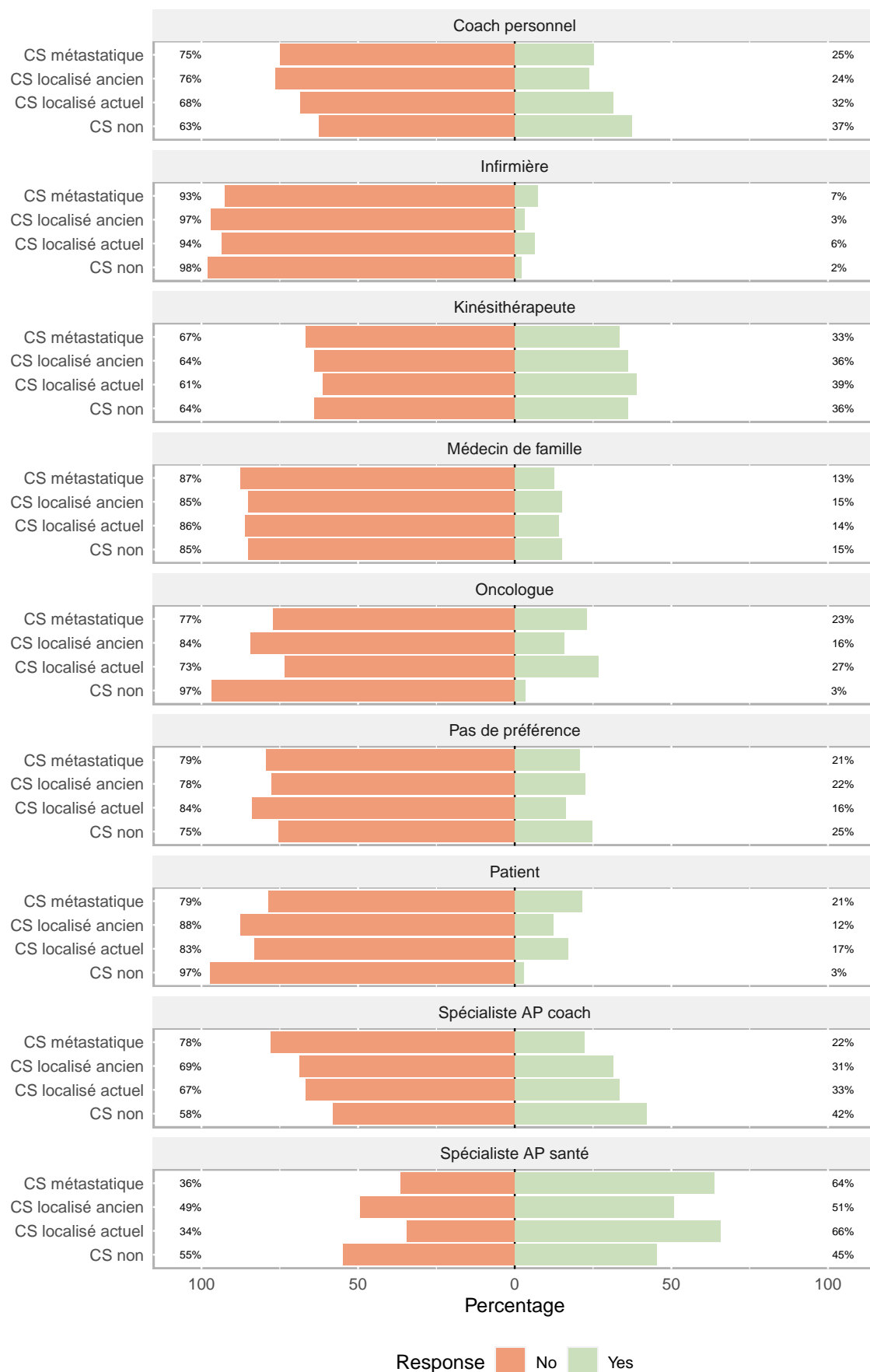


FIGURE 4 – De qui les participantes souhaitent obtenir des informations sur la pratique d'une activité physique (en %).

