

Définition des trajectoires de soins pour les personnes ayant déclaré un trouble de l'humeur

– Rapport méthodologique –

Matea Belan, Ph.D.
Stéphanie Fontaine, M.Sc., patiente-partenaire
Christian Kouakou, Ph.D.
Maude Laberge, Ph.D.
Thomas Poder, Ph.D.

Octobre 2022

Auteurs de correspondance :

Maude Laberge, Ph.D
Université Laval
Département d'opérations et systèmes
de décision
maude.laberge@fsa.ulaval.ca

Thomas G Poder, Ph.D.
Université de Montréal
École de santé publique – Département de
gestion, d'évaluation et de politique de santé
thomas.poder@umontreal.ca

PARTENAIRES PRINCIPAUX

Instituts de recherche en santé du Canada
Ministère de la Santé et des Services sociaux
Fonds de recherche du Québec – Santé
RUISSS Université McGill
RUISSS Université de Montréal
RUISSS Université de Sherbrooke
RUISSS Université Laval

Université Laval
Université McGill
Université de Montréal
Université de Sherbrooke

Stratégie de recherche axée sur le patient
SRAP
Le patient d'abord



IRSC CIHR
Instituts de recherche en santé du Canada Canadian Institutes of Health Research

Table des matières

1.	<i>Introduction</i>	4
2.	<i>Contexte</i>	4
3.	<i>Méthodologie</i>	5
3.1.	Détermination de la cohorte de patients d'intérêt	5
3.2.	Détermination de la période de suivi	6
3.3.	Variables d'intérêt utilisées	7
3.3.1.	Analyse de séquences	7
3.3.2.	Analyse de classes et de profils latents	8
3.4.	Détermination des trajectoires	8
3.4.1.	Analyse de séquences	8
3.4.2.	Analyse de classes et profils latents	8
3.5.	Constitution et traitement de la base de données	9
3.5.1.	Analyse de séquences	9
3.5.2.	Analyse de classes et profils latents	11
3.6.	Analyse des données	12
3.6.1.	Analyse de séquences	12
3.6.2.	Analyse de classes et profils latents	14
3.6.3.	Analyse des coûts	15
4.	<i>Résultats</i>	16
4.1.	Description de l'échantillon	16
4.2.	Résultats obtenus par l'analyse de séquences	20
4.2.1.	Données descriptives des séquences	20
4.2.2.	Résultats de l'analyse par grappes	22
4.2.3.	Résultats des coûts à partir de l'analyse de séquences	27
4.3.	Résultats obtenus par l'analyse de classes et profils latents	27
4.3.1.	ACL : Tous les services médicaux	27
4.3.2.	APL : Services médicaux	29
4.3.3.	APL : Hospitalisations	34
5.	<i>Interprétation et utilisation des méthodes</i>	38
5.1.	Méthode de l'analyse de séquences	38
5.2.	Méthode de l'analyse de classes et de profils latents	39
5.3.	Comparaison des trois méthodes d'analyse	40
5.3.1.	Analyse de séquences vs Analyse de classes latentes et de profils latents	40
5.3.2.	Analyse de classes latentes vs Analyse de profils latents	41
6.	<i>Discussion</i>	41
7.	<i>Conclusion</i>	42
8.	<i>Fiches techniques des méthodes</i>	42

9. Algorithmes utilisés	43
9.1. Traitement des données et sélection des individus.....	43
9.2. Algorithme de l'analyse de séquences et classes latentes	44
Références	44
Annexes 46	
Annexe 1. Bases de données, variables et transformations utilisées pour analyse de séquences	46
Annexe 2. Bases de données, variables et transformations utilisées pour analyse de classes latentes.....	50
Annexe 3. Coûts totaux moyens par individu pour les analyses de séquences, de classes latentes et de profils latents.	51

Abréviations

AAO : Analyse d'appariement optimal

ACL : Analyse de classes latentes

AIC : *Akaike information criteria*

AL : Analyse latente

APL : Analyse de profils latents

AS : Analyse de séquences

BIC : *Bayesian information criteria*

ESCC : Enquête sur la santé dans les collectivités canadiennes

IC : Intervalle de confiance

IIQ : Intervalle interquartile

IMC : Indice de masse corporelle

IPC : Indice de prix à la consommation

NIRRU : Niveau d'intensité relative des ressources utilisées

ONU : Organisation des Nations Unies

RAMQ : Régie de l'assurance maladie du Québec

s.e. : Erreur standard

SSA : Système de santé apprenant

TS : Trajectoire de soins

Glossaire

Analyse de classes latentes	L'analyse de classes latentes est une technique d'analyse statistique qui consiste en une segmentation de l'échantillon observé à partir des données discrètes recueillies. Le but de la méthode est de mettre en évidence des classes ou sous-groupes essentiellement homogènes.
Analyse de profils latents	L'analyse de profils latents est une technique d'analyse statistique qui consiste en une segmentation de l'échantillon observé à partir des données continues recueillies. Le but de la méthode est de mettre en évidence des classes ou sous-groupes essentiellement homogènes. La différence fondamentale entre ces deux méthodes repose sur la nature discrète ou continue des données.
Analyse de séquences	L'analyse de séquences est une technique d'analyse statistique qui permet de décrire les trajectoires des individus à travers une suite ou succession d'états ou séquences au fil du temps. Le but de la méthode est de mettre en évidence et regrouper les individus ayant des états essentiellement similaires.
Troubles de l'humeur	Les troubles de l'humeur sont des troubles mentaux essentiellement caractérisés par une perturbation de l'humeur. Les personnes souffrant de troubles de l'humeur peuvent connaître des périodes relativement longues et intenses d'émotions négatives. Ces troubles incluent la dépression, le trouble bipolaire, la manie ou la dysthymie.
Trajectoire de soins	La trajectoire de soins est la série de soins de santé que peuvent recevoir des personnes souffrant d'une condition médicale quelconque dans une période définie.

1. Introduction

Depuis quelques années, l'évaluation des trajectoires de soins (TS) pour les maladies chroniques gagne en importance. Le concept de « trajectoire de soins » représente la séquence individuelle de l'utilisation des soins de santé sur une période de temps définie¹. L'évaluation des TS permet d'identifier celles qui sont les plus efficaces, c'est-à-dire celles qui minimisent l'utilisation potentiellement inappropriée des services et des soins de santé, tout en maximisant les résultats sur la santé des individus. Ceci relève d'une grande importance pour les décideurs dans la conception et la réorganisation des soins de santé et des systèmes de santé, mais également pour les patients dans l'amélioration de leur expérience au sein du système de santé.

La santé mentale et l'amélioration des soins et services rendus à la population souffrant de maladies mentales, est un enjeu sociétal de plus en plus important et souvent négligé². Pourtant, sa prévalence est alarmante partout à travers le monde³. Au Canada, plus spécifiquement, selon l'Association canadienne pour la santé mentale, dans une année donnée, une personne sur cinq présente une problématique de santé mentale et dès l'âge de 40 ans, près de 50 % des Canadiens auront eu ou auront une maladie mentale⁴. Les coûts totaux directs des maladies mentales, incluant les coûts relatifs aux soins médicaux et soins hospitaliers, aux médicaments et aux services communautaires et sociaux, étaient estimés au Canada à 42,3 milliards de dollars en 2011⁵.

Parmi les maladies mentales, les troubles de l'humeur (TH) représentent une catégorie de maladies caractérisées par des perturbations émotives pouvant se manifester par des sentiments envahissants et durables de dépression, de manie, ou des deux. Deux maladies principales composent les TH, soit la dépression majeure, qui affecte environ 5,4 % de la population canadienne, et le trouble bipolaire qui touche près de 1 % de la population⁴.

L'optimisation des soins de santé pour les problématiques de santé mentale est devenue un enjeu prioritaire, non seulement pour les gouvernements provinciaux et fédéral au Canada, mais également au niveau international⁶. En ce sens, l'évaluation de TS typiques des personnes atteintes de maladies mentales représente une première étape en ce sens. Elle permet d'observer les contacts des usagers avec le système de santé et leur utilisation des services de soins. Pour ce faire, nous avons mis en place cette étude de cohorte dont l'objectif est de décrire et d'identifier les TS typiques des personnes atteintes de TH, dans une population québécoise représentative.

2. Contexte

Dans le cadre de la transition vers la phase 2 de l'Unité de soutien au système de santé apprenant (SSA) du Québec, deux projets en lien avec les TH ont été proposés. L'objectif de ces projets est de définir les principales TS des personnes souffrant de TH et de déterminer les coûts de ces trajectoires. L'évaluation et l'analyse des trajectoires dans le domaine de la santé est relativement récent et a été rendu possible particulièrement avec le développement de technologies informatiques puissantes, permettant des calculs complexes. De ce fait, encore peu de techniques méthodologiques et statistiques existent pour faire l'analyse des TS et déceler les TS typiques. À

notre connaissance, deux groupes de méthodes ont particulièrement été utilisées dans le domaine : l'analyse de séquences (AS) et l'analyse latente (AL) (c.-à-d. analyse de classes latentes et analyse de profils latents (ACL et APL)). Notre étude a pour but d'appliquer ces techniques pour l'identification des TS typiques pour les personnes vivant avec un TH et de comparer les résultats issus de ces méthodes entre eux.

Les données utilisées proviennent de la cohorte TorSaDE (Trajectoires Santé – Données Enrichies). Cette cohorte a été constituée en faisant un jumelage entre les données du Québec provenant des Enquêtes sur la santé dans les collectivités canadiennes (ESCC) et les données administratives de la Régie de l'assurance maladie du Québec (RAMQ), sur une période de 20 ans (1996 à 2016). Les données relatives à l'utilisation des services en santé disponibles concernent seulement les hospitalisations, les consultations en chirurgie d'un jour, à l'urgence, et avec les médecins omnipraticiens et spécialistes, ainsi que les prescriptions de médicaments couvertes par la RAMQ. De ce fait, les informations relatives à la consultation d'autres professionnels de la santé, tels que psychologue ou travailleur social, ou auprès des services communautaires ne sont pas disponibles. La définition des TS demeure tout de même utile en l'absence de ces sources de soins, puisqu'elle permet de voir le fardeau de la condition sur le système public de soins et services de santé. De telles informations peuvent être utiles aux décideurs pour la réorganisation et l'optimisation des soins et services.

Ce rapport présente un aperçu des différents aspects conceptuels, méthodologiques et analytiques du projet sur la définition des TS des personnes ayant déclaré être atteintes de TH dans l'ESCC.

3. Méthodologie

3.1. Détermination de la cohorte de patients d'intérêt

La cohorte initiale est composée des personnes ayant complété au moins une fois l'ESCC entre 2007 et 2016 et ayant donné leur accord pour le jumelage de leurs données de l'ESCC avec leurs données de la RAMQ.

L'identification des personnes atteintes de TH reposent sur deux possibles approches, soit : 1) l'approche de déclaration par la personne concernée en répondant à la question suivante de l'ESCC : « Rappelez-vous, nous sommes intéressés aux problèmes de santé diagnostiqués par un professionnel de la santé. Êtes-vous atteint.e d'un trouble de l'humeur tel que la dépression, le trouble bipolaire, la manie ou la dysthymie? » ; ou 2) l'approche du diagnostic obtenu par une tierce partie, c'est-à-dire un médecin, pouvant être utilisée par le biais des diagnostics émis lors des visites médicales disponibles dans la base de données de la RAMQ. Notre choix d'identification a été celui de l'auto-déclaration par la personne concernée, car elle correspond à la situation vécue par le patient au moment de la complétion de l'ESCC et permet l'utilisation des différentes données fournies par les participants lors de l'ESCC (telles qu'utilisées dans la section des résultats). L'utilisation des diagnostics disponibles par la RAMQ a été écartée, car les experts médicaux consultés estiment que la qualité de la donnée pourrait être sujette à un risque d'erreur pour l'identification des personnes avec un TH. Ceci s'explique par le fait que, dans la base de

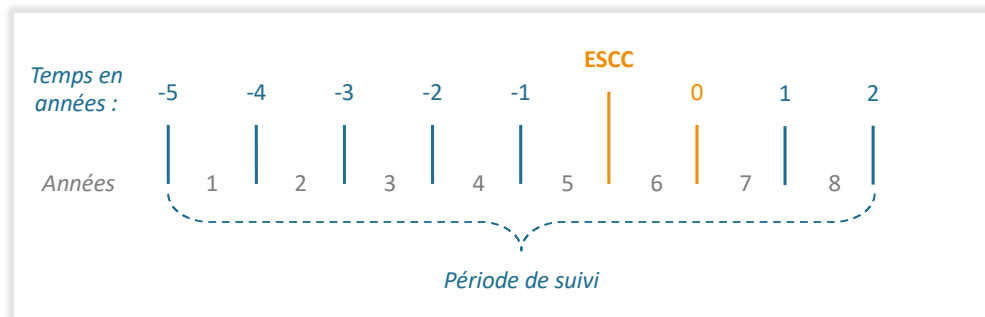
données de la RAMQ, chaque visite médicale est associée à un seul diagnostic. Cependant, il est fort possible qu'un individu consulte pour plus d'une problématique. Par exemple, un individu avec un TH pourrait aussi consulter pour son diabète. Dans ce cas, le diagnostic associé à la visite médicale serait la problématique de diabète et non le TH. Ceci pourrait faire en sorte que nous sous-estimions l'échantillon de personnes avec un TH dans certaines situations. Nous pourrions aussi le surestimer dans d'autres situations, comme dans le cas où une personne consulterait pour une dépression mineure et passagère. Par ailleurs, les risques d'erreurs de saisies des diagnostics sont également possibles.

3.2. Détermination de la période de suivi

Pour l'AS et l'ACL, la période de suivi des patients couvre 7 années, soit les 5 années antérieures à la date de la déclaration du TH à l'ESCC (nommée « date index ») et les 2 années après cette date. Par exemple, pour un individu ayant déclaré avoir un TH à son ESCC le 21 octobre 2009, la période de suivi sera du 21 octobre 2004 (5 années avant la date index) au 21 octobre 2011 (2 années après la date index). Ce choix se justifie par le fait que nous disposons des données de la RAMQ jusqu'à 2016, soit 2 ans après la date de la dernière série d'ESCC utilisée, ainsi que par la volonté de suivre chaque individu sur une période relativement longue afin d'analyser leur TS, soit 5 ans avant la date index.

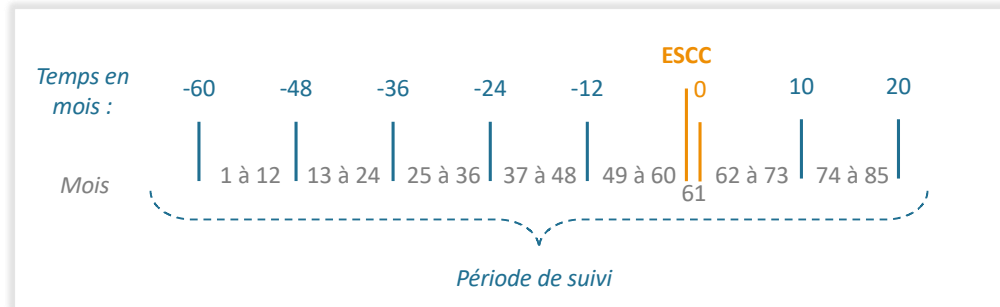
Pour l'APL, nous avons considéré l'année de déclaration du TH à l'ESCC comme date index. Ainsi, pour un individu ayant déclaré avoir un TH à son ESCC en 2009 (date index), nous avons considéré les années de 2004 à 2008 (5 années avant), l'année 2009, puis les années 2010 et 2011 (2 années après), soit 8 années en tout. Puisque pour l'APL, la période considérée est l'année, nous avons 8 années de suivi en incluant l'année index (voir figure 1).

Figure 1. Période de suivi pour la définition des trajectoires de soins pour l'APL.



Dans le cas de l'AS et l'ACL, la période étant mensuelle, la période totale de suivi est de 85 mois, soit 7 années et 1 mois; le mois supplémentaire étant compté pour le mois de la date index (voir figure 2).

Figure 2. Période de suivi pour la définition des trajectoires de soins pour l'AS et l'ACL.



3.3. Variables d'intérêt utilisées

3.3.1. Analyse de séquences

Avant d'effectuer la construction de la base de données finale servant à effectuer l'AS, nous avons ciblé les différentes bases de données et variables nous permettant d'avoir les informations nécessaires. Le tableau A1, à l'annexe 1, présente un résumé des bases de données, des variables et des transformations effectuées sur celles-ci pour la construction de la base de données finale pour l'AS.

Les variables d'intérêts pour notre AS sont les hospitalisations, les consultations à l'urgence, et les consultations médicales ambulatoires avec un médecin spécialiste ou avec un omnipraticien. Nous avons ainsi défini cinq états, selon l'ordre de priorités suivant :

1. Hospitalisation ou consultation à l'urgence
2. Consultation ambulatoire avec un médecin psychiatre
3. Consultation ambulatoire avec un médecin spécialiste autre qu'en psychiatrie
4. Consultation ambulatoire avec un omnipraticien
5. Aucune consultation

Puisque les hospitalisations étaient peu fréquentes, soit en moyenne $5,38 \pm 7,60$ hospitalisations par individu sur la période de 7 années de suivis autour de la date index, nous avons décidé de combiner les hospitalisations et les consultations à l'urgence. En effet, puisque qu'en moyenne les individus avait une hospitalisation par année et demie de suivi, cela aurait engendré plusieurs mois avec aucune occurrence, ce qui peut nuire à l'efficacité de l'AS. Une proportion de 81,2 % des individus avait eu au moins une hospitalisation, et 95,4 % avaient eu au moins une consultation à l'urgence.

3.3.2. Analyse de classes et de profils latents

Concernant l'ACL, nous avons dichotomisé chacune des variables d'utilisation des services médicaux et construit un ensemble de 5 variables d'intérêt, identique à celui de l'AS : hospitalisation et visite à l'urgence, consultation avec un psychiatre, consultation avec un médecin spécialiste (autre que psychiatre), consultation avec un médecin omnipraticien et aucun service. Cette dernière catégorie (« aucun service ») a été ajoutée pour prendre en compte les personnes n'ayant utilisé aucun service au cours de la période de suivi.

Pour l'APL, nous avons principalement utilisé 2 variables d'intérêt: le nombre de services médicaux utilisés et le nombre d'hospitalisations (ceci exclu les visites aux urgences). Par la suite, plusieurs autres variables socio-économiques ont été utilisées comme co-variables dans le modèle.

3.4. Détermination des trajectoires

Dans cette section, nous présentons la détermination des TS selon le type d'analyse, soit l'AS et l'AL.

3.4.1. Analyse de séquences

Dans le cadre de l'AS, la définition des TS typiques se fait au moyen de l'analyse par grappes (*cluster analysis* en anglais), qui nous permet de regrouper des trajectoires jugées similaires. La similarité des trajectoires est évaluée à partir de distances calculées entre chacune des séquences individuelles au moyen de l'analyse d'appariement optimal (AAO ; *Optimal Matching Analysis* en anglais). De ce fait, les séquences sont regroupées entre elles de manière à maximiser la similarité entre les séquences qui composent la même grappe (homogénéité intra-grappe), ainsi que la différence entre les grappes de séquences elles-mêmes (hétérogénéité inter-grappe).

3.4.2. Analyse de classes et profils latents

Pour l'ACL, nous avons utilisé un modèle adapté aux données dichotomiques et polytomiques. Aucune variable explicative n'a été incluse dans l'analyse. Après l'estimation du modèle, nous avons déterminé le nombre optimal de classes à partir de la comparaison des statistiques de tests associées à chaque modèle estimé. Principalement, nous avons retenus les critères d'information *Akaike* (AIC) et Bayésien (BIC), ainsi que de la logique analytique des données.

Concernant l'APL, considérant la nature longitudinale des données, un modèle de la famille des combinaisons de croissance (GMM) a été utilisé afin de pouvoir tenir compte des potentielles hétérogénéités intra-groupes. Le nombre de profils dans l'analyse a été déterminé de façon identique à l'ACL, c'est-à-dire en comparant les statistiques de tests associées à chaque modèle de régression, principalement les critères comme ceux utilisés pour l'ACL. Plusieurs variables d'intérêt ont été considérées, en particulier les services médicaux utilisés et le nombre d'hospitalisations. Les caractéristiques socio-économiques des participants ont également été incluses dans l'analyse.

3.5. Constitution et traitement de la base de données

3.5.1. Analyse de séquences

Pour effectuer l'AS, il est nécessaire de construire une base de données constituée d'une ligne par individu et d'une colonne pour chaque fréquence d'observation des états d'intérêt. Nous avons construit cette base de données finale en cinq étapes expliquées ci-dessous.

La première étape consiste à identifier des états ou des événements que nous voulons observer durant la période de suivi, et à établir un ordre de priorité pour ceux-ci (voir section 2.3.1.).

La seconde étape comprend la création d'une base de données contenant les dates de début et de fin pour chacun des événements visés, que nous pouvons nommer « Fichier des services reçus ». Le tableau 1 présente un exemple de ce type de fichier.

Tableau 1. Exemple de fichier des services reçus			
id	type_service	date_debut	date_fin ^α
1	Hospitalisation	1998-01-27	1998-02-04
1	Consultation avec médecin omnipraticien	2001-07-08	2001-07-08
2	Urgence	1999-09-20	1999-09-20
...			
^α À noter que pour les services dont la durée est de moins d'une journée, la date de fin sera la même que la date de début.			

La troisième étape comprend la création d'une base de données contenant les dates de la période de suivi. Il faut ici avoir une ligne par individu et par période d'observation des états.

Nous avons opté pour une fréquence mensuelle d'observation des états, ceci afin d'assurer une diversité dans les états répertoriés, surtout en tenant compte de l'étendue élevée de la période de suivi (7 années). Une fréquence quotidienne ou hebdomadaire aurait certainement représenté trop d'états où les individus n'auraient eu aucun contact avec le système de santé. La démarche d'analyse et les résultats de l'AS n'auraient pas été optimaux dans ce contexte.

Pour faciliter les analyses, nous avons tenu compte des mois du calendrier, plutôt que des durées mensuelles (p. ex. 30 jours = un mois). Ainsi, un individu ayant complété l'ESCC en date du 12 novembre 2013 aura une période de suivi allant du 12 novembre 2008 au 12 novembre 2015. Puisque nous avons tenu compte du mois du calendrier pour cet individu, nous avons un suivi allant du mois de novembre 2008 au mois de novembre 2015, soit 85 mois au total. De ce fait, notre base de données contient 85 lignes par individu. Le tableau 2 présente un exemple de type de base de données.

Tableau 2. Exemple de fichier des dates de suivi		
ID	mois	date
1	1	1996-04

1	2	1996-05
1	3	1996-06
1	...	
1	85	2003-04
2	1	1998-12
2	...	
2	85	2005-12
mois = Mois #1 à #j ; j étant la période de suivi divisée en mois du calendrier (p. ex. 85 mois).		
date = Permet de faire le lien entre la date index et la période de suivi.		

La quatrième étape consiste à jumeler le fichier avec les services reçus avec celui des dates de la période de suivi. Le tableau 3 présente un exemple de base de données complète, après jumelage de ces deux fichiers.

Tableau 3. Exemple de base de données combinant les fichiers avec les services reçus et les dates de la période de suivi								
Fichier des dates			Fichier des services					
id	mois	date	urg	hospit	psych	md_autre_spec	md_fam	Priorité
1	1	1996-04	1					urg
1	2	1996-05		1	1			hospit
1	3	1996-06			1	1		psych
1
1	85	2003-04						aucun
2

La cinquième et dernière étape implique un transfert de la base de données en format horizontal (une ligne par individu) plutôt qu'en format vertical (plusieurs lignes par individu), telle que créée à la quatrième étape. Ainsi, la base de données finale pour l'AS (tableau 4) est composée d'une ligne par individu et d'une colonne par fréquence d'observation (mensuelle dans notre étude), dans laquelle l'état prioritaire est indiqué et codifié.

Tableau 4. Exemple de base de données finale utilisée pour effectuer l'analyse de séquences							
id	mois_1	mois_2	mois_3	...	mois_84	mois_85	
1	1	2	2	...	4	4	
2	4	4	4	...	3	3	
...							
1 = urgence ou hospitalisation, 2 = consultation médicale avec un psychiatre (« psych »), 3 = consultation avec autre médecin spécialiste, autre que psychiatre (« md_autre_spec »), 4 = consultation médicale avec un médecin omnipraticien (« md_fam ») et 5 = aucune consultation.							

3.5.2. Analyse de classes et profils latents

La disposition des données diffère selon l'objectif de faire une analyse de classes ou de profils latents.

La constitution de la base pour l'ACL est similaire à celle de l'AS. La base de données des services médicaux a été utilisée et constitue la base d'intérêt. Pour chacun des services médicaux reçus par les répondants, nous avons déterminé, selon les codes associés, s'il s'agissait d'une consultation auprès d'un médecin omnipraticien, d'un psychiatre, d'un médecin spécialiste autre que psychiatre, d'un séjour hospitalier ou d'une visite à l'urgence. En plus d'avoir identifié ces catégories, une 5^e catégorie a été créée pour les répondants n'ayant utilisé aucun de ces services. Les années de suivi ont été décomposées en mois, et, pour chaque mois, nous avons inscrit « 0 » comme valeur si le service n'avait pas été utilisé, et « 1 » sinon. Si pour un mois donné, un répondant avait « 0 » partout pour les 4 principales variables, alors la variable « Aucun » prenait la valeur « 1 » ce qui veut dire qu'aucun service n'avait été utilisé, sinon « Aucun » prenait la valeur « 0 ».

La base de données d'utilisation des services médicaux ainsi reconstituée a été fusionnée avec celle des participants de la base ESCC, afin de ne retenir que les répondants adultes (18 ans et plus) ayant un TH. Au total, nous avons 4 421 individus pour 85 périodes de temps (mois), soit un total de 375 785 individus-mois.

Pour l'APL, la base de données contenant les informations socio-économiques des participants a été couplée avec la base de données de la RAMQ afin de constituer une base générale de données socio-économiques (base `escs_part`). Nos variables d'intérêt étaient le nombre de services médicaux reçus ainsi que le nombre de diagnostics reçus lors de séjours hospitaliers.

Tout d'abord, nous avons, pour chacune des variables d'intérêt, calculé le nombre moyen de fois que ces services ont été utilisés (nombre de services médicaux utilisés, nombre de diagnostics) sur la période d'étude allant de 5 années avant la date index à 2 années après. Ainsi, nous avons calculé le nombre moyen de services médicaux de toutes sortes en dehors des séjours hospitaliers. Pour les séjours hospitaliers, nous avons 5 variables : les hospitalisations, les diagnostics reçus, les services hospitaliers dans lesquels le répondant a été consulté lors de son séjour hospitalier, les interventions reçues, et finalement un éventuel séjour aux soins intensifs.

Chaque séjour hospitalier était identifié par un numéro unique à partir duquel nous avons pu retracer et regrouper l'ensemble du parcours hospitalier du répondant et les différents actes administrés et services utilisés. À partir de la base de données principale du séjour hospitalier, nous avons fusionné les données provenant des bases secondaires du séjour hospitalier (diagnostic, services médicaux, soins intensifs, interventions).

Les moyennes ont été calculées pour chaque répondant et pour chacune des années avant et après la date index. Nous avons fusionné chacune des bases de séjour hospitalier avec la base `escs_part` afin de constituer des ensembles distincts de données contenant l'information nécessaire aux estimations. Cinq bases ont ainsi été créées :

La base `escs_smed` : base des services médicaux reçus.

La base `escc_diag` : base des diagnostics reçus pendant le séjour hospitalier

La base `escc_serv` : base des services médicaux reçus pendant le séjour hospitalier

La base `escc_interv` : base des interventions reçues pendant le séjour hospitalier

La base `escc_intens` : base retraçant le passage aux soins intensifs pendant le séjour hospitalier

Ces 5 bases de données ont constitué le fondement de notre APL qui s'appuie essentiellement sur les services médicaux reçus et les diagnostics reçus en séjour hospitalier. Au total, nous avons 4 421 individus pour 8 périodes de temps (années), soit un total de 35 368 individus-années.

3.6. Analyse des données

3.6.1. Analyse de séquences

L'AS est une technique statistique qui permet d'étudier les successions d'états ou d'évènements. Elle compare des séquences chronologiques d'états au sein d'un modèle conceptuel global, où les observations ne sont pas considérées comme indépendantes dans le temps.

Puisque l'AS est une technique d'analyse relativement récente⁷ dans le domaine de la santé, encore peu de logiciels pour les analyses statistiques contiennent les fonctions nécessaires pour la réaliser. Les deux logiciels les plus utilisés sont Stata (avec ses progiciels SADI ou SQ-Ados 2.0) et R (avec le progiciel TraMineR). Dans le cadre de notre étude, nous avons opté pour le progiciel TraMineR pour effectuer les étapes de l'AAO et de l'analyse par grappes, tandis que Stata a été utilisé pour la construction des bases de données ainsi que pour les analyses descriptives.

Le terme trajectoire, dans le contexte de l'AS, représente la séquence qui est définie par l'enchaînement d'états ou d'évènements. Chaque individu présente une trajectoire (ou séquence) unique. Cependant, il est possible de repérer des ressemblances (ou des différences) entre les séquences individuelles le long de sous-séquences communes entre les individus.

Pour déterminer une TS dans le cadre de l'AS, il faut déterminer la période de suivi et la fréquence à laquelle nous allons observer l'état dans lequel se situe l'individu (p. ex. par jour, par semaine, par mois), les différents types de contacts pertinents avec le système de santé (qui seront considérés comme les différents états dans la séquence), ainsi que la priorité que nous voulons attribuer à chacun de ces contacts/états.

Comme mentionné à la section 2.3.1., nous avons retenu les cinq états suivants, selon l'ordre de priorité présenté :

1. Hospitalisation ou consultation à l'urgence
2. Consultation ambulatoire avec un médecin psychiatre
3. Consultation ambulatoire avec un médecin spécialiste autre que psychiatre
4. Consultation ambulatoire avec un omnipraticien
5. Aucune consultation

Chacune des séquences individuelles est composée de 85 éléments, chacun représentant un mois durant la période de suivi, et un état, selon la priorisation établie plus tôt. À titre d'exemple, si un individu a eu une hospitalisation et une consultation avec un psychiatre au mois #20, l'état indiqué dans le mois #20 sera l'hospitalisation.

Une fois les états et leur priorité définis, ainsi que la fréquence d'observation de ces états, déterminés, nous pouvons procéder à l'AS à partir de la base de données finale (tableau 4 ; données fictives présentées).

L'étape suivante consiste à comparer les séquences, pour évaluer les similarités et les différences entre celles-ci. Pour ce faire, nous comparons les séquences de chaque paire de l'échantillon entre elles et nous calculons le nombre d'opérations élémentaires (nommés *indels*) de substitutions, d'insertions ou de suppressions nécessaires pour transformer une séquence (source) en une autre (cible).

Le nombre d'opérations élémentaires nécessaires pour transformer une séquence en une autre est appelé la distance entre les deux séquences; plus la distance est grande, plus les séquences sont différentes l'une de l'autre. Il existe plusieurs stratégies de transformation, mais la plus utilisée est celle de l'AAO, qui utilise des techniques de programmation dynamique pour déterminer la combinaison spécifique d'opérations élémentaires requises pour transformer une séquence en une autre, à un coût minimal (c.-à-d. en termes d'opérations). De manière générale, il est admis que les opérations de suppression et d'insertion sont établies à 50 % de celui de la substitution (modification), car cette dernière implique deux opérations sous-jacentes. En effet, pour substituer un état par un autre, nous devons tout d'abord supprimer l'élément qui est différent entre les deux séquences (au coût de 1) et insérer l'élément correspondant à la séquence cible (au coût de 1).

Le tableau 5 présente un exemple de deux séquences, composées de 6 temps de mesures, pour lesquelles nous voulons calculer la distance, c'est-à-dire le coût nécessaire pour transformer la séquence B en la séquence A. Le coût total est de 6, puisque nous avons une suppression (au temps 3), une insertion (au temps 4) et deux modifications (aux temps 5 et 6).

Tableau 5. Opérations de transformations et coûts pour comparaison de séquences						
Séq. / État	1	2	3	4	5	6
A	H	H	-	A	A	U
B	H	H	U	-	U	H
Transformation	E	E	S	I	M	M
Coût	0	0	1	1	2	2
E : équivalent ; aucun coût associé. I : insertion ; coût de 1. M : modification ; coût de 2. S : suppression ; coût de 1. H : Hospitalisations. U : consultation à l'urgence. A : Consultation avec un autre médecin spécialiste.						

Une fois le calcul des distances pour chaque paire de séquences obtenue, nous pouvons effectuer une analyse par grappes (*cluster analysis*), qui nous permet de regrouper les séquences pour construire des typologies de séquences similaires. La méthode la plus utilisée est celle de Ward⁷, dont l'algorithme permet de regrouper des séquences pour minimiser la variance à l'intérieur de celles-ci, et maximiser les différences entre les groupes créés.

Pour déterminer le nombre optimal de groupes à retenir, il est utile de présenter un dendrogramme. Celui-ci utilise une méthode hiérarchique agglomérée en considérant chaque séquence au départ pour ensuite regrouper des séquences en fonction de leur distance, de manière hiérarchique. Une fois l'analyse par grappes complétée, nous pouvons ensuite utiliser cette nouvelle variable comme une variable catégorique pour effectuer des comparaisons entre les groupes de séquences déterminés par l'analyse.

3.6.2. Analyse de classes et profils latents

Tout le traitement des données a été fait avec le logiciel Stata.

L'analyse des données pour l'ACL s'est faite avec RStudio tandis que nous avons utilisé Stata pour l'APL.

Concernant l'ACL, nous avons utilisé le package `poLCA` dans RStudio qui permet l'analyse des classes basée sur des données dichotomiques. Précisons que la méthode d'estimation (par maximum de vraisemblance) de la fonction `poLCA` ne supporte pas les valeurs nulles (0). De ce fait, nous avons retransformé les données comme suit : 1 = utilisation du service ; 2 = pas d'utilisation du service. Nous avons fait plusieurs estimations en modifiant à chaque fois le nombre de classes allant de 2 à 5. Après comparaisons des critères d'informations (AIC et BIC), nous avons retenus le modèle à 4 classes. Nous avons par la suite caractérisé les 4 classes à partir de leurs probabilités et des variables.

Pour déterminer les profils des répondants (APL), nous avons estimé un modèle de combinaisons finies (*finite mixture model*). Ce modèle exige la spécification du nombre de profils (groupes latents) à estimer et la forme de la fonction de distribution⁸. Pour la fonction de distribution, compte tenu de la nature des données, nous avons retenu la fonction binomiale négative pour le nombre de services médicaux et de diagnostics. Une distribution de Poisson aurait également pu être retenue, mais cela n'a pas été fait, car le modèle binomial négatif est une généralité et le modèle de Poisson est un cas spécifique de ce dernier. Par ailleurs, le modèle de Poisson fait l'hypothèse de l'égalité entre la moyenne et la variance, ce qui n'était pas vraisemblable pour nos données en raison de la surdispersion (figure 10 et 14).

Comme pour l'ACL, nous avons estimé plusieurs modèles différents en changeant à chaque fois le nombre de profils. L'objectif était de retenir le modèle qui avait le plus de sens dans le contexte de l'étude avec des paramètres optimaux pour la décision statistique⁹. Nous avons calculé à chaque estimation les AIC et BIC afin de déterminer le modèle de nombre de profils qui correspondait le mieux aux données. Finalement nous avons retenu un modèle à 4 profils pour les services médicaux et à 2 profils pour les hospitalisations.

Après l'estimation économétrique du modèle, nous avons calculé deux paramètres principaux : les moyennes marginales et les probabilités a posteriori.

La moyenne marginale permet d'obtenir la moyenne de la variable d'intérêt respectivement pour chaque profil latent. Nous avons donc calculé 4 moyennes marginales pour les services médicaux et 2 moyennes marginales pour les hospitalisations. La probabilité a posteriori permet de

déterminer la probabilité qu'a un participant d'appartenir à l'un des profils latents à l'issue de l'estimation du modèle de régression. Cette probabilité représente donc l'importance de chacun des profils dans la partition de l'échantillon.