	CS non	CS localisé actuel	CS localisé ancien	CS métastatique
Où pratiquer ?				
A l'extérieur	1008 (81.0%)	418 (78.9%)	282 (71.2%)	108 (80.0%)
A la maison	611 (49.1%)	268 (50.6%)	169 (42.7%)	62 (45.9%)
Au travail	164 (13.2%)	32 (6.0%)	16 (4.0%)	5 (3.7%)
Dans une salle	621 (49.9%)	221 (41.7%)	177 (44.7%)	57 (42.2%)
Dans un hôpital	310 (24.9%)	215 (40.6%)	136 (34.3%)	58 (43.0%)
Pas de préférence	97 (7.8%)	62 (11.7%)	44 (11.1%)	12 (8.9%)
Comment pratiquer ?				
Seule	513 (41.2%)	223 (42.1%)	144 (36.4%)	42 (31.1%)
En binôme	563 (45.2%)	208 (39.2%)	147 (37.1%)	45 (33.3%)
Avec des patients	192 (15.4%)	196 (37.0%)	113 (28.5%)	59 (43.7%)
Avec des non-malades	395 (31.7%)	207 (39.1%)	167 (42.2%)	51 (37.8%)
Pas de préférence	351 (28.2%)	115 (21.7%)	94 (23.7%)	32 (23.7%)
Intensité des séances				
Faible	54 (4.3%)	26 (4.9%)	23 (5.8%)	8 (5.9%)
Modérée	849 (68.2%)	399 (75.3%)	304 (76.8%)	100 (74.1%)
Élevée	286 (23.0%)	88 (16.6%)	53 (13.4%)	22 (16.3%)
Pas de préférence	56 (4.5%)	17 (3.2%)	16 (4.0%)	5 (3.7%)
Organisation des séances				
Flexible	560 (45.0%)	223 (42.1%)	161 (40.7%)	50 (37.0%)
Planifiée	558 (44.8%)	242 (45.7%)	169 (42.7%)	67 (49.6%)
Pas de préférence	127 (10.2%)	65 (12.3%)	66 (16.7%)	18 (13.3%)
Quand pratiquer				
Matin	348 (28.0%)	203 (38.3%)	161 (40.7%)	60 (44.4%)
Midi	120 (9.6%)	46 (8.7%)	13 (3.3%)	7 (5.2%)
Après-midi	141 (11.3%)	72 (13.6%)	69 (17.4%)	23 (17.0%)
Soirée	403 (32.4%)	123 (23.2%)	89 (22.5%)	24 (17.8%)
Pas de préférence	233 (18.7%)	86 (16.2%)	64 (16.2%)	21 (15.6%)

TABLE 1 – Préférences des participantes sur la pratique d'une activité physique.