Au regard des valeurs des moyennes marginales dans chacun des profils, il est ensuite possible de caractériser les profils et de les représenter visuellement¹⁰.

3.6.3. Analyse des coûts

Après la détermination des TS selon chacune des méthodes (AS, ACL et APL), nous avons calculé les coûts médicaux directs, selon la perspective du système de santé.

Les catégories de coûts disponibles sont les coûts relatifs aux hospitalisations, aux visites en chirurgie d'un jour (seulement à partir de 2010), aux visites à l'urgence, aux visites médicales (avec omnipraticien ou médecin spécialiste) ainsi que les frais pharmaceutiques pour les personnes couvertes par le régime public de la RAMQ, soient les personnes âgées de 65 ans et plus ou en situation de précarité financière.

Le tableau 6 présente chacune de ces catégories de coûts et leur description, la base de données d'origine, les années disponibles et la procédure de valorisation monétaire. Le niveau d'intensité relative des ressources utilisées (NIRRU) est une mesure attribuée par le Ministère de la santé et des services sociaux (MSSS) pour chacune des visites d'hospitalisation et de chirurgie d'un jour (à partir de 2010 seulement). La valeur du NIRRU peut être inférieure ou supérieure à 1. Il suffit ensuite d'attribuer la valeur monétaire correspondant à un NIRRU de l'année de référence (2019 dans notre cas). La valeur monétaire d'un NIRRU était de 6 026,06 \$ pour une hospitalisation, 1 368,30 \$ pour une chirurgie d'un jour et de 355,60 \$ pour une visite à l'urgence.

Tableau 6. Catégories de coûts pour l'analyse des coûts selon les trajectoires de soins

Catégorie de coûts	Description	Base de données d'origine	Années disponibles	Procédure de valorisation
Hospitalisations	Coûts directs et indirects calculés à partir du NIRRU associé à l'hospitalisation, excluant la rémunération du médecin.	APR-DRG	1996 à 2016	Valeur monétaire d'un NIRRU pour hospitalisation équivaut à 6 026,06 \$ en 2019. Multiplication du NIRRU obtenu par le MSSS par la valeur monétaire d'un NIRRU.
Chirurgie d'un jour	Coûts directs et indirects calculés à partir du NIRRU associé à la chirurgie d'un jour, excluant la rémunération du médecin.	MED-ÉCHO	2010 à 2016	Valeur monétaire d'un NIRRU pour chirurgie d'un jour équivaut à 1 368,30 \$ en 2019. Multiplication du NIRRU obtenu par le MSSS par la valeur monétaire d'un NIRRU.
Visites à l'urgence	Estimation des coûts directs et indirects calculés à partir du NIRRU associé à une visite à	RAMQ	1996 à 2016	Estimation effectuée, car valeur réelle du NIRRU par individu par visite à l'urgence n'est pas connu. Nous avons

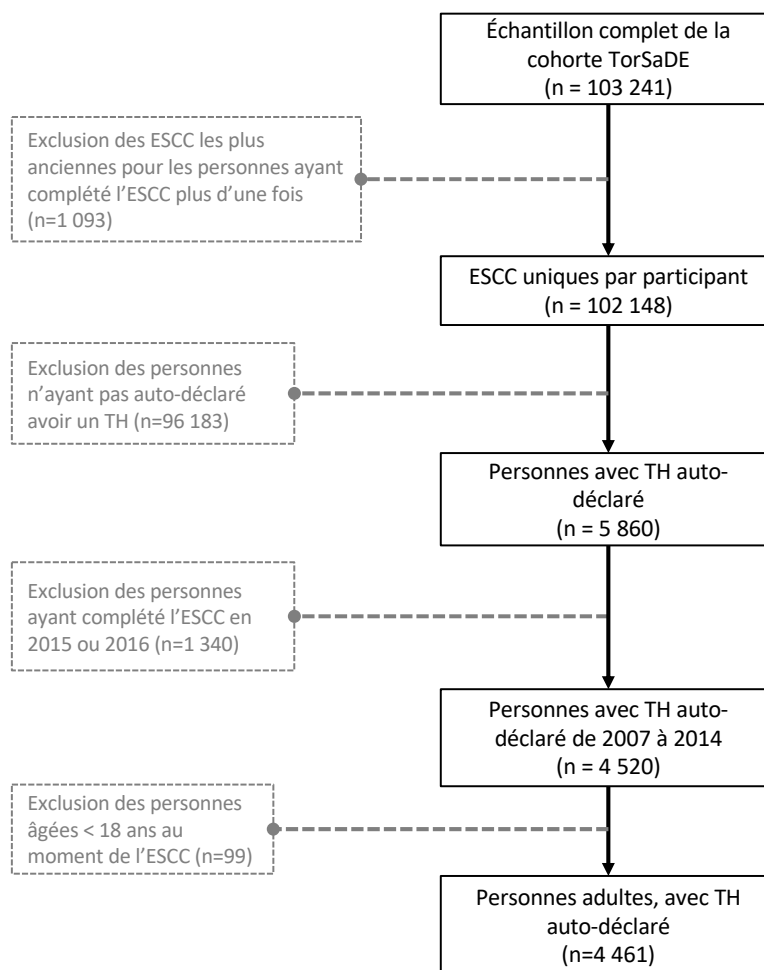
	l'urgence, excluant la rémunération du médecin.			attribué la valeur monétaire d'un NIRRU pour visite à l'urgence (355,60 \$) à chaque visite à l'urgence répertoriée.
Visites médicales	Montant facturé par le médecin à la RAMQ pour l'acte médical effectué lors de la visite.	RAMQ	1996 à 2016	Dollars courants de l'année facturée, transformés en dollars constants de 2019 au moyen de l'IPC.
Services pharmaceutiques	Frais de service et du coût du médicament couvert par la RAMQ.	RAMQ	1996 à 2016	Dollars courants de l'année facturée, transformés en dollars constants de 2019 au moyen de l'IPC.
IPC : Indice du prix à la consommation ; NIRRU : Niveau d'intensité relative des ressources utilisées ; RAMQ : Régie de l'assurance maladie du Québec.				

4. Résultats

4.1. Description de l'échantillon

La figure 3 présente le diagramme de flux des participants. L'échantillon complet de la cohorte TorSaDE compte un peu plus de 103 000 entrées d'ESCC, parmi lesquelles 1 093 ont été complétées par des individus ayant complété l'ESCC plus d'une fois. En gardant seulement l'entrée de l'ESCC la plus récemment complétée pour ces individus, nous avons 102 148 entrées d'ESCC uniques. Parmi celles-ci, 5 860 répondants ont auto-déclaré avoir un TH. Après élimination des individus ayant complété l'ESCC en 2015 et 2016, puisque nous avons besoin de 2 années de données médico-administratives après la complétion de l'ESCC, nous retenons 4 461 adultes avec TH.

Figure 3. Diagramme de flux des participants.



Le tableau 7 présente la répartition des participants adultes ayant complété une ESCC entre 2007 et 2014, inclusivement, selon la présence d'un TH auto-déclaré (ESCC) ou diagnostiqué (RAMQ).

Près de 10 000 personnes (9 765) ont eu au moins une consultation médicale associée à un diagnostic de TH selon la base de données de la RAMQ et durant la période de suivi retenue (5 ans pré- et 2 ans post-ESCC). Parmi celles-ci, un peu plus de 70 % d'entre elles n'ont pas déclaré avoir reçu un diagnostic de TH (6 969 / 9 765 = 71,4 %)

Parmi les personnes ayant auto-déclaré avoir un TH, 36,8%, (1 625 / 4 421) d'entre elles n'avaient pas de diagnostic de TH selon la RAMQ.

Tableau 7. Répartition des participants adultes selon un trouble de l'humeur auto-rapporté ou un diagnostic dans les données de la RAMQ.				
		Données de la RAMQ		TOTAL
		Pas de code diagnostic de TH (0)	Présence d'un code diagnostic de TH (1)	
Données de l'ESCC	TH non auto-déclaré (0)	62 748	6 969	69 717
	TH auto-déclaré (1)	1 625	2 796	4 421

TOTAL	64 373	9 765	74 138
--------------	---------------	--------------	---------------

Le tableau 8 présente les données descriptives de l'échantillon, composé de 4 421 adultes québécois, ayant déclaré avoir un TH dans leur ESCC complétée entre 2007 et 2014, comparativement à l'échantillon complet de l'étude (cohorte TorSaDE, soient 80 651 personnes ayant répondu aux ESCC entre 2007 et 2014). L'âge moyen des personnes atteintes de TH est de 50,7 ans \pm 15,7, dont 60,8 % sont des femmes, comparativement à 48,2 ans \pm 20,2 dans l'ensemble de la cohorte TorSaDE, composée à 50,6 % de femmes. L'indice de masse corporelle (IMC) semble un peu plus élevé chez les personnes atteintes de TH (27,2 kg/m²), équivalent à la catégorie de « surpoids », que celui de la population générale (25,6 kg/m²) qui se rapproche plutôt d'un poids sain. Près d'un tiers des individus ont déclaré être célibataire ou ne jamais avoir été marié au moment de la complétion de leur enquête, tant chez les personnes atteintes de TH que celles de la population générale, tandis qu'un peu plus d'entre elles étaient mariées (35,6 %) comparativement à l'échantillon avec TH (28,1 %). Le revenu annuel par répondant semblait être plus élevé dans la population générale, avec 37,9 % des individus ayant déclaré plus 40 000\$ par année, comparativement à 26,7 % chez les personnes avec TH.

Le niveau de stress perçu a été mentionné comme étant présent dans la vie quotidienne et au travail par une majorité des individus, tandis que les états de santé générale et de santé mentale ont été rapportés comme étant généralement positifs. En effet, un peu plus de 80 % des personnes ont déclaré ressentir du stress dans leur vie quotidienne (80,9 %) et au travail (83,2 %), alors qu'ils ont indiqué avoir une perception favorable de leur santé mentale et de leur santé globale (67,8 % et 65,0 % les ont qualifiées de bonne à excellente, respectivement).

Quant à l'utilisation des services de santé retenus, elle est plus élevée chez les personnes atteintes de TH, particulièrement pour les consultations avec un psychiatre et avec un médecin omnipraticien. Nous rapportons peu d'hospitalisations, avec un nombre médian de 3 [intervalle interquartile (IIQ) : 1-7], malgré une proportion élevée (81,2 %) d'individus ayant eu au moins une hospitalisation selon les données disponibles de la RAMQ chez les personnes atteintes de TH. Les consultations à l'urgence étaient plus fréquentes (18 [8-36]), avec 95,1 % des répondants ayant eu au moins une consultation à l'urgence durant leur période de suivi de 7 années. En ce qui concerne les consultations ambulatoires, le nombre médian était de 104 consultations [62-166] avec un médecin omnipraticien, de 63 consultations [31-117] avec un médecin spécialiste (autre que psychiatre) et de 3 consultations [0-32] avec un psychiatre.

Tableau 8. Données descriptives de l'échantillon avec TH et échantillon TorSaDE			
Caractéristiques	Échantillon TH (n = 4 421)	Échantillon TorSaDE (n = 80 651)	Valeur p
Âge, années	50,7 \pm 15,7	48,2 \pm 20,2	<0,001
Sexe, %			
Femmes	60,8 %	50,6 %	<0,001
Hommes	39,2 %	49,4 %	<0,001
Indice de masse corporelle, kg/m²	27,2 \pm 6,04	25,6 \pm 5,18	<0,001
Poids insuffisant (IMC<18)	1,76 %	2,88 %	<0,001
Poids normal (18 \geq IMC<25)	37,6 %	46,8 %	<0,001

Surpoids (25≥IMC<30)	30,1 %	30,3 %	0,7574
Obésité classe I (30≥IMC<35)	16,1 %	11,2 %	<0,001
Obésité classe II (35≥IMC<40)	5,52 %	2,80 %	<0,001
Obésité classe III (IMC≤40)	8,99 %	6,00 %	<0,001
Niveau d'éducation, %			
Inconnu	5,86 %	7,35 %	0,0148
Secondaire non complété	21,5 %	22,7 %	0,0007
Diplôme d'études secondaires	14,0 %	12,7 %	0,0742
Diplôme professionnel ou CÉGEP	25,6 %	22,3 %	<0,001
Baccalauréat	21,0 %	21,5 %	0,0109
Études universitaires supérieures	12,1 %	13,5 %	<0,001
Statut matrimonial, %			
Marié	28,1 %	35,8 %	<0,001
Union libre	18,6 %	20,8 %	<0,001
Veuf	4,98 %	4,86 %	0,0522
Séparé	4,15 %	2,54 %	<0,001
Divorcé	12,6 %	6,02 %	<0,001
Célibataire/Jamais marié	31,6 %	30,0 %	0,0098
Revenu annuel par ménage, %			
Moins de 20 000 \$	21,9 %	10,2 %	<0,001
20,000 à moins de 40 000 \$	25,1 %	21,1 %	<0,001
40,000 à moins de 60 000 \$	19,2 %	19,6 %	0,3936
60,000 à moins de 80 000 \$	12,8 %	16,2 %	<0,001
80 000 \$ et plus	21,1 %	32,9 %	<0,001
Revenu annuel par répondant, %			
Moins de 20 000 \$	44,6 %	32,2 %	<0,001
20,000 à moins de 40 000 \$	28,7 %	30,1 %	0,5720
40,000 à moins de 60 000 \$	14,7 %	19,6 %	<0,001
60,000 à moins de 80 000 \$	7,63 %	10,1 %	<0,001
80 000 \$ et plus	4,39 %	8,02 %	<0,001
Pays de naissance, %			
Canada et Amérique du Nord	91,4 %	86,1 %	<0,001
Amérique Centrale et du Sud, Bermudes et Antilles	1,87 %	3,20 %	<0,001
Europe	3,44 %	4,62 %	<0,001
Asie et Océanie	1,71 %	3,25 %	<0,001
Afrique	1,54 %	2,82 %	<0,001

Tableau 8. Données descriptives de l'échantillon avec TH et échantillon TorSaDE (suite)			
Caractéristiques	Échantillon TH (n = 4 421)	Échantillon TorSaDE (n = 80 651)	Valeur p
Perception du stress – Vie quotidienne, %			
Pas du tout stressé	6,53 %	14,0 %	<0,001
Pas tellement stressé	12,6 %	21,3 %	<0,001
Un peu stressé	34,2 %	38,8 %	<0,001
Assez stressé	34,5 %	22,4 %	<0,001
Extrêmement stressé	12,2 %	3,57 %	<0,001

Perception du stress – Travail, %			
Pas du tout stressé	5,92 %	8,89 %	<0,001
Pas tellement stressé	11,0 %	15,6 %	<0,001
Un peu stressé	31,8 %	39,4 %	<0,001
Assez stressé	38,0 %	30,7 %	<0,001
Extrêmement stressé	13,4 %	5,50 %	<0,001
Perception de la santé générale, %			
Mauvaise	9,59 %	2,01 %	<0,001
Passable	25,5 %	8,23 %	<0,001
Bonne	36,5 %	30,4 %	<0,001
Très bonne	22,6 %	36,3 %	<0,001
Excellente	5,87 %	23,1 %	<0,001
Perception de la santé mentale, %			
Mauvaise	7,92 %	0,56 %	<0,001
Passable	24,9 %	3,39 %	0,0462
Bonne	41,6 %	20,5 %	<0,001
Très bonne	19,1 %	35,8 %	<0,001
Excellente	6,56 %	39,8 %	<0,001
HUI-3^ß, score	0,68 ± 0,30	0,88 ± 0,18	<0,001
Nombre de consultations médicales d'intérêt, moyenne ± écart-type ; médiane [intervalle interquartile]			
Hospitalisations	5,38 ± 7,61	3,16 ± 4,81	<0,001
	3 [1 – 7]	2 [0 – 4]	<0,001 ^α
Urgence	28,5 ± 35,4	17,2 ± 23,7	<0,001
	18 [8 – 36]	10 [4 – 22]	<0,001 ^α
Psychiatre	40,4 ± 101,8	5,01 ± 33,5	<0,001
	3 [0 – 32]	0 [0 – 0]	<0,001 ^α
Autre médecine spécialiste	92,9 ± 112,5	71,1 ± 90,9	<0,001
	63 [31 – 117]	44 [18 – 92]	<0,001 ^α
Médecin omnipraticien	127,5 ± 103,2	75,9 ± 76,3	<0,001
	104 [62 – 166]	55 [26 – 101]	<0,001 ^α
Proportion d'individus ayant eu au moins une consultation selon le type de service, %			
Hospitalisations	81,2 %	68,2 %	<0,001
Urgence	95,4 %	88,5 %	<0,001
Psychiatre	59,8 %	14,8 %	<0,001
Autre médecine spécialiste	99,6 %	98,1 %	<0,001
Médecin omnipraticien	99,7 %	99,3 %	0,1517
^ß Score non disponible pour individus ayant complété le cycle 2011-2012 de l'ESCC (n = 19 966).			
^α Comparaison effectuée par le Test t de Wilcoxon, pour échantillon unique.			

4.2. Résultats obtenus par l'analyse de séquences

4.2.1. Données descriptives des séquences

Les résultats de l'AS permettent de constater que les trajectoires individuelles sont très différentes les unes des autres. En effet, parmi les 4 421 répondants, il y a 4 407 séquences différentes, soit un niveau de concentration de 99,6 % (4 405 séquences différentes / 4 421 répondants). Seuls 14 individus partagent une séquence commune, soit celle reflétant aucune

utilisation des services de santé, et 2 autres répondants partagent la séquence avec une seule consultation avec leur médecin omnipraticien au 8^e mois de suivi.

Les caractéristiques descriptives des séquences nous permettent d’observer que la moitié des répondants avaient 4 à 5 états différents composant leur séquence, et qu’en moyenne les séquences étaient composées de $41,2 \pm 14,0$ épisodes, c’est-à-dire des sous-séquences composées d’éléments identiques subséquents. L’élément le plus fréquemment répété de manière consécutive à l’intérieur des séquences était l’état d’absence d’utilisation des services de santé, avec une moyenne de $15,7 \text{ mois} \pm 5,96$ consécutifs, suivi des consultations avec un médecin omnipraticien ($10,4 \text{ mois} \pm 5,74$), avec un médecin spécialiste autre que psychiatre ($8,19 \pm 5,84$), d’une hospitalisation ou consultation à l’urgence ($4,21 \pm 3,97$), et enfin avec un psychiatre ($2,73 \pm 5,65$; 0 [0-2]).

La figure 4 présente la moyenne individuelle de mois passés dans chaque état. Ainsi, les répondants n’ont utilisé aucun service de santé durant 50 mois en moyenne, soit près de 60 % du temps composant la séquence ($50 / 85 = 58,8\%$). Le service de santé le moins fréquent était encore une fois les consultations avec un psychiatre, avec une moyenne 4,47 mois (5,25 %), suivi des hospitalisations et des consultations à l’urgence à 5,41 (6,36 %). Comme attendu, l’état le plus fréquent pour chacun des mois composant les séquences était celui d’absence d’utilisation des services de santé (figure 5), qui représentait plus de 50 % des types de contacts avec le système de santé.

Figure 4. Moyenne de temps passé (en mois) par état

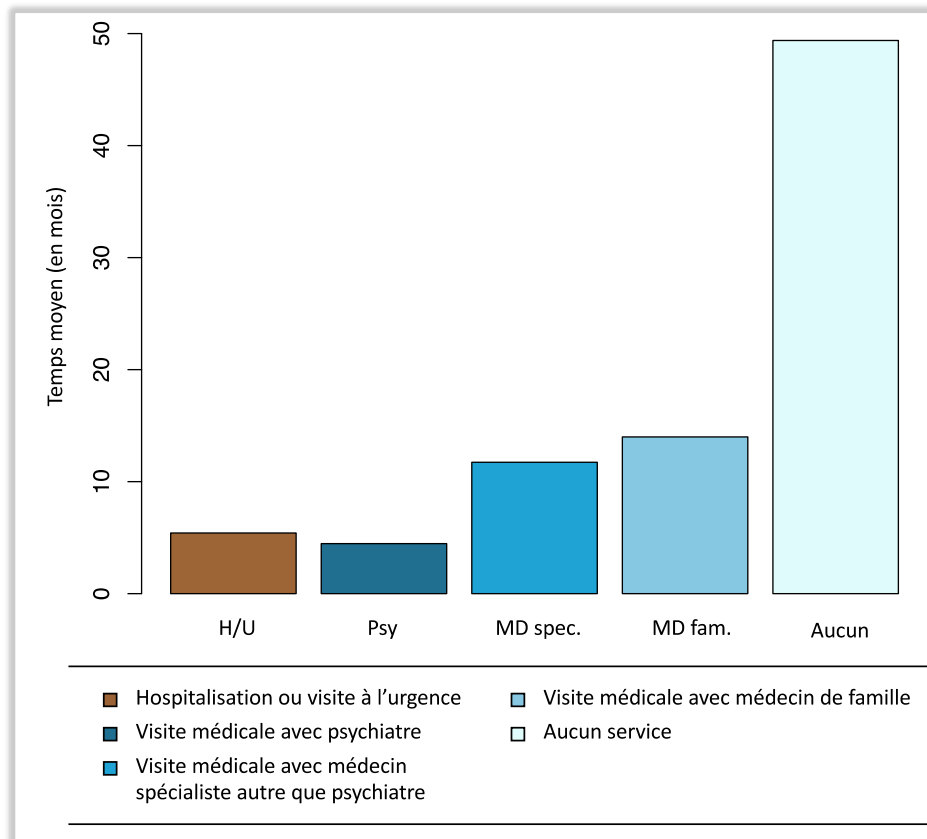
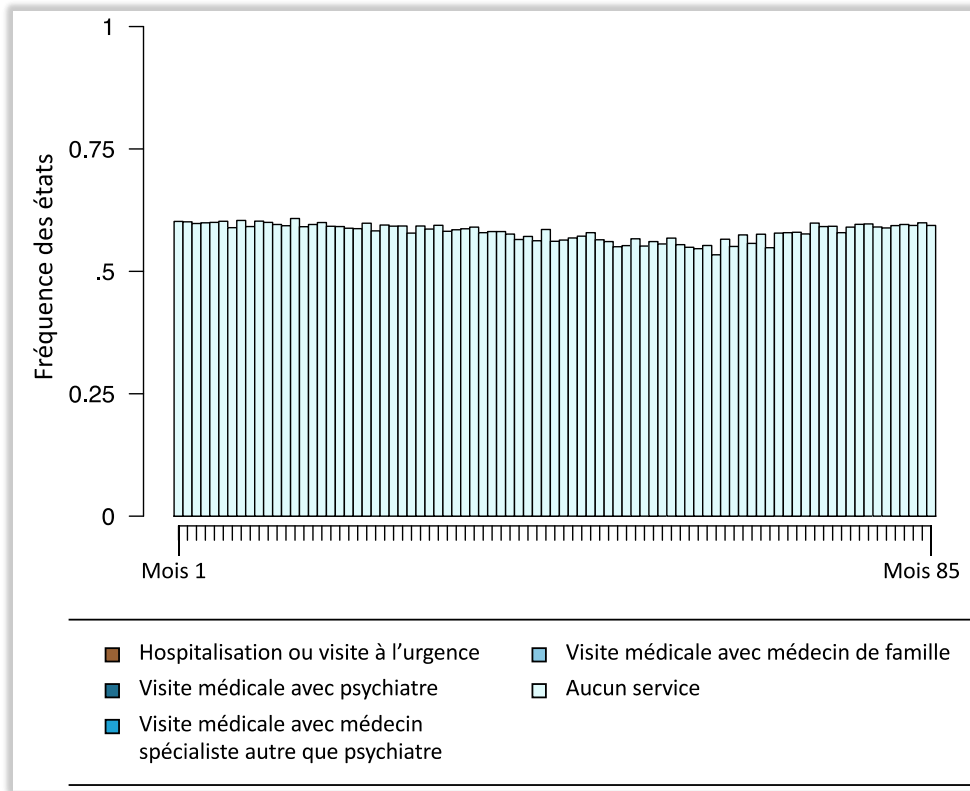


Figure 5. Proportion de l'état le plus fréquent par mois de suivi

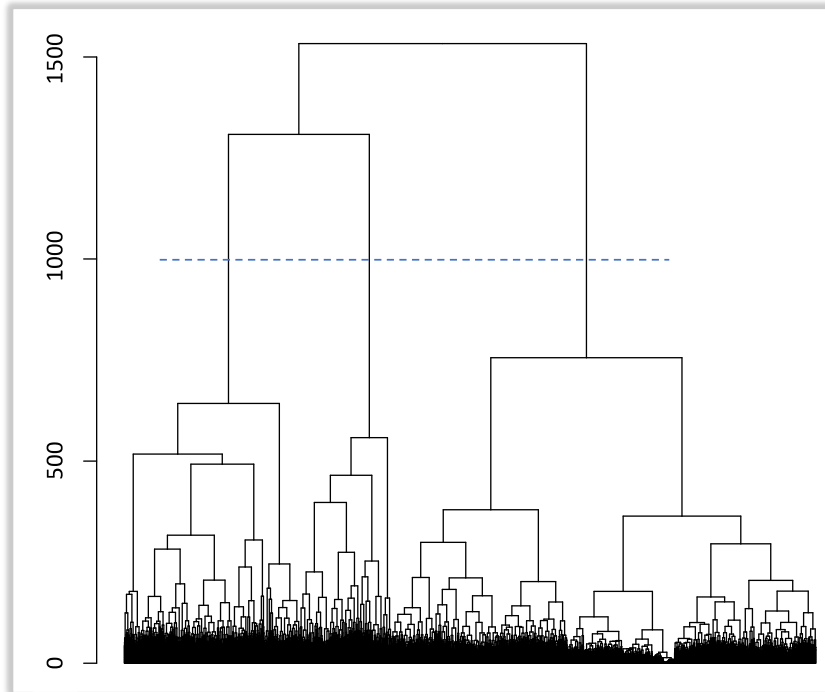


4.2.2. Résultats de l'analyse par grappes

Une fois construites, les séquences ont été comparées à partir de l'AAO, desquelles le calcul de distance a permis d'effectuer l'analyse par grappes. Pour identifier le nombre optimal de groupes à retenir suite à l'analyse par grappes, deux aides visuelles ont été produites.

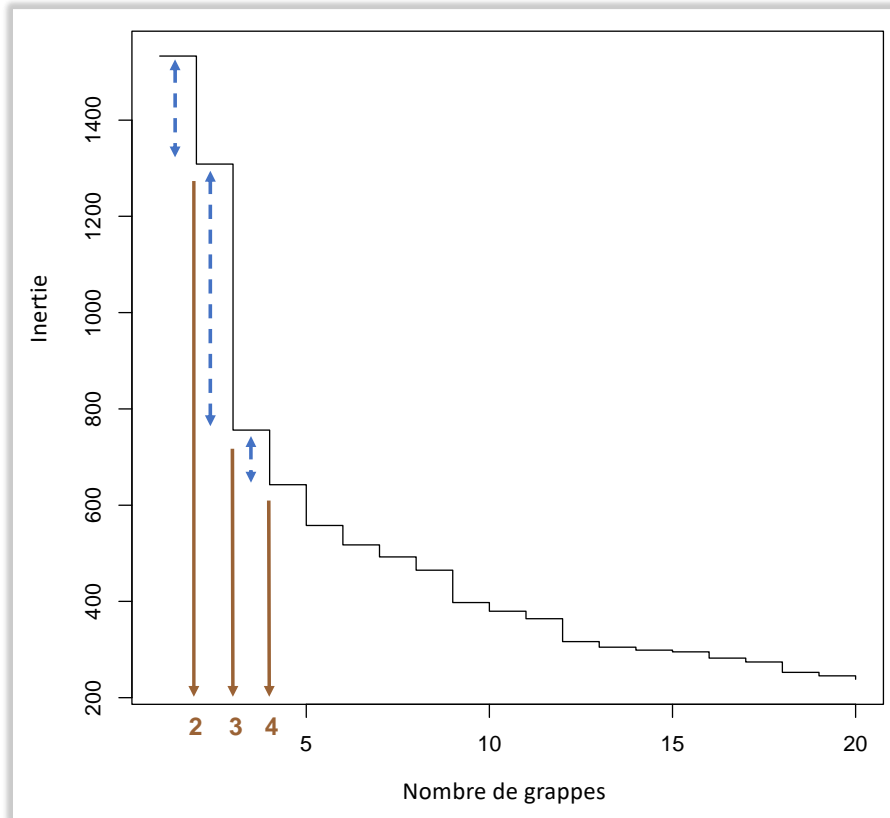
La première présente un dendrogramme (figure 6) qui nous permet de visualiser les distances entre les séquences (complètement au bas de la figure) et les distances entre les regroupements de séquences similaires au fur et à mesure que celles-ci sont jumelées. Chaque plateau représente un regroupement. Pour identifier le nombre optimal de groupes, il suffit de passer sur une ligne horizontale dans le diagramme, là où nous observons la plus grande distance verticale entre les différents plateaux (ligne pointillée sur la figure 6).

Figure 6. Dendrogramme issu de l'analyse par grappes après analyse de séquences



La deuxième aide visuelle présente un graphique des sauts d'inertie (figure 7). Celui-ci présente sur l'axe des y l'inertie, qui représente de manière générale la différence, en termes de variance et de covariance, entre les groupes. L'objectif étant de choisir un nombre de groupes qui va maximiser la différence entre ceux-ci, mais minimiser la différence entre les individus qui les composent. Lorsque l'on regarde la figure 7, nous observons qu'il y a une forte augmentation de l'inertie lorsque l'on passe de 3 à 2 groupes, et celle-ci augmente, mais de manière plus faible lorsque nous passons de 4 à 3 groupes. Ces observations, combinées à celles obtenues avec le dendrogramme, suggèrent que le nombre optimal de groupes dans notre échantillon est de 3.

Figure 7. Graphique des sauts d'inertie



La figure 8 présente la fréquence mensuelle des types de contacts avec le système de santé par répondant. Il est possible d'observer une différence nette d'utilisation des services de santé entre les trois groupes obtenus suite à l'analyse par grappes. Le type 1 (n = 2 714) est le groupe le plus large. Les individus composant ce groupe ont eu peu de contacts avec le système de santé durant la période de suivi de 85 mois, avec seulement environ 30 % des répondants ayant eu un contact avec le système de santé. Le type 2 (n = 1 120) semble avoir une fréquence plus élevée d'hospitalisations et de consultations avec un médecin omnipraticien ou un médecin spécialiste autre qu'un psychiatre. Enfin, le troisième groupe (n = 587) est surtout composé de grands utilisateurs des services spécialisés d'un psychiatre.

Une comparaison des données descriptives entre les trois types d'utilisateurs (tableau 9) permet de constater qu'une plus grande proportion des individus du type 3 avait obtenu un diplôme universitaire d'études supérieures (16,3 % vs 10,9 % et 11,7 % pour les types 2 et 1, respectivement), était célibataire ou jamais marié (42,2 % vs 19,9 % et 33,6 %), gagnait moins de 20 000\$ par année (52,3 % vs 45,3 % et 42,7 %) et avait déclaré percevoir leur santé mentale comme étant mauvaise (14,0 % vs 7,16 % et 6,97 %) ou passable (36,0 % vs 23,5 % et 23,1 %).

Figure 8. Distribution mensuelle des répondants selon le type de contact avec le système de santé (états)

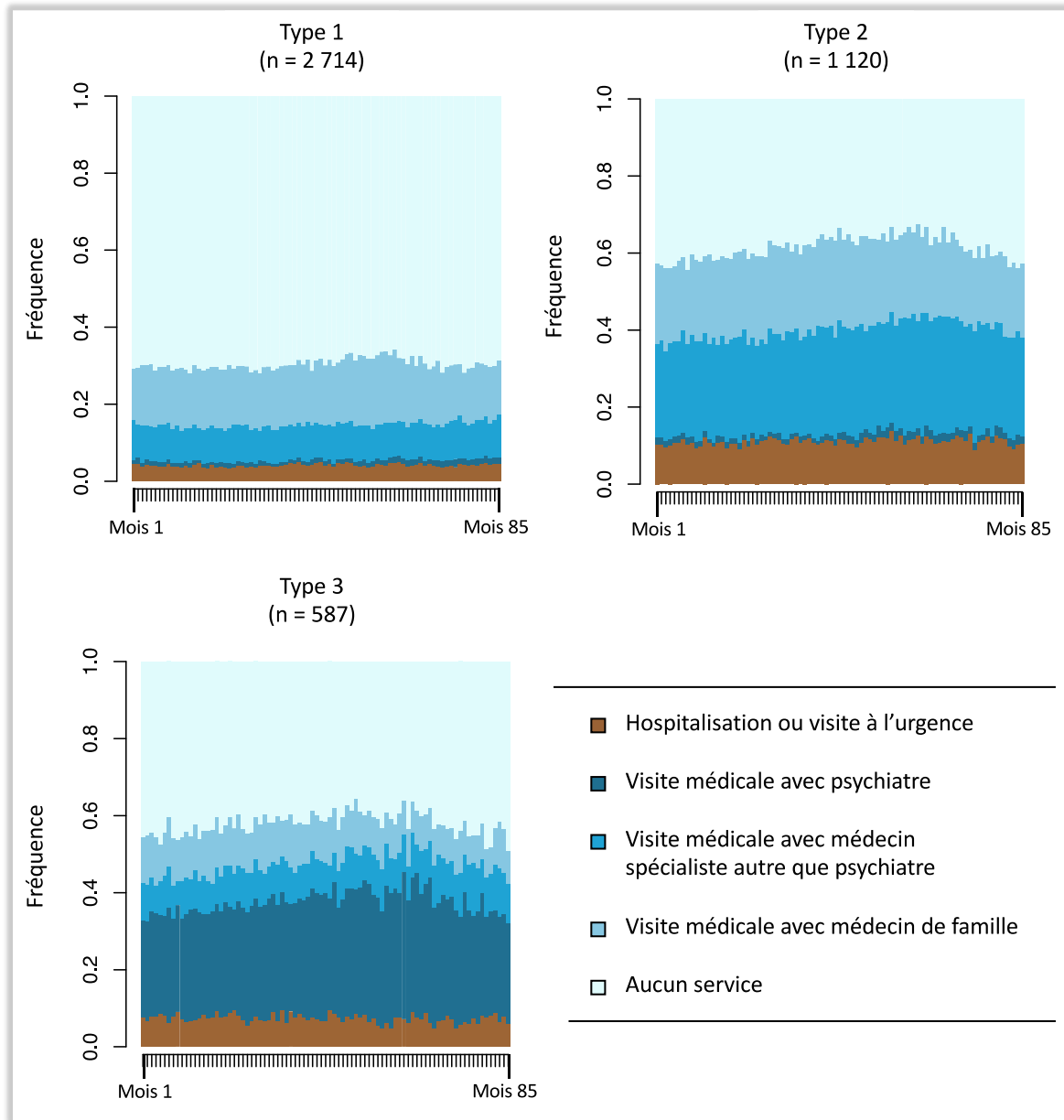


Tableau 9. Comparaison des données descriptives entre groupes issus de l'analyse par grappes

Caractéristiques	Type 1 (n = 2 714)	Type 2 (n = 1 120)	Type 3 (n = 587)	Valeur p
Âge, années	48,6 ± 15,5	56,5 ± 15,3	49,3 ± 14,3	< 0,001
Sexe, %				< 0,001
Femmes	56,6 %	72,6 %	60,4 %	
Hommes	43,4 %	27,4 %	39,6 %	
Indice de masse corporelle, kg/m ²	26,6 ± 5,71	28,2 ± 6,65	27,7 ± 6,03	< 0,001
Niveau d'éducation, %				< 0,001