Caractéristiques des classes

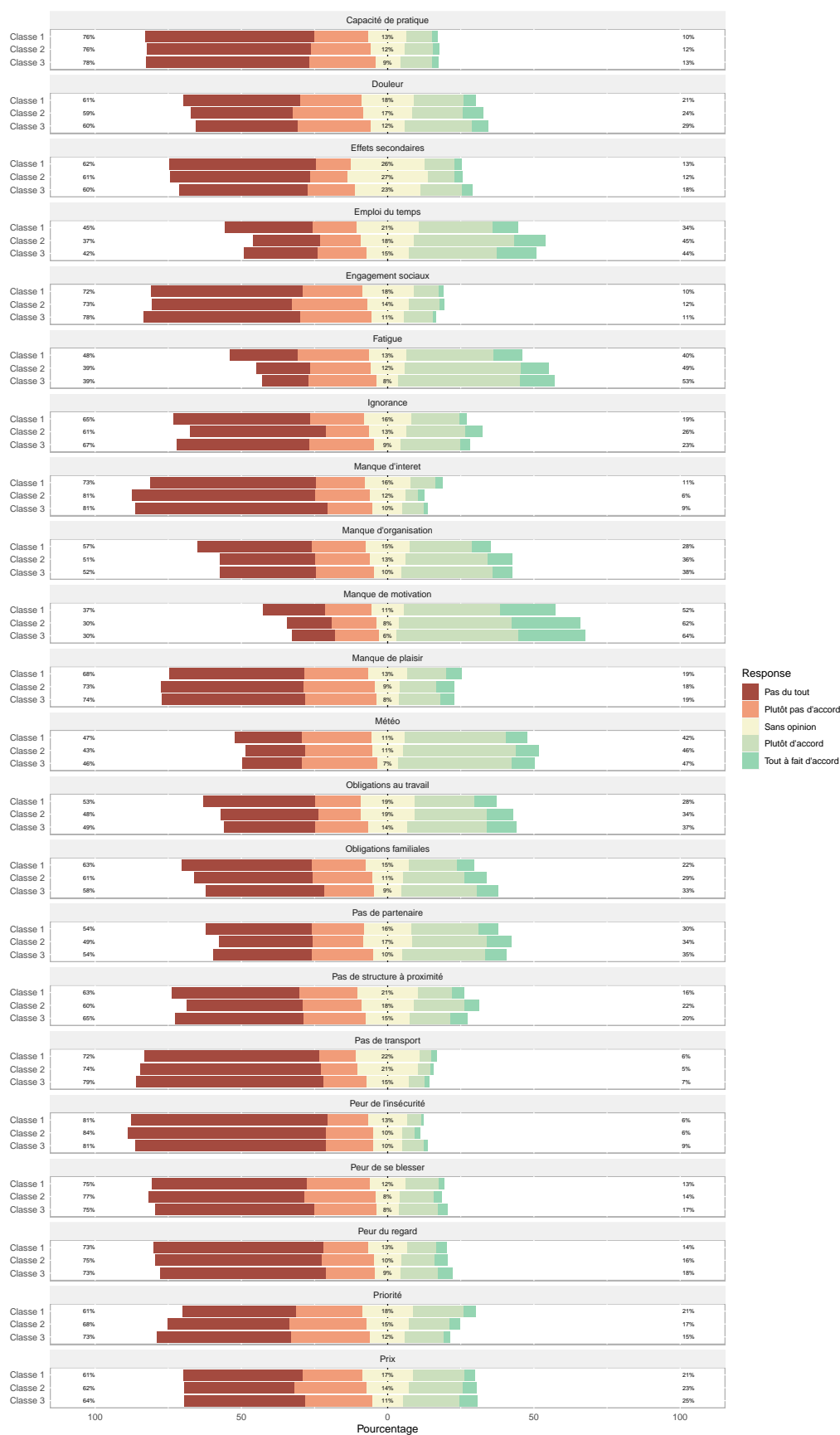


FIGURE 5 – Barrières à l'activité physique en fonctions des classes issues du partitionnement.