Inconnu	6,16 %	3,21 %	9,04 %
Secondaire non complété	20,6 %	27,2 %	15,7 %
Diplôme d'études secondaires	14,8 %	13,4 %	11,1 %
Diplôme professionnel ou CÉGEP	25,2 %	25,9 %	26,8 %
Baccalauréat	21,6 %	19,4 %	21,2 %
Études universitaires supérieures	11,7 %	10,9 %	16,3 %
Statut matrimonial, %			< 0,001
Marié	27,1 %	34,8 %	20,8 %
Union libre	18,9 %	17,2 %	15,0 %
Veuf	3,83 %	9,34 %	2,97 %
Séparé	4,42 %	2,94 %	4,80 %
Divorcé	11,1 %	15,8 %	14,3 %
Célibataire/Jamais marié	33,6 %	19,9 %	42,2 %
Revenu annuel par ménage, %			< 0,001
Moins de 20 000 \$	20,4 %	22,7 %	27,5 %
20,000 à moins de 40 000 \$	23,8 %	30,1 %	23,0 %
40,000 à moins de 60 000 \$	18,7 %	20,5 %	19,2 %
60,000 à moins de 80 000 \$	13,5 %	10,5 %	13,6 %
80 000 \$ et plus	23,7 %	16,2 %	16,7 %
Revenu annuel par répondant, %			< 0,001
Moins de 20 000 \$	42,7 %	45,3 %	52,3 %
20,000 à moins de 40 000 \$	27,5 %	34,8 %	24,2 %
40,000 à moins de 60 000 \$	15,2 %	12,7 %	16,0 %
60,000 à moins de 80 000 \$	8,94 %	5,45 %	5,00 %
80 000 \$ et plus	5,65 %	1,80 %	2,60 %
Pays de naissance, %			0,383
Canada et Amérique du Nord	91,8 %	93,2 %	86,7 %
Amérique Centrale et du Sud, Bermudes et Antilles	1,94 %	1,46 %	2,27 %
Europe	3,13 %	3,58 %	4,73 %
Asie et Océanie	1,87 %	0,62 %	2,85 %
Afrique	1,28 %	1,13 %	3,55 %

Tableau 9. Comparaison des données descriptives entre groupes issus de l'analyse par grappes (suite)

Perception du stress – Vie quotidienne, %			0,530
Pas du tout stressé	6,40 %	7,57 %	5,38 %
Pas tellement stressé	11,9 %	12,6 %	16,11 %
Un peu stressé	34,3 %	32,7 %	36,5 %
Assez stressé	35,6 %	34,6 %	29,0 %
Extrêmement stressé	11,9 %	12,6 %	13,1 %

Perception du stress – Travail, %				0,316
Pas du tout stressé	5,02 %	10,9 %	4,82 %	
Pas tellement stressé	11,1 %	9,37 %	12,4 %	
Un peu stressé	31,5 %	28,3 %	39,2 %	
Assez stressé	38,0 %	39,6 %	34,6 %	
Extrêmement stressé	14,3 %	11,9 %	8,95 %	
Perception de la santé générale, %				< 0,001
Mauvaise	6,06 %	18,6 %	11,20 %	
Passable	21,9 %	33,0 %	30,3 %	
Bonne	38,1 %	29,1 %	41,1 %	
Très bonne	26,5 %	16,9 %	13,1 %	
Excellente	7,43 %	2,39 %	4,30 %	
Perception de la santé mentale, %				< 0,001
Mauvaise	6,97 %	7,16 %	14,0 %	
Passable	23,1 %	23,5 %	36,0 %	
Bonne	43,7 %	40,0 %	33,6 %	
Très bonne	19,5 %	22,1 %	12,0 %	
Excellente	6,76 %	7,22 %	4,41%	
Indice de la qualité de vie (HUI-3)^β, score	0,73 ± 0,27 (n=1927)	0,59 ± 0,33 (n=832)	0,67 ± 0,30 (n=432)	< 0,001
^β Score non disponible pour individus ayant complété le cycle 2011-2012 de l'ESCC (n = 1 111).				

4.2.3. Résultats des coûts à partir de l'analyse de séquences

Pour chacun des individus inclus dans l'AS, nous avons calculé les coûts médicaux directs, séparés selon les catégories de coûts présentés à la section 2.6.3.

L'Annexe 3A présente les coûts totaux moyens par individu, selon l'attribution au type de trajectoire définie par l'AS. Il est possible de constater que dans la majorité des cas, des différences significatives peuvent être observées entre les types d'utilisateurs, sur la base des coûts moyens individuels.

4.3. Résultats obtenus par l'analyse de classes et profils latents

Pour l'ACL, nous présentons tous les services utilisés. Concernant l'APL, considérant la nature des données utilisées, les résultats sont présentés pour les services médicaux et les hospitalisations. Puisque l'APL utilise une combinaison de données continues, il nous a semblé pertinent d'effectuer les analyses séparément pour les services médicaux et les hospitalisations. Cela dans l'optique de capter la relative intensité dans l'utilisation des soins de santé, que ce soit en ambulatoire (services médicaux) qu'en séjour hospitalier (hospitalisations).

4.3.1. ACL : Tous les services médicaux

Le tableau 11 présente l'évolution des critères d'information statistique pour différentes classes. Les critères AIC et BIC sont minimaux dans la situation avec 4 classes. Nous avons donc retenu le modèle à 4 classes.

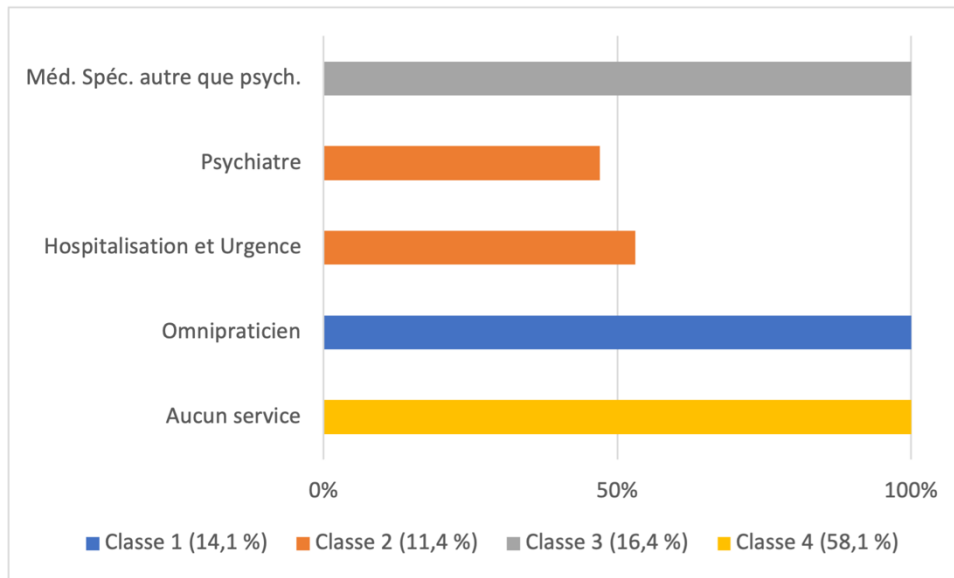
Tableau 11. Critères statistiques post-estimation			
	3 classes	4 classes	5 classes
AIC	1 096 158	971 747.3	971 759.3
BIC	1 096 342	971 996.5	972 073.5
G²	183 506.9	59 084.44	59 084.44
X²	157 468	42 762	42 762
DL	14	8	2

Dans le tableau 12, nous voyons que la classe 1 correspond aux personnes ayant eu des services auprès des médecins omnipraticiens. La classe 2 correspond d'une part aux personnes ayant eu une consultation avec un psychiatre et aux personnes ayant été hospitalisées ou qui ont visité le service des urgences. Dans la classe 3, la totalité des personnes ont consulté d'autres médecins spécialistes, et finalement la classe 4, avec plus de la moitié des répondants correspond aux personnes n'ayant eu recours à aucun services médicaux. Le regroupement des classes 1 et 3 permet de se rapprocher de la classe 2 obtenue dans l'AS. L'ACL permet donc de distinguer deux sous-groupes dans la catégorie des consommateurs divers (médecin omnipraticien et autre médecin spécialiste).

Tableau 12. Caractérisation des classes				
	Classe 1 (14.1 %) n = 623	Classe 2 (11.4 %) n = 504	Classe 3 (16.4 %) n = 725	Classe 4 (58.1 %) n = 2 569
Aucun service	0	0	0	100 %
Autre médecin spécialiste	0	0	100 %	0
Hospitalisation ou urgence	0	53.4 %	0	0
Médecin omnipraticien	100 %	0	0	0
Psychiatre	0	46.6 %	0	0

La répartition post-estimation (figure 9) présente une répartition quasi-exclusive des classes entre les services médicaux, excepté la classe 2 qui comprend une combinaison des services hospitaliers et d'urgence, et des services d'un psychiatre.

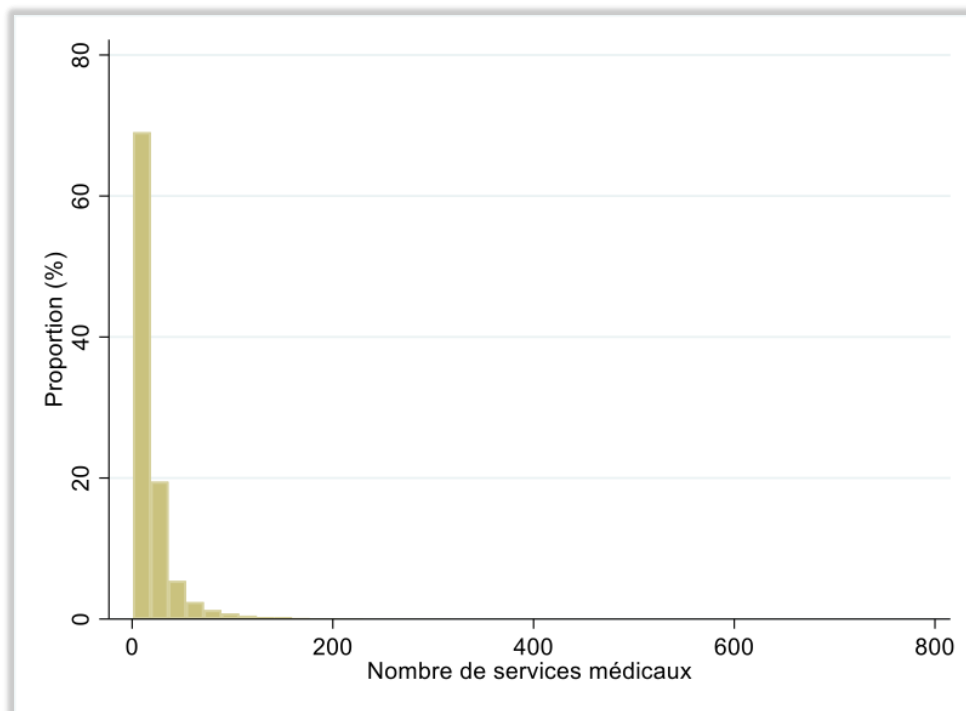
Figure 9. Histogramme des classes latentes



4.3.2. APL : Services médicaux

La figure 10 montre la répartition de l'utilisation des services médicaux. Près de 70 % des individus n'avaient eu recours à aucun service médical au cours de la période de suivi de 8 années.

Figure 10. Répartition du nombre de services médicaux

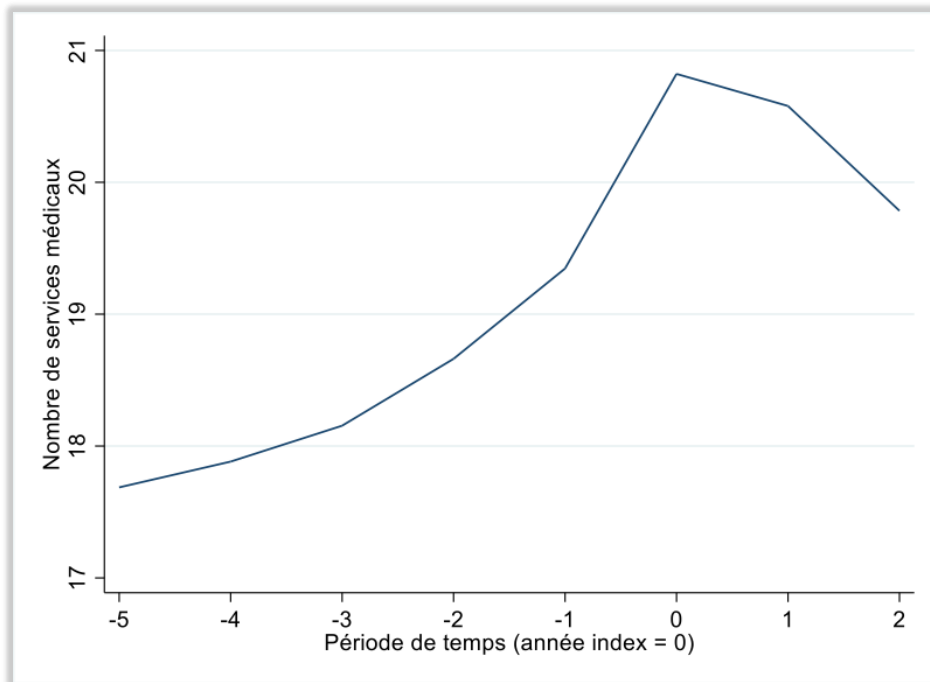


La figure 11 présente l'évolution du nombre de services médicaux reçus par les individus ayant reçu au moins un service médical avant et après la date index. En moyenne et par année, le

nombre de services médicaux a évolué à la hausse, passant de 17 à 20 services, avec un pic observé à l'année de l'enquête ESCC. Une tendance à la baisse est ensuite observée après la date index, mais l'observation est faite sur 2 années et ne permet pas de tirer de conclusions même si cela semble symétrique.

Les résultats du tableau 13 montrent que les femmes dans les profils 3 et 4 avaient une probabilité statistiquement plus grande d'avoir des services médicaux par rapport aux hommes. Dans les profils 1 et 2, c'est le résultat contraire avec les hommes ayant une probabilité plus grande d'utiliser plus de services médicaux ; le coefficient de la classe 1 n'est toutefois pas statistiquement significatif. L'âge et le revenu ont un impact positif sur la consommation de services médicaux, toutefois les individus du profil 1 ont consommé moins de services médicaux lorsque leur revenu augmentait. Le statut matrimonial semble avoir un effet mitigé sur la consommation des services médicaux. Les personnes seules (veuves, séparées, divorcées, célibataires, jamais mariées) consommaient plus de services médicaux dans les profils 1 et 2, et moins dans les profils 3 et 4 par rapport aux personnes en couple (mariées, conjoints de fait). Il y a peut-être un effet d'interaction avec le genre que nous vérifierons plus tard dans nos analyses. De plus, le fait d'avoir au moins 1 enfant de moins de 12 ans dans le ménage avait un impact très positif et significatif sur la consommation de services médicaux, du moins pour les profils 2, 3 et 4. Le niveau d'étude et la taille du ménage n'avaient dans l'ensemble pas d'effet significatif.

Figure 11. Évolution annuelle du nombre de services médicaux



Lorsque nous regardons les caractéristiques descriptives des profils, nous pouvons constater que le profil 1 correspond aux individus présentant une forte consommation de services médicaux avec une consommation annuelle moyenne de plus de 41 services (sd : 16) soit plus de 3 par mois,

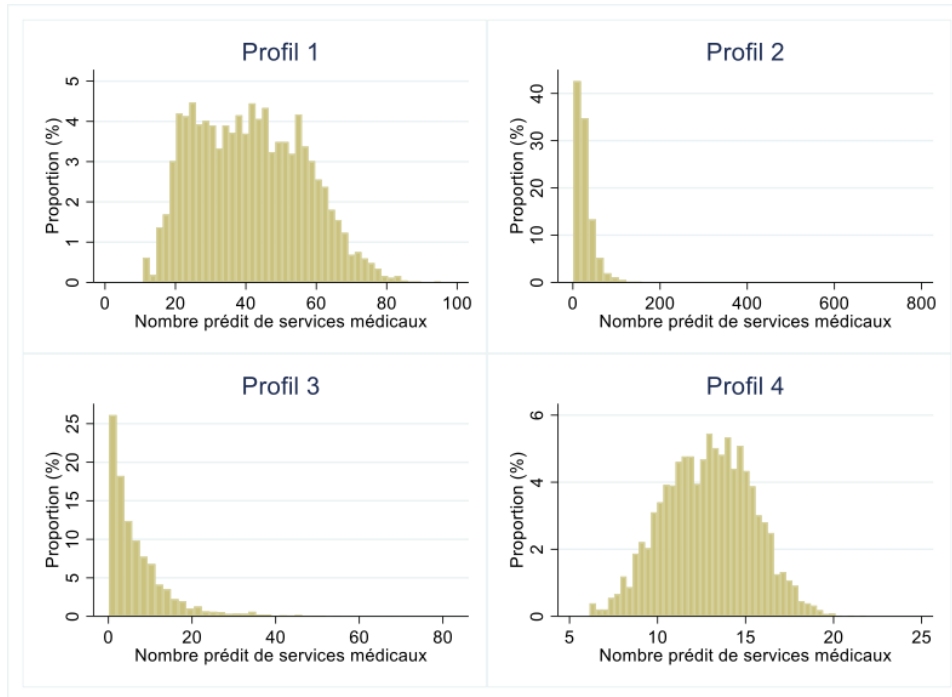
et correspond à 17,6% de l'échantillon total. Le profil 2 suit avec une consommation annuelle moyenne de 33 services (sd : 25) soit plus de 2 services par mois et représente environ 5,1% de l'échantillon. Le profil 3 est celui consommant le moins de services médicaux, à 7 par année, soit près de 2 par trimestre ; cela correspond à 16,9% de l'échantillon. Finalement le profil 4, le plus large avec plus 60% de l'échantillon, a une consommation d'un service en moyenne par mois.