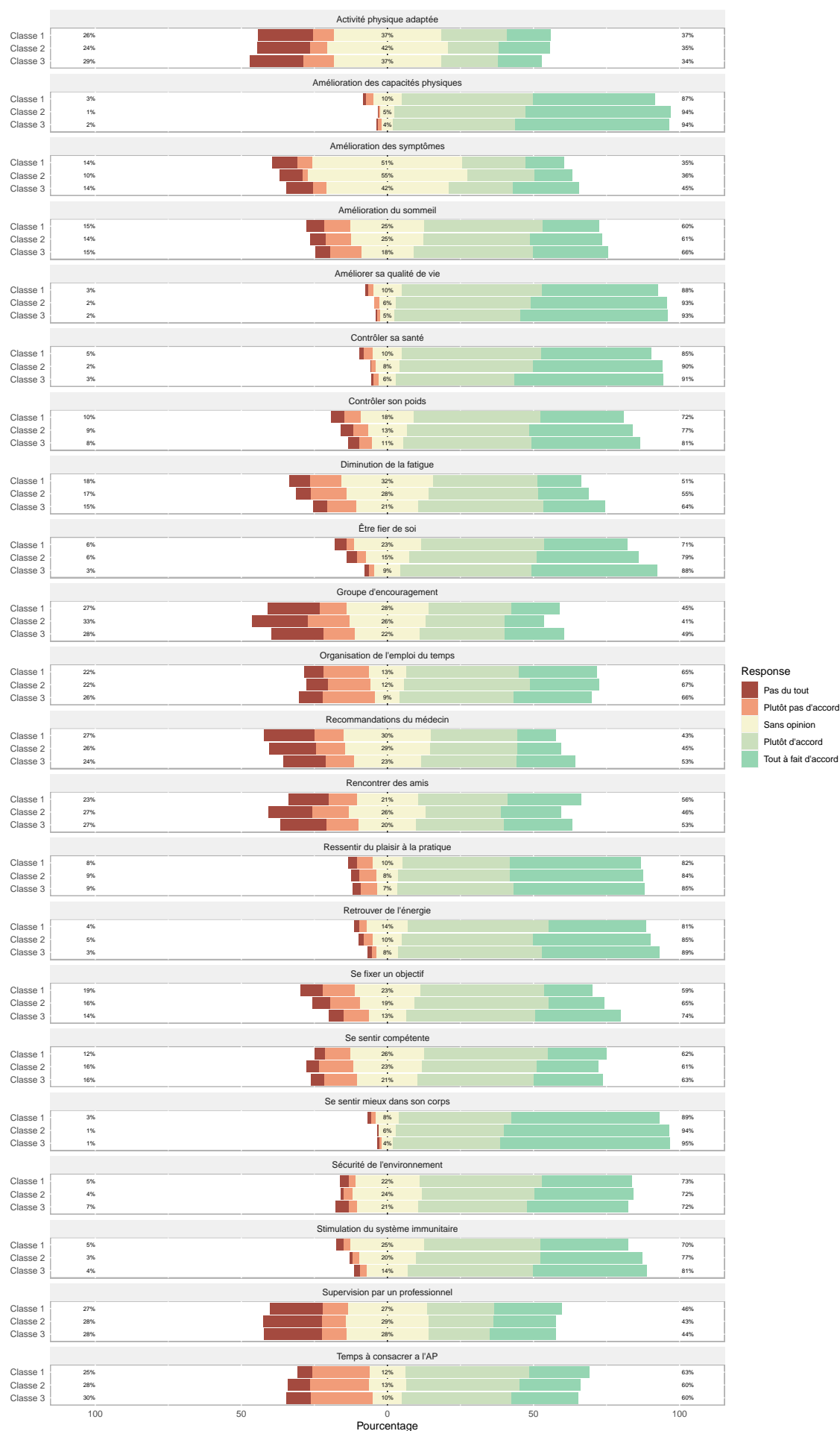


FIGURE 6 – Facilitateurs à l'activité physique en fonctions des classes issues du partitionnement.

	Classe 1	Classe 2	Classe 3
Utilisation des réseaux			
Facebook	58.6%	65.0%	78.2%
Instagram	32.6%	44.8%	53.4%
Linkedin	24.7%	33.6%	39.0%
Twitter	7.8%	9.0%	14.0%
Aucun	31.0%	20.9%	10.3%
Utilisation des traqueurs			
Apple	1.9%	0.9%	5.1%
Fitbit	4.4%	4.2%	10.0%
Garmin	4.4%	4.2%	8.3%
Podomètre téléphone	26.8%	31.3%	41.8%
Aucun	55.1%	54.3%	25.6%
Programme AP			
Via un traqueur	0.9%	0%	100%
Via une application	0.7%	60.6%	80.3%
Via une site web	0.2%	74.5%	61.5%

TABLE 2 – Utilisation des technologies (% de participantes ayant répondu oui).

D Indices de pénibilités de Capgemini

Description des indices

le volet s'affiche lors d'un clic ou d'un passage de souris sur les catégories (cf page 1)