Tableau 13. Résultats pour la régression du nombre de services médicaux

		Profil 1 (n = 6 225)		Profil 2 (n = 1 804)		Profil 3 (n = 5 977)		Profil 4 (n = 21 362)	
		Coeff. (s.e.)	Valeur <i>p</i>	Coeff. (s.e.)	Valeur <i>p</i>	Coeff. (s.e.)	Valeur <i>p</i>	Coeff. (s.e.)	Valeur <i>p</i>
Sexe	Homme ^β								
	Femme	-0,02 (0,04)	0,629	-0,56 (0,13)	0,000	1,39 (0,14)	0,000	0,16 (0,02)	0,000
Âge		0,01 (0,00)	0,000	0,01 (0,00)	0,125	0,06 (0,00)	0,000	0,01 (0,00)	0,000
Revenu		-0,21 (0,02)	0,000	0,04 (0,05)	0,352	0,08 (0,04)	0,032	-0,07 (0,01)	0,000
Statut matrimonial	En couple ^β								
	Veuf, séparé, divorcé	0,09 (0,05)	0,065	0,09 (0,21)	0,683	-0,32 (0,13)	0,013	0,03 (0,03)	0,368
	Célibataire, jamais marié	0,02 (0,05)	0,715	0,49 (0,12)	0,000	-0,26 (0,14)	0,063	-0,05 (0,02)	0,037
Enfant < 12 ans	Réf : Aucun enfant								
	Au moins 1 enfant	-0,30 (0,08)	0,000	0,76 (0,22)	0,001	0,45 (0,17)	0,007	0,12 (0,03)	0,000
Plus haut niveau d'éducation	Études secondaires partielles ou non complétées ^β								
	Diplôme d'études secondaires, sans études postsecondaires	0,05 (0,05)	0,376	-0,32 (0,16)	0,042	-0,03 (0,13)	0,849	-0,05 (0,03)	0,072
	Études postsecondaires partielles	-0,52 (0,13)	0,000	1,09 (0,23)	0,000	0,36 (0,17)	0,035	-0,03 (0,04)	0,381
	Diplôme postsecondaire ou grade universitaire	-0,07 (0,04)	0,109	-0,20 (0,13)	0,120	0,13 (0,11)	0,228	0,04 (0,02)	0,111
Taille du ménage		0,11 (0,03)	0,000	0,00 (0,07)	0,965	-0,09 (0,07)	0,183	0,01 (0,01)	0,545
Valeur prédite	Moyenne	41,29 (15,98)		33,11 (24,62)		7,48 (8,90)		12,63 (2,55)	
	Probabilité	17,6 %		5,1 %		16,9 %		60,4 %	
		n = 35 368 ; AIC = 267487,5 ; BIC = 267919,3							
β Catégorie de référence									

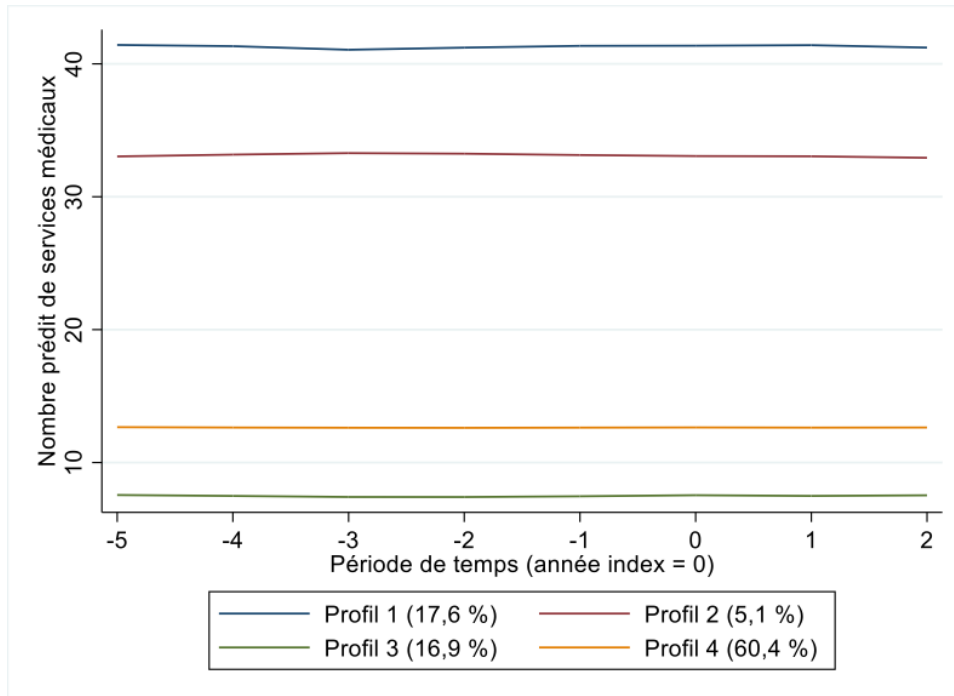
La répartition post-estimation de l'utilisation des services médicaux (figure 12) nous permet d'observer une forte concentration des services médicaux proche de 0 dans le profil 3, et dans le profil 2, un étalement à droite pour des valeurs plus grandes de consommation de services médicaux, ce qui lui donne un écart-type plus grand. Les profils 1 et 4 ont des distributions plus concentrées autour de leur moyenne, quoique le profil 1 soit plus fortement concentrée vers 0.

Figure 12. Répartition du nombre de services médicaux par profils



Un écart moyen d'au moins 20 services médicaux par année sépare les profils 1 et 2 des profils 3 et 4, ces derniers représentant plus des trois-quarts de l'échantillon total (figure 13). Nous aurions ainsi également pu retenir deux grands profils latents.

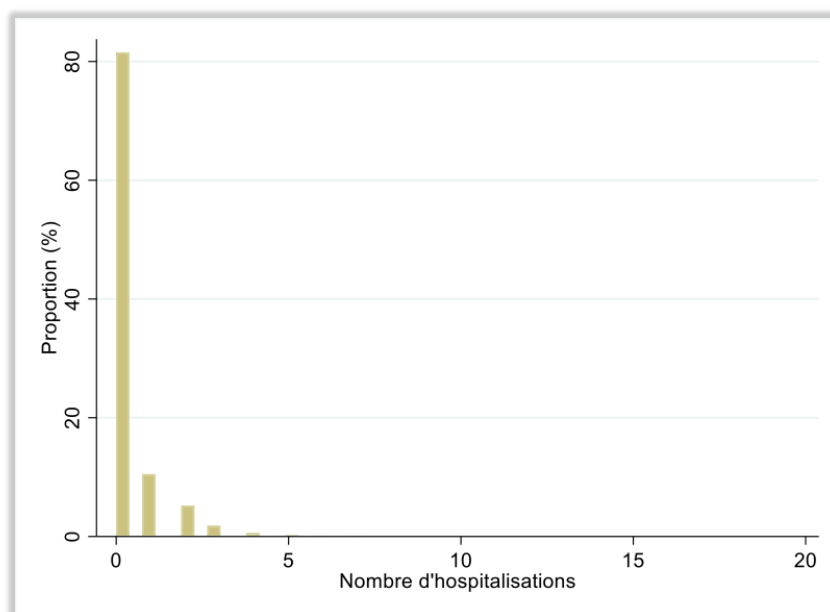
Figure 13. Évolution annuelle du nombre de services médicaux par profils



4.3.3. APL : Hospitalisations

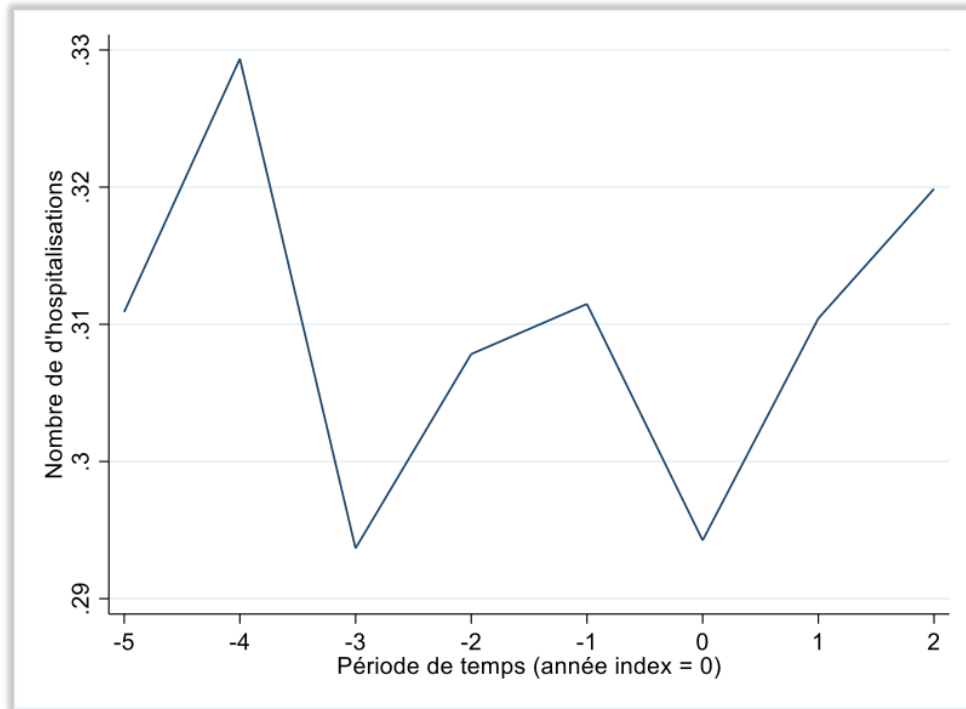
Au total, moins de 20 % des répondants ont été hospitalisés au moins 1 fois durant la période de suivi, contre plus de 80 % qui n'ont pas été hospitalisés (figure 14).

Figure 14. Répartition du nombre annuel moyen d'hospitalisations



Le nombre d'hospitalisations pour l'ensemble de l'échantillon est très faible et est en moyenne inférieure à 1 par année par individu (figure 15).

Figure 15. Répartition du nombre annuel moyen d'hospitalisations



Pour les hospitalisations, nous avons retenu deux profils. Le profil 1 correspond à celui des personnes qui ont eu quelques hospitalisations avec une moyenne d'une hospitalisation par année et représente un peu moins du tiers de l'échantillon total; et le profil 2 correspond au groupe des personnes en ayant très peu ou n'ayant jamais été hospitalisés, et représente un peu plus des deux tiers de notre échantillon. La raison étant l'absence d'utilisation de services médicaux. Les résultats de la régression pour le nombre de diagnostics semblent cohérents avec ceux des services médicaux.

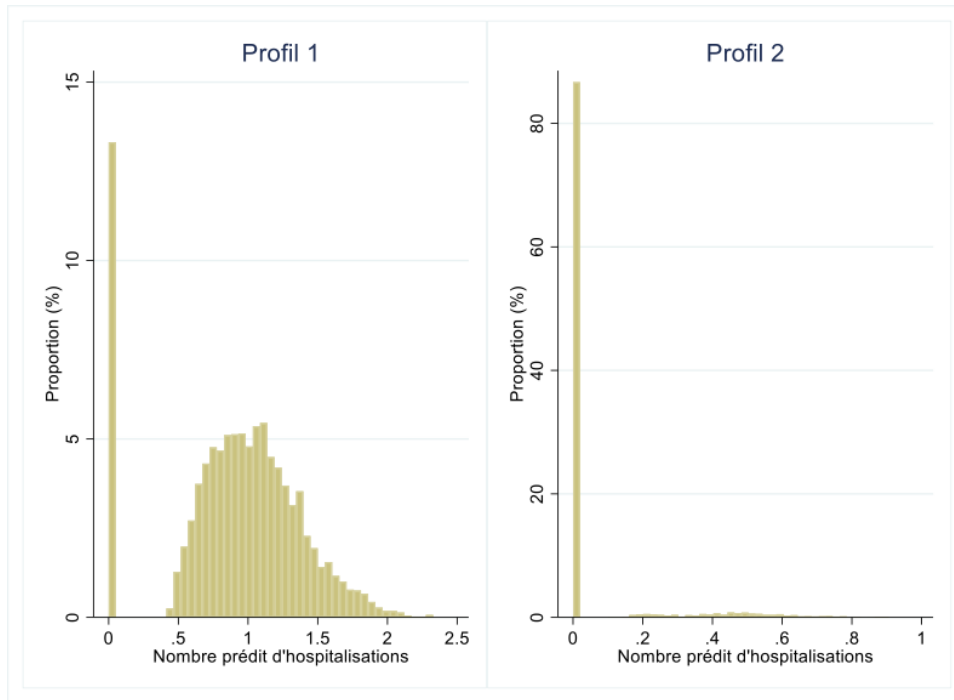
Tableau 14. Résultats pour la régression du nombre d'hospitalisations

		Profil 1 (n = 10 186)		Profil 2 (n = 25 182)	
		Coeff. (s.e.)	Valeur <i>p</i>	Coeff. (s.e.)	Valeur <i>p</i>
Sexe	Homme ^β				
	Femme	-0,14 (0,03)	0,000	0,67 (0,08)	0,000
Âge		0,01 (0,00)	0,000	-0,02 (0,00)	0,000
Revenu		-0,05 (0,01)	0,000	-0,01 (0,03)	0,675

Tableau 14. Résultats pour la régression du nombre d'hospitalisations (suite)					
		Profil 1 (n = 10 186)		Profil 2 (n = 25 182)	
		Coeff. (s.e.)	Valeur <i>p</i>	Coeff. (s.e.)	Valeur <i>p</i>
Statut matrimonial	En couple ^β				
	Veuf, séparé, divorcé	0,009 (0,03)	0,009	0,04 (0,09)	0,649
	Célibataire, jamais marié	-0,05 (0,04)	0,229	-0,08 (0,90)	0,320
Enfant < 12 ans	Aucun enfant (réf.)				
	Au moins 1 enfant	-33,64 (1e+04)	1,00	-18,60 (840)	0,982
Plus haut niveau d'éducation	Études secondaires partielles ou non complétées ^β				
	Diplôme secondaire sans études postsecondaires	-0,018 (0,04)	0,000	-0,06 (0,11)	0,591
	Études postsecondaires partielles	-0,18 (0,05)	0,000	-0,40 (0,14)	0,004
	Diplôme postsecondaire ou grade universitaire	-0,16 (0,03)	0,000	-0,16 (0,09)	0,063
Taille du ménage		0,06 (0,02)	0,005	0,07 (0,03)	0,013
Valeur prédite					
	Moyenne	0,91 (0,003)		0,06 (0,001)	
	Probabilité	28,8 %		71,2 %	
n = 35 368 ; AIC = 47 797,8 ; BIC = 47 992,58					
β Catégorie de référence					

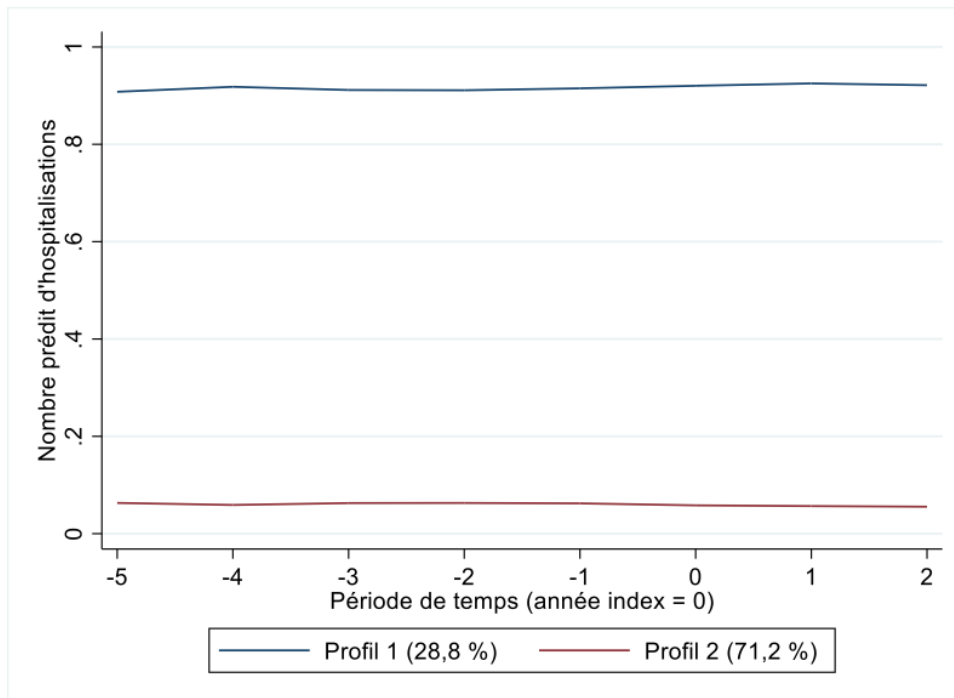
La répartition post-estimation montre une concentration des hospitalisations autour de 1 pour le profil 1, avec environ 14 % des membres n'ayant eu aucune hospitalisation (figure 16). Dans le profil 2, la distribution reste concentrée à 0 à plus de 80 %, ce qui signifie que plus de 80% des membres de ce profil n'ont eu aucune hospitalisation.

Figure 16. Répartition du nombre annuel moyen d'hospitalisations par profils



Un écart moyen d'environ 1 hospitalisation par année sépare les profils 1 et 2, ce dernier représentant plus des deux tiers de l'échantillon total (figure 17).

Figure 17. Évolution annuelle du nombre d'hospitalisations par profils



5. Interprétation et utilisation des méthodes

5.1. Méthode de l'analyse de séquences

L'AS a pour avantage de tenir compte de l'aspect temporel des observations, en étudiant la succession d'états ou d'évènements. Dans le cadre de notre étude, nous nous sommes intéressés à l'évaluation de l'utilisation des services de santé auprès d'une population ayant auto-déclaré avoir un TH.

Après la construction des séquences selon les états prioritaires retenus, il nous a été possible de distinguer trois groupes d'utilisateurs des services de santé ayant un TH.

Le premier groupe (type 1) est composé d'individus ayant eu une très faible utilisation des services de santé, avec plus de 50 % des individus n'ayant eu aucun contact avec le système de santé durant la période de suivi de 7 années. Lorsqu'ils ont fait appel à des services médicaux, la grande majorité des consultations médicales ont eu lieu auprès d'un médecin omnipraticien (environ 20 % des répondants par mois).

Le deuxième groupe (type 2) est quant à lui composé d'individus ayant une utilisation modérée des services de santé, avec environ 60 % des individus ayant eu au moins un contact avec le système de santé mensuellement. La fréquence des répondants ayant eu une consultation avec un médecin spécialiste autre que psychiatre, semble similaire ou légèrement plus élevée que celle avec un médecin omnipraticien, toutes deux autour de 20-25 %. Les hospitalisations ou consultations à l'urgence semblent plus élevées également comparativement aux utilisateurs de type 1, tandis que la fréquence des consultations avec un psychiatre semble similaire.

Le troisième groupe (type 3) quant à lui représente les répondants ayant une utilisation modérée des services de santé de manière générale, mais avec une utilisation élevée des services avec un psychiatre. En effet, lorsque nous observons la fréquence totale moyenne par mois, nous observons qu'environ 60 % des individus ont eu au moins un contact avec le système de santé. Par contre, lorsque nous regardons la distribution des types de contacts, nous observons que la plus forte proportion des consultations appartient à celles avec un psychiatre, soit entre 20 à 30 % en moyenne par mois, comparativement à moins de 5 % pour les utilisateurs de type 1 et 2.

Il est important de rappeler que dans le cadre de l'AS, il est nécessaire d'établir un ordre de priorité des états ou des évènements d'intérêt. Dans le contexte de notre étude, cela se traduit par le fait qu'un individu ne peut avoir qu'un type de consultation (incluant l'absence de consultation) par mois. Il est donc possible qu'un individu ait eu une consultation avec un médecin omnipraticien et un psychiatre dans le même mois. Par contre, l'AS ne retiendra que la consultation prioritaire, soit celle avec le psychiatre dans cette situation. De ce fait, l'AS ne reflète pas l'intensité de l'utilisation des services de santé via la quantité d'utilisation, mais plutôt par la gravité de celle-ci par la détermination de l'ordre de priorité des états, allant du plus intense (hospitalisation ou consultation à l'urgence) au moins intense (aucune consultation).

Enfin, il est aussi pertinent de rappeler que nous avons choisi une fréquence d'évaluation des états qui est mensuelle, pour nous assurer d'avoir une variété dans l'utilisation des services et

pour pouvoir converger les données vers l'identification de trajectoires typiques. Il serait possible d'effectuer le même exercice d'AS mais avec une fréquence d'évaluation plus élevée (p. ex. : hebdomadaire ou quotidienne) et de comparer les données, et ainsi ajouter de la robustesse à nos résultats.

5.2. Méthode de l'analyse de classes et de profils latents

L'analyse des classes et des profils latents est utile pour la détermination de sous-groupes inobservés à partir des caractéristiques observées dans la population. Le choix du modèle d'estimation doit toujours représenter la structure des données. L'estimation des modèles peut se faire avec ou sans covariables dans l'ACL comme dans l'APL. Dans le cas d'un modèle avec covariables, il faut s'assurer d'inclure les variables qui sont le mieux susceptibles d'expliquer la variable dépendante et de contribuer au mieux à la prédiction des classes.

Trois éléments permettent de déterminer le meilleur modèle et d'utiliser les résultats des analyses:

- i) **Les critères statistiques d'information sur le modèle** : Ce sont les statistiques post-estimation permettant de mieux informer sur l'acceptabilité d'un modèle relativement à un autre. Plusieurs autres statistiques comme le test de ratio de vraisemblance sont également utilisées pour apprécier la qualité du modèle.
- ii) **La moyenne marginale prédite des classes** : À l'issue de l'estimation du modèle économétrique, le calcul de la moyenne marginale permet de déterminer pour chaque classe la valeur moyenne de la variable dépendante de la régression. Cette moyenne permet très facilement de caractériser chaque classe au regard de la valeur de sa moyenne marginale.
- iii) **La probabilité prédite d'appartenance à une classe** : Cette statistique a pour utilité de déterminer l'importance relative de chaque classe dans l'ensemble de l'échantillon à l'étude.

Les méthodes d'ACL et d'APL conviennent pour l'analyse des données transversales (cross sectional) ou des données longitudinales. Elles permettent de déterminer à partir d'un ensemble de variables, les caractéristiques inobservées ou méconnues d'un ensemble de la population. À partir de la caractérisation des classes ou sous-groupes, il est plus aisé de pouvoir mettre en œuvre des politiques ciblées afin de mieux répondre à un problème de santé latent observé. Une analyse descriptive classique des données ne permettrait pas de mettre en évidence toutes les subtilités inobservées (latentes) dans une population.

Bien qu'elle soit très utile pour l'analyse approfondie d'un échantillon de patients, l'utilisation de la méthode (ACL ou APL) nécessite de disposer de compétences relativement avancées en statistiques et d'être apte à travailler avec des logiciels statistiques. De fait, la disposition de la base de données dans le format approprié pour l'analyse des classes latentes représente une part non négligeable du travail, surtout pour les données longitudinales. Par ailleurs la compréhension du modèle de régression et des hypothèses sous-jacentes ainsi que l'interprétation des résultats demeurent un autre défi à l'utilisation de la méthode.

L'ACL a montré une répartition de l'échantillon en 4 groupes distinctement caractérisé selon les services médicaux. Le groupe des personnes n'ayant utilisé aucun service médical reste le plus nombreux avec près de 60 % de la taille de l'échantillon.

Les résultats de l'analyse de profils ont montré que globalement notre échantillon est composé de 4 groupes d'utilisateurs des services médicaux, et de deux groupes pour les hospitalisations. Les personnes n'ayant pas utilisé des services médicaux et n'ayant pas été hospitalisé durant toute la période de suivi sont les plus nombreux, avec plus de 60 % pour les services médicaux et plus de 70% pour les hospitalisations. Cela traduit une faible utilisation des soins de santé de façon générale pour la cohorte suivie. Une observation générale pouvant être faite est que notre cohorte peut être classée en 3 grands groupes. Le premier groupe qui est celui qui contient le plus de personnes est celui des personnes n'ayant consommé aucun soin de santé (services médicaux et hospitalisations). Le deuxième groupe, moins nombreux que le premier, contient les personnes ayant une consommation modérée des soins de santé. Finalement le troisième groupe, moins nombreux, est celui des personnes ayant consommé une grande quantité de soins de santé à travers un grand nombre de services médicaux et des hospitalisations fréquentes et de durées relativement longues.

En définitive, et peu importe la méthode, le groupe « Aucun service » demeure le groupe le plus large.

5.3. Comparaison des trois méthodes d'analyse

5.3.1. Analyse de séquences vs Analyse de classes latentes et de profils latents

L'analyse de séquence (AS) modélise les successions d'événements survenus dans une période donnée, tandis que l'analyse de classes latentes (ACL) et de profils latents (APL) pourrait s'appliquer à des mesures répétées ou non. L'AS par son procédé non paramétrique, fait reposer une importante partie des choix analytiques sur les décisions du chercheur, alors que l'ACL et l'APL par leur procédé paramétrique ne laissent pas nécessairement le choix au chercheur d'intervenir dans la priorisation des choix importants pour la définition des TS.

Dans tous les cas, ces deux techniques, d'une part l'AS et d'autre part l'ACL et l'APL, permettent de mettre en évidence des groupes inobservés qui prévalent au sein d'une population.

Les résultats obtenus à l'issue de l'AS, l'ACL et l'APL semblent très proches. D'abord l'AS a permis d'établir 3 classes d'utilisateurs des services médicaux ; et les résultats de l'ACL permettent d'en arriver aux mêmes conclusions. En effet, bien que l'ACL aient permis d'établir 4 classes, le regroupement des classes 1 et 3 permet de se rapprocher de la classe 2 obtenue dans l'AS. Ainsi, nous pouvons dire que l'AS et l'ACL donnent des résultats similaires. Concernant l'APL, la comparaison est moins évidente à cause de la constitution de la base des données, la périodicité des mesures ainsi que des variables d'intérêt (nombre de services médicaux et d'hospitalisations).

5.3.2. Analyse de classes latentes vs Analyse de profils latents

Comme mentionnée plus haut, la principale différence entre l'ACL et l'APL repose d'abord et avant tout sur la nature des variables d'intérêt; l'ACL nécessitant des variables catégorielles tandis que l'APL utilise des variables continues. Ces deux méthodes utilisent des procédés paramétriques similaires et les critères de décisions sont les mêmes.

Dans notre analyse, nous n'avons pas utilisé de covariables pour l'ACL afin de la rendre plus proche de l'AS et de faciliter les comparaisons. Pour l'APL, nous avons utilisé des covariables susceptibles d'influencer la détermination des profils des individus.

Bien que l'APL ne nous permet pas d'établir une comparaison assez directe avec l'ACL, elle a le mérite de présenter des résultats sur l'intensité dans l'utilisation des services médicaux et des hospitalisations.

6. Discussion

Différents choix méthodologiques ont dû être faits afin de mener à bien les analyses. Comme tout type de projet et méthodologie de recherche, certaines limites sont présentes.

La première limite concerne la sélection des individus constituant notre cohorte de personnes atteintes de TH, reposant sur l'auto-déclaration du répondant quant à avoir reçu ou non un diagnostic de TH. Certains pourraient questionner la validité de l'auto-déclaration d'une maladie pour identifier les personnes atteintes de maladies mentales. Certaines personnes pourraient ne pas vouloir déclarer la présence d'une maladie mentale, en raison de la stigmatisation qu'elles peuvent vivre. Cependant, Stuart et al., dans leur analyse comparative de l'auto-déclaration d'une dépression à l'entrevue clinique structurée, considérée comme l'étalon d'or pour identifier les personnes atteintes de maladies mentales, ont rapporté que l'auto-déclaration avait une validité acceptable, avec une sensibilité à 61,0 % et une spécificité à 89,5 %¹¹. Comme mentionné au début de ce rapport, une alternative à l'utilisation des données auto-rapportées aurait été l'utilisation des données de diagnostic disponibles dans la base de données de la RAMQ. Par contre, l'utilisation de ces données médico-administrative pour l'identification de personnes avec des conditions d'intérêt rapporte son lot de limites, comme celle où un seul diagnostic est émis par visite médicale, même en la présence d'une personne avec plusieurs comorbidités.

La deuxième limite à notre étude concerne le regroupement des TH sous une même catégorie. En effet, la formulation de la question sur les TH de l'ESCC ne permet pas de distinguer les individus selon la maladie (ex : dépression vs trouble bipolaire). Bien qu'ils soient regroupés sous l'appellation commune des TH, ces troubles n'ont pas le même impact sur la vie des patients au quotidien et certains se distinguent par leur nature chronique (ex : trouble bipolaire), ce qui peut conséquemment influencer leur trajectoire de soins.

La dernière limite se rapporte aux données médico-administratives disponibles pour l'analyse de l'utilisation des soins et services de santé. Malheureusement, les données issues des banques de données de la RAMQ sont principalement de nature médicale, puisqu'elles incluent seulement les

services couverts par la RAMQ, et non les services payés par les usagers ou les services couverts par les assurances privées. Il est bien attendu que les patients souffrants de TH utilisent des soins de santé non couverts par le régime de santé public, comme les consultations chez les psychologues et d'autres professionnels de la santé (autre que le médecin). En outre, certains patients utilisent les services de soutien offerts par les organismes communautaires. Les services reçus dans le régime privé, ainsi que certains services couverts par le régime public ne sont pas dans les bases de données de la RAMQ (tel que les services dispensés par des travailleurs sociaux en CLSC) échappent à notre analyse en raison de leur absence dans les sources des données disponibles pour la cohorte TorSaDE.

Enfin, il est aussi pertinent de mentionner que chacune des méthodes d'analyses utilisées possède en elle-même des avantages et des limites. Le principal avantage de l'AS est sa considération pour l'interdépendance des états ou des événements à l'intérieur d'une même trajectoire de soins. Elle tient compte de la succession des états, plutôt que seulement leur nombre d'apparition. Cependant, l'AS demeure sensible aux données manquantes. En effet, si plusieurs périodes d'observations demeurent sans états ou événements, l'AS risque de ne pas pouvoir converger ou représenter adéquatement les trajectoires de soins. Il est donc nécessaire de faire une priorisation des états d'intérêts, ce qui, malheureusement, peut diminuer la précision des trajectoires.

L'ACL et l'APL sont deux méthodes qui ont l'avantage d'estimer de façon paramétrique les groupes latents, ce qui permet d'éviter les biais qui pourraient être liés aux préférences du chercheur. De plus, elles permettent de considérer plusieurs variables d'intérêt de façon simultanée afin de déterminer la classe d'appartenance. Le principal inconvénient des méthodes d'ACL et d'APL est qu'elles ne permettent d'estimer que des effets moyens. Dans le cas où il y aurait une absence d'utilisation des services médicaux, la présence d'un grand nombre de 0 pourrait empêcher l'estimation optimale des classes ou profils latents. Par ailleurs, ces méthodes ont l'inconvénient d'estimer un grand nombre de paramètres qui peuvent être difficiles à interpréter.

7. Conclusion

Le but de ce rapport est de présenter les résultats de l'analyse des TS des patients souffrants de TH au sein de la population québécoise. Plusieurs méthodes d'analyses ont été utilisées afin de mettre en évidence les soins de santé utilisés, mais aussi l'intensité de cette utilisation.

8. Fiches techniques des méthodes

	Analyse de classes/profils latent(e)s	Analyse de séquences
Approche	Modélise les mesures (répétées) d'une variable d'intérêt donné.	Modélise les séquences d'états ou d'évènements qui surviennent sur une période donnée.
Fonctionnement		
Technique statistique	Approche paramétrique ou semi-paramétrique permettant	Approche non-paramétrique

	l'hétérogénéité à l'intérieur des classes ou sous-groupes	
Type de données	Longitudinale et/ou transversales	Données longitudinales
Type de variables	Catégorielles et/ou continues	Catégorielles
Inclusion de covariables	Oui	Oui/Non : Possibilité de stratifier selon des variables ou des caractéristiques catégoriques (p. ex. : le sexe)
Logiciels statistiques disponibles	R, SAS, Stata, Mplus	R : TraMineR Stata : SADI ou SQ-Ados 2.0
Avantages		
	L'avantage de la méthode de l'ACL et de l'APL est la possibilité de considérer plusieurs variables d'intérêt de façon simultanée dans une seule estimation afin de déterminer la classe d'appartenance. Elle permet également de considérer l'hétérogénéité existante entre les groupes et les individus à travers la prise en compte d'effets fixes. L'estimation faite de façon paramétrique rend le choix du nombre de classes indépendante du chercheur.	L'avantage principal de l'AS réside dans son approche holistique, c'est-à-dire qu'elle permet de tenir compte des TS des individus dans leur globalité plutôt que de s'arrêter à des événements ou des moments particuliers. La trajectoire, en termes de durée, de moment et d'enchaînement des états, représente l'unité d'analyse de l'AS. Ainsi, elle ne considère pas les états ou événements composant une séquence comme étant indépendants, contrairement aux études transversales (une seule mesure à un moment précis).
Inconvénients		
	Le principal inconvénient est que l'ACL et l'APL ne permet d'estimer que des effets moyens. Par ailleurs, la méthode a l'inconvénient d'estimer un grand nombre de paramètres qui peuvent être difficiles à interpréter.	La validité des analyses peut être compromise lorsque la longueur des séquences n'est pas identique entre les individus. Ceci est particulièrement vrai pour les données manquantes.

9. Algorithmes utilisés

L'algorithme présenté dans cette section a été développé sous le logiciel Stata. La logique algorithmique reste toutefois applicable à tout logiciel d'analyse statistique de données. Les procédures générales liées à la sélection des individus et au calcul des coûts restent identiques pour les deux méthodes d'analyse ; seules les étapes directement reliées à l'application des méthodes LCA ou SA diffèrent.

9.1. Traitement des données et sélection des individus

Les individus ayant répondu « Oui » à la question « Êtes-vous atteint.e d'un trouble de l'humeur tel que la dépression, le trouble bipolaire, la manie ou la dysthymie? » dans l'enquête sur la santé

dans les collectivités canadiennes (ESCC) ont été identifiés et extraits de la grande base de données de l'ESCC. Les identifiants multiples ont été supprimés pour ne garder qu'un seul individu, avec sa donnée la plus récente. Nous avons ensuite éliminé les individus mineurs (âgés de moins de 18 ans au moment de l'ESCC) et ceux qui avaient complété l'ESCC en 2015 et 2016, pour être en mesure de couvrir la période de 2 années après l'ESCC pour les données médico-administratives, sachant que celles-ci étaient disponibles jusqu'à 2016. Cette base de données, qui constituait notre échantillon de base, avait ainsi 4 421 individus. Nous avons combiné cette base de données à celle des participants, contenant les données socio-économiques des répondants, ainsi qu'aux bases de données de la RAMQ pour les informations relatives aux hospitalisations, consultations à l'urgence et autres consultations médicales.

9.2. Algorithme de l'analyse de séquences et classes latentes

Voir les fichiers de syntaxe associés.

Références

1. Vanasse, A., Courteau, M. & Ethier, J.-F. The '6W' multidimensional model of care trajectories for patients with chronic ambulatory care sensitive conditions and hospital readmissions. *Public Health* 157, 53–61 (2018).
2. Unies, N. Le Conseil des droits de l'homme débat du droit à la santé, s'agissant plus particulièrement de la santé mentale, et de la solidarité internationale. <https://www.ohchr.org/fr/press-releases/2017/06/human-rights-council-holds-interactive-dialogue-physical-and-mental-health> (2017).
3. Datta, J. & Petticrew, M. Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Public Health* 13, 568 (2013).
4. mentale, A. canadienne pour la santé. Faits saillants sur la santé mentale et la maladie mentale. *Faits saillants sur la santé mentale et la maladie mentale* https://cmha.ca/fr/brochure/faits-saillants-sur-la-maladie-mentale/#_edn1 (2021).
5. Canada, C. de la santé mentale du. *La nécessité d'investir dans la santé mentale au Canada*. 1–32 <https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2017-03/la%20nesessite%20dinvestir%20dans%20la%20sante%20mentale%20au%20canada.pdf> (2017).
6. santé., O. mondiale de la. L'OMS souligne qu'il est urgent de transformer la santé mentale et les soins qui lui sont consacrés. <https://www.who.int/fr/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care> (2022).
7. Untitled. 10, (2018).
8. Nguéfacq, H. L. N. *et al.* Trajectory Modelling Techniques Useful to Epidemiological Research: A Comparative Narrative Review of Approaches. *Clinical Epidemiology* 1205–1222.

9. Untitled. *EJP* 965–976 (2017) doi:10.1002/ejp.998.

10. Gavin, van der N., Valéria, L. P., J.J.M., C., Math & J.P., van B., Gerard. An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. 43, 100323 (2020).

11. Stuart, A. L. *et al.* Comparison of self-report and structured clinical interview in the identification of depression. *Compr Psychiat* 55, 866–869 (2014).