 Accessibilité aux soins	 Addiction	 Economie	 Etat de santé
<p>Cet indice doit être interprété comme la facilité pour un patient d'accéder à des soins. Plus la valeur de cet indice est élevée, plus il sera difficile pour ce patient d'accéder à des soins.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Temps d'accès moyen à un service d'urgence (SU) ou à un service mobile d'urgence et de réanimation (SMUR); - Temps d'accès moyen à un service d'urgence (SU), à un service mobile d'urgence et de réanimation (SMUR) ou à un médecin correspondant du SAMU (MCS); - Temps d'accès moyen à HéliSMUR; - Temps d'accès moyen à un HéliSC (hélicoptère de la sécurité civile); - APL au médecin généralistes (nombre moyen de consultation de médecine générale auquel ont accès les habitants d'une commune, compte tenu de la structure par âge de la population); - Score d'accessibilité aux soins de premiers recours (SCALE index). <p>Cet indice est donné à l'échelle communale.</p>	<p>Cet indice doit être interprété comme le niveau d'addiction d'un territoire donné. Les addictions concernent le tabac, l'alcool et les drogues « dures ». Plus la valeur de cet indice est élevée, plus il le territoire dans lequel le patient vit sera sujet à l'addiction.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Proportion de fumeurs quotidiens; - Ventes de Stéribox® (trousse de réduction des risques infectieux liés à l'injection intraveineuse de drogue); - Consultations en CSAPA pour addiction à l'alcool (Centres de Soins, d'Accompagnement et de Prévention en Addictologie); - Consultations en CSAPA pour addiction sans substance; - Consultations en CSAPA pour addiction aux opiacés et stimulants. <p>Cet indice est donné à l'échelle départementale.</p>	<p>Cet indice doit être interprété comme une description de la situation économique du lieu de vie d'un patient. Plus la valeur de cet indice est élevée, Cet indice s'inspire de l'indice européen FDep (french deprivation index).</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Pourcentage de bacheliers dans les plus de 15 ans; - Revenu médian par unité de consommation dans le ménage; - Taux de chômeurs; - Pourcentage d'ouvriers dans la population active. <p>Cet indice est donné à l'échelle de l'iris.</p>	<p>Cet indice doit être interprété comme une description de l'état de santé moyen de la population du lieu de vie du patient. Plus la valeur de cet indice est élevée, plus la population vivant dans son territoire est en mauvaise santé.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Taux de mortalité standardisé des 0 à 64 ans; - Taux de suicide; - Pourcentage de la population admis en ALD (affection longue durée, liste des pathologies incluses en suivant ce lien); - Espérance de vie à la naissance; - Espérance de vie à 65 ans. <p>Cet indice est donné à l'échelle départementale.</p>

FIGURE 7 – Descriptions des indices de pénibilités de Capgemini, 1ère partie.

Description des indices

le volet s'affiche lors d'un clic ou d'un passage de souris sur les catégories (cf page 1)




 Matériel	 Sécurité	 Socio démographie
<p>Cet indice doit être interprété comme la description de la situation matérielle (type et confort des logements, équipement etc.) du lieu de vie d'un patient. Plus la valeur de cet indice est élevée, plus la situation d'un point de vue matériel du territoire de vie d'un patient sera mauvaise.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Taux de ménages non propriétaires; - Taux de ménages sans salle de bain privative; - Part de logement en situation de sur occupation; - Taux de ménages sans voiture; - Nombre de résidences principales de type appartement; - Nombre de résidences principales construites entre 1946 et 1990; - Nombre de logements vacants; - Nombre de résidences principales de 2 pièces au maximum; - Nombre de commerces de proximité; - Nombre de d'équipements de sport et de loisir; - Nombre d'équipements pour l'enseignement. <p>Cet indice est donné à l'échelle de l'iris.</p>	<p>Cet indice doit être interprété comme la la description de la situation sécuritaire du lieu de vie d'un patient. Plus la valeur de cet indice est importante, plus la situation sécuritaire de son lieu de vie est mauvaise.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Taux de cambriolages de logement; - Taux de coups et blessures volontaires; - Taux de dégradations et destructions; - Taux de vols d'accessoires sur véhicules; - Taux de vols sans violence sur des personnes; - Taux de vols violents sans arme; - Taux de vols dans les véhicules. <p>Cet indice est donné à l'échelle départementale.</p>	<p>Cet indice doit être interprété la description de la situation socio démographique du lieu de vie d'un patient. Plus la valeur de cet indice est élevée, plus la situation est défavorable.</p> <p>Les variables qui composent cet indice sont les suivantes:</p> <ul style="list-style-type: none"> - Taux de chômeurs; - Taux de familles monoparentales; - Densité de population; - Taux de personnes de 15 ou plus vivant seules; - Taux de famille avec 4 enfants ou plus de moins de 25 ans; - Taux de personnes de plus de 15 ans non scolarisées titulaire d'au plus un BEPC (brevet d'études du premier cycle) ou DNB (diplôme national du brevet); - Taux de personnes appartenant à la catégorie socio professionnelle 5 (employés et personnels de service) ou à la catégorie socio professionnelle 6 (ouvriers qualifiés); - Taux de personne appartenant à la catégorie socio professionnelle 8: autres personnes sans activité professionnelle. <p>Cet indice est donné à l'échelle départementale.</p>

FIGURE 8 – Descriptions des indices de pénibilités de Capgemini, 2ème partie.