Annexes

Annexe 1. Bases de données, variables et transformations utilisées pour analyse de séquences

Tableau A1. Résumé des bases de données, des variables et des transformations utilisées pour la construction de la base de données pour l'analyse de séquences				
Variable originale	Base de données	Variable renommée	Libellé et transformation (si applicable)	Définition
CCC_280	escs_2007_2016_cvm	escs_th	Est atteint de troubles de l'humeur (auto-reporté).	1 : Oui 0 : Non
adm_yoi	escs_2007_2016_cvm	annee_escs	Année à laquelle l'entrevue de l'ESCC a été complétée.	Année (2007 à 2016)
gen_07	escs_2007_2016_cvm	stress_percu	Stress perçu dans la vie.	1 : Pas du tout stressé 2 : Pas tellement stressé 3 : Un peu stressé 4 : Assez stressé 5 : Extrêmement stressé
gen_09	escs_2007_2016_cvm	stress_travail	Évaluation personnelle du stress au travail.	1 : Pas du tout stressé 2 : Pas tellement stressé 3 : Un peu stressé 4 : Assez stressé 5 : Extrêmement stressé
gendhdi	escs_2007_2016_cvm	sante_percue	Santé perçue.	1 : Mauvaise 2 : Passable 3 : Bonne 4 : Très bonne 5 : Excellente
gendmhi	escs_2007_2016_cvm	sante_mentale_percue	Santé mentale perçue.	1 : Mauvaise 2 : Passable 3 : Bonne 4 : Très bonne 5 : Excellente

Tableau A1. Résumé des bases de données, des variables et des transformations utilisées pour la construction de la base de données pour l'analyse de séquences (suite)

huidhsi	escs_2007_2016_cvm	hui_index	Indice de l'état de santé (HUI3).	Valeur de 0 (état équivalent à la mort) à 1 (santé parfaite).
hwtdbmi	escs_2007_2016_cvm	bmi	Indice de masse corporelle.	Kg/m ²
dhh_ms	escs_2007_2016_cvm	statut_marital	État matrimonial.	1 : Marié 2 : Union libre 3 : Veuf 4 : Séparé 5 : Divorcé 6 : Célibataire ou jamais marié
WTS_S_ajust	escs2009_2010_poids_ajustes escs2011_2012_poids_ajustes escs2013_2014_poids_ajustes escs2015_2016_poids_ajustes	ponderation	Poids accordé à chaque individu selon Statistique Canada pour ajuster pour la représentativité de la population canadienne.	
dat_interview_escs	participants_2007_2016	date_escs	Date exacte à laquelle l'entrevue de l'ESCC a été réalisée.	Date AAAA-MM-JJ
sexe_srap	participants_2007_2016	sexe	Sexe du répondant.	F : Femme H : Homme
age_interv_escs	participants_2007_2016	age	Âge du répondant à la date à laquelle l'ESCC a été complétée.	Âge en année
niveau_educ_srap	participants_2007_2016	niv_educ	Niveau de scolarité du répondant.	0 : inconnu 1 : N'a pas gradué du secondaire 2 : Diplôme d'études secondaires obtenu 3 : Diplôme professionnel ou du CÉGEP 4 : Baccalauréat universitaire 5 : Études supérieures (après baccalauréat)

Tableau A1. Résumé des bases de données, des variables et des transformations utilisées pour la construction de la base de données pour l'analyse de séquences (suite)

revenu_menage_srap	participants_2007_2016	rev_men	Revenu annuel du ménage en CAD.	1 : moins de 20 000 \$/année 2 : 20 000 à 40 000 \$/année 3 : 40 000 à 60 000 \$/année 4 : 60 000 à 80 000 \$/année 5 : Plus de 80 000 \$/année
revenu_perso_srap	participants_2007_2016	rev_perso	Revenu annuel du répondant en CAD.	1 : moins de 20 000 \$/année 2 : 20 000 à 40 000 \$/année 3 : 40 000 à 60 000 \$/année 4 : 60 000 à 80 000 \$/année 5 : Plus de 80 000 \$/année
pays_naissance_srap	participants_2007_2016	pays_naiss	Pays de naissance du répondant.	1 : Canada et Amérique du Nord 2 : Amérique centrale et du Sud, Antilles et Bermudes 3 : Europe 4 : Asie et Océanie 5 : Afrique
dat_admis_sej_hosp	sej_hosp_gen	date_debut_hosp	Date d'admission lors d'une hospitalisation.	Date AAAA-MM-JJ
dat_depar_sej_hosp	sej_hosp_gen	date_fin_hosp	Date de départ de l'utilisateur lors d'une hospitalisation.	Date AAAA-MM-JJ
dateservice	serv_medicaux1996_2016_gen	date_srv_med	Date de la consultation médicale.	
type_etab	serv_medicaux1996_2016_gen	type_service	Type de service reçu, selon les états prioritaires retenus à l'étude ; urgence.	Si [type_etab] = « 0X7 », alors [type_service] = « urgence »
smod_cod_spec	serv_medicaux1996_2016_gen	type_service	Type de service reçu, selon les états prioritaires retenus à l'étude ; consultation médicale avec un psychiatre. Le code « 28 » pour la variable originale [smod_cod_spec] équivaut à une consultation chargée par un médecin psychiatre.	Si [smod_cod_spec] = 28, alors [type_service] = « psych »

Tableau A1. Résumé des bases de données, des variables et des transformations utilisées pour la construction de la base de données pour l'analyse de séquences (suite)

smod_cod_spec	serv_medicaux1996_2016_gen	type_service	Type de service reçu, selon les états prioritaires retenus à l'étude ; consultation médicale avec médecin spécialiste autre que psychiatre. Pour la variable originale [smod_cod_spec], le code « 00 » indique que la consultation a été chargée par un médecin omnipraticien et le code « 28 » indique qu'elle a été chargée par un psychiatre.	Si [smod_cod_spec] ≠ 28 & [smod_cod_spec] ≠ 00 & [type_service] ≠ « urgence », alors [type_service] = « autre md spec »
smod_cod_spec	serv_medicaux1996_2016_gen	type_service	Type de service reçu, selon les états prioritaires retenus à l'étude ; consultation médicale avec médecin omnipraticien. Pour la variable originale [smod_cod_spec], le code « 00 » indique une consultation chargée par un médecin omnipraticien.	Si [smod_cod_spec] ≠ 28 & [smod_cod_spec] ≠ 00 & [type_service] ≠ « urgence », alors [type_service] = « autre md spec »

Annexe 2. Bases de données, variables et transformations utilisées pour analyse de classes latentes

Tableau A2. Résumé des bases de données, des variables et des transformations utilisées pour la construction de la base de données pour l'analyse de classes latentes			
Description	Variables originales	BD	Transformations
Sexe	SEXE_SRAP	participants_2007_2016	Renommer. Recoder F en 1 et M en 0 dans une nouvelle variable « sex ».
Âge	AGE_INTERV_ESCC	participants_2007_2016	Renommer simplement
Revenu	REVENU_MENAGE_SRA P	participants_2007_2016	Renommer. Recoder en 5 catégories en prenant le milieu de chacune des classes de revenu. Le milieu des classes a été utilisé comme revenu moyen l'individu. Recoder ensuite en 2 catégories, ceux qui gagnent moins ou plus de 40 000 \$.
Statut marital	DHH_MS	escs_2007_2016_cvm	Renommer et recoder. Regrouper les valeurs 1, 2 : « En couple ». Regrouper des valeurs 3, 4, 5 : « Veuf, séparé, divorcé ». Renommer la valeur 6 : « Célibataire, jamais marié »
Enfant de moins de 12 ans	DHHDL12	escs_2007_2016_cvm	Renommer. Regrouper la valeur 0 : « Aucun enfant ». Regrouper les valeurs de 1 et plus : « Au moins 1 enfant ».
Plus haut niveau d'éducation	EDUDR04	participants_2007_2016.dta	Renommer simplement.
Taille du ménage	dhhhsz	participants_2007_2016	Renommer simplement.
Nombre de services médicaux		serv_medicaux1996_2016_gen	Variable créée et calculée comme étant le nombre de fois qu'un même identifiant « NOINDIV_SRAP » apparaît dans la base des services médicaux.
Nombre de séjours hospitaliers		sej_hosp_gen.dta	Variable créée et calculée comme étant le nombre de fois qu'un même identifiant « NOINDIV_SRAP » apparaît dans la base des séjours hospitaliers.
Période de temps			Variable créée dans chacune des bases. C'est l'écart annuel par rapport à l'année de l'ESCC ou date index.

Annexe 3. Coûts totaux moyens par individu pour les analyses de séquences, de classes latentes et de profils latents.

Tableau A. Coûts totaux moyens par individu, selon le type d'utilisateur défini à partir de l'analyse de séquences.				
	Type 1 (n = 2 714)	Type 2 (n = 1 120)	Type 3 (n = 587)	Valeur p
Hospitalisations	74 347,05 \$ ± 305 664,00	185 279,20 \$ ± 512 424,50	153 057,00 \$ ± 516 841,20	2 vs 1 : < 0,001 3 vs 1 : < 0,001 3 vs 2 : 0,253
	0 \$ [0 – 3 522,83]	6 702,18 \$ [0 – 120 199,40]	0 \$ [0 – 38 363,02]	
Chirurgie d'un jour	7 725,47 \$ ± 35 737,59	9 818,80 \$ ± 33 610,94	5 564,68 \$ ± 26 303,01	2 vs 1 : 0,090 3 vs 1 : 0,513 3 vs 2 : 0,038
	0 \$ [0 – 0]	0 \$ [0 – 291,45]	0 \$ [0 – 0]	
Urgences	2 672,63 \$ ± 3 538,84	7 583,81 \$ ± 9 495,63	5 353,99 \$ ± 6 442,55	2 vs 1 : < 0,001 3 vs 1 : < 0,001 3 vs 2 : < 0,001
	1 422,40 \$ [355,60 – 3 556,00]	4 978,40 \$ [1778,00 – 9 601,20]	3 200,40 \$ [1 066,80 – 7 467,60]	
Services médicaux	3 030,44 \$ ± 2 372,30	9 148,58 \$ ± 7 474,77	8 997,31 \$ ± 6 677,50	2 vs 1 : < 0,001 3 vs 1 : < 0,001 3 vs 2 : 0,814
	2 515,45 [1 483,66 – 3 895,96]	7 022,74 [4 948,60 – 10 655,50]	7 191,72 [4 325,52 – 11 438,92]	
Services pharmaceutiques	7 108,08 \$ ± 15 672,20	19 068,85 \$ ± 28 697,93	18 275,41 \$ ± 26 670,43	2 vs 1 : < 0,001 3 vs 1 : < 0,001 3 vs 2 : 0,745
	1 603,78 \$ [0 – 7 821,55]	10 516,24 \$ [693,10 – 26 259,20]	8 842,21 \$ [350,17 – 27 464,87]	
Coût total	94 433,68 \$ ± 309 189,60	230 899,20 \$ ± 514 675,30	191 248,40 \$ ± 519 904,40	2 vs 1 : < 0,001 3 vs 1 : < 0,001 3 vs 2 : 0,129
	10 601,97 \$ [4 384,47 – 36 306,86]	54 452,60 \$ [20 972,53 – 226 549,70]	37 681,20 \$ [15 619,84 – 138 048,50]	

Comparaisons inter-groupe effectuées au moyen du test non-paramétrique Kruskal Wallis.

Tableau B. Coûts totaux moyens par individu, selon le type d'utilisateur défini à partir de l'analyse de classes latentes.

	Type 1 (n = 760)	Type 2 (n = 762)	Type 3 (n = 283)	Type 4 (n = 2 661)	Valeur p
Hospitalisations	147 441.60 ± 496 061.40	113 406.20 ± 321 204.50	129 518.90 ± 441 427.00	101 404.90 ± 388 700.10	2 vs 1 : 0.650 3 vs 1 : 0.932 4 vs 1 : 0.028
	2 024.15 [0 – 49 870.77]	0 [0 – 14 049.16]	0 [0 – 27 093.17]	0 [0 – 8 713.08]	3 vs 2 : 0.949 4 vs 2 : 0.886 4 vs 3 : 0.729
Chirurgie d'un jour	7 852.09 ± 29 201.14	7 995.27 ± 49 292.00	4 376.34 ± 21 680.67	7 857.05 ± 30 854.59	Tous non significatives
	0 [0 – 0]	0 [0 – 0]	0 [0 – 0]	0 [0 – 0]	
Urgences	5 438.81 ± 7 826.56	4 319.93 ± 6 338.27	8 098.12 ± 10 553.40	3 584.20 ± 5 125.11	2 vs 1 : 0.003 3 vs 1 : < 0.001 4 vs 1 : < 0.001
	2 844.80 [1 066.80 – 6 756.40]	2 133.60 [711.20 – 5 689.60]	4 622.80 [1 778.00 – 10 312.40]	2 133.60 [355.60 – 4 622.80]	3 vs 2 : < 0.001 4 vs 2 : 0.022 4 vs 3 : < 0.001
Services médicaux	7 694.03 ± 7 730.50	5 960.96 ± 5 533.60	7 937.04 ± 7 908.67	4 311.81 ± 4 339.21	2 vs 1 : < 0.001 3 vs 1 : 0.934 4 vs 1 : < 0.001
	5 535.97 [3 305.95 – 8 855.97]	4 255.92 [2 673.53 – 7 220.62]	5 795.19 [3 182.75 – 9 727.63]	3 076.17 [1 674.68 – 5 339.48]	3 vs 2 : < 0.001 4 vs 2 : < 0.001 4 vs 3 : < 0.001
Services pharmaceutiques	18 005.09 ± 30 510.29	12 481.10 ± 18 551.58	18 594.66 ± 26 572.47	8 927.51 ± 18 897.77	2 vs 1 : < 0.001 3 vs 1 : 0.983 4 vs 1 : < 0.001
	8 193.95 [152.12 – 24 141.21]	4 329.97 [0 – 18 599.31]	8 880.30 [1 372.91 – 26 763.76]	1 910.17 [0 – 9 664.96]	3 vs 2 : 0.001 4 vs 2 : < 0.001 4 vs 3 : < 0.001
Coût total	186 431.60 ± 500 719.90	144 163.40 ± 328 134.80	168 525.00 ± 444 858.20	126 085.50 ± 392 169.40	2 vs 1 : 0.117 3 vs 1 : 0.934 4 vs 1 : 0.002
	35 855.21 [11 621.93 – 153 817.90]	21 627.08 [8 476.62 – 104 569.40]	35 495.99 [13 636.93 – 115 856.10]	14 816.44 [5 153.98 – 77 597.43]	3 vs 2 : 0.851 4 vs 2 : 0.700 4 vs 3 : 0.411

Comparaisons inter-groupe effectuées au moyen du test non-paramétrique Kruskal Wallis.

Tableau C. Coûts totaux moyens par individu, selon le type d'utilisateur défini à partir de l'analyse de profils latentes.

	Type 1 (n = 1 760)	Type 2 (n = 2 661)	Valeur <i>p</i>
Hospitalisations	130 282.20 ± 421 095.80 0 [0 – 27 801.06]	101 404.90 ± 388 700.10 0 [0 – 8 713.08]	< 0.001
Chirurgie d'un jour	7 444.07 ± 38 521.62 0 [0 – 0]	7 857.05 ± 30 854.59 0 [0 – 0]	0.0226
Urgences	5 314.00 ± 7 765.40 2 844.80 [711.20 – 6 756.40]	3 584.20 ± 5 125.11 2 133.60 [355.60 – 4 622.80]	< 0.001
Services médicaux	6 976.55 ± 6 946.60 4 984.53 [3 010.40 – 8 283.77]	4 311.81 ± 4 339.21 3 076.17 [1 674.68 – 5 339.48]	< 0.001
Services pharmaceutiques	15 693.18 ± 25 565.97 6 449.89 [70.20 – 22 574.62]	8 927.51 ± 18 897.77 1 910.17 [0 – 9 664.96]	< 0.001
Coût total	165 709.90 ± 426 402.30 29 609.07 [10 325.05 – 133 030.90]	126 085.50 ± 392 169.40 14 816.44 [5 153.98 – 77 597.43]	< 0.001
<i>Comparaisons inter-groupe effectuées au moyen du test U de Mann-Whitney.</i>			

Annexe 4. Répartition de l'échantillon selon les classes et le type de comorbidités.

Type d'analyse	Comorbidité physique			Comorbidité mentale	
	0 (N = 1080)	1 – 3 (N = 2552)	4 et + (N = 789)	0 (N = 13)	1 – 3 (N = 4408)
Analyse de séquence	Type 1 (N = 2714)	28,89%	59,61%	11,61%	0,29% 99,71%
	Type 2 (N = 1120)	10,09%	54,02%	35,89%	0,36% 99,64%
	Type 3 (N = 587)	31,18%	56,56%	12,27%	0,17% 99,83%
	p-value	0,0000			0,795
Analyse de classe latente	Classe 1 (N = 760)	16,45%	54,61%	28,95%	0,26% 99,74%
	Classe 2 (N = 762)	18,24%	59,19%	22,57%	0,13% 99,87%
	Classe 3 (N = 238)	28,99%	55,04%	15,97%	0,84% 99,16%
	Classe 4 (N = 2661)	28,07%	58,44%	13,49%	0,30% 99,70%
p-value	0,0000			0,371	
Analyse de profil latent	Profil 1 (N = 1760)	18,92%	56,65%	24,43%	0,28% 99,72%

Profil 2 (N = 2661)	28,07%	58,44%	13,49%		0,30%	99,70%
p-value	0,0000				0,921	

Comorbidité physique: douleur articulaire, asthme, apnée du sommeil, fibromyalgie, arthrite, ostéoporose, maux de dos, hypertension, cholestérol, migraine, MPOC, diabète, maladie cardiaque, cancer, ulcère d'estomac, maladie vasculaire cérébrale, incontinence urinaire, allergie, trouble intestinal, scoliose, fatigue chronique, sensibilité chimique, bronchite, emphysème, BPCO

Comorbidité mentale: démence, anxiété

Rapport méthodologique de l'analyse des trajectoires
de soins pour les troubles de l'humeur

Belan, M.; Kouakou, C.;
Laberge, M.; Poder, T., 2022

Financé par

