

Projets en analyse de données longitudinales de l'UE STA302

Master II Biostatistique 2021-2022

8/11/2021

Déroulé des projets

Le projet de l'UE STA302 sera réalisé par groupes de 3 personnes autour de 5 sujets décrits précisément dans la suite du document.

Les analyses pourront être réalisées sous SAS ou R (ou les deux).

Dans tous les cas, chaque groupe devra fournir un rapport de 5 pages (aucun dépassement autorisé) structuré de manière très classique :

- introduction:
 - une brève introduction (je ne demande pas une revue de la littérature - il s'agit d'un travail statistique avant tout)
 - un objectif clairement énoncé de l'analyse statistique
- méthodes:
 - une brève description des données
 - une justification de l'analyse statistique envisagée
 - la description des modèles (équation à donner précisément pour au moins 1 modèle)
 - le plan d'analyse
- résultats:
 - la description des données incluant des informations sur le suivi longitudinal (e.g. nombre de mesures par sujet)
 - les résultats de l'analyse statistique
- conclusion:
 - une petite discussion des résultats et des analyses réalisées en insistant sur les limites et ce qui pourrait/devrait être fait. Je n'attends pas une comparaison des résultats épidémiologiques avec la littérature mais une discussion sur l'analyse statistique.

Un maximum de 5 figures/tableaux/schémas pourra être inclus en fin de rapport et ne sera pas comptabilisé dans les 5 pages (le nombre de tableaux ne pourra pas dépasser 2). Chaque figure/tableau/schéma inclus devra être précisément décrit dans le texte. Ne pas oublier que les graphiques sont très utiles pour décrire les données complexes ET visualiser les résultats.

La présentation des résultats est laissée au choix des étudiants. Une interprétation rigoureuse des analyses est attendue.

J'insiste sur le fait que la méthode et les résultats doivent être extrêmement bien décrits et structurés, et que l'intro doit rester brève. De plus, j'attends des interprétations chiffrées des estimations en graphe/tableau mais aussi dans le texte. Je ne veux pas uniquement des "c'est significatif", "ce n'est pas significatif". Le lecteur doit bien comprendre les résultats à la lecture du texte.

Ma grille d'évaluation est la suivante:

- énoncé objectif clair
- modèle choisi adapté pour répondre à l'objectif, correctement justifié et correctement décrit
- stratégie analyse pertinente et bien décrite
- description des données assez informative
- description des résultats correcte et appropriée
- interprétation des estimations correctes.
- graphes et tableaux correctement décrits et appropriés
- recul sur les résultats, sur les analyses et leurs limites
- conclusion
- **qualité de la rédaction et respect des consignes**

Le rapport sera à envoyer par mail en pdf pour le **dimanche 5 décembre 2021**. Un script commenté et éditable doit permettre de reproduire 100% des analyses. Il sera envoyé sous format éditable (.rmd et/ou .sas).

Chaque groupe devra aussi préparer pour le **jeudi 16 décembre (14h)** une présentation orale de son projet en 10 mn détaillant notamment l'approche statistique et les résultats. Les étudiants devront se partager le temps de parole. S'en suivra une discussion des résultats, limites et difficultés du projet avec les autres groupes. **Chaque groupe devra poser des questions aux autres groupes pendant la discussion.**

Le projet sera principalement noté sur la base du rapport écrit et servira de contrôle continu pour l'UE STA 302. Les présentations orales ont un double objectif:

- préparer à la restitution synthétique d'analyses statistiques, comme cela est demandé à tout statisticien en poste, et à plus court terme en soutenance de stage;
- permettre de revoir et discuter les notions essentielles dans l'analyse des données longitudinales avant l'examen final. **Je vous invite à préparer vos questions en amont.**

Je propose une séance d'une heure par semaine pour répondre aux questions des groupes:

- mardi 16/11 à 17h - même salle que le cours d'avant (ED29).
- lundi 22/11 à 17h - salle info et zoom pour les apprentis
- lundi 29/11 à 17h - salle info

Sujet 1: Evaluation de l'effet de primo-passation de tests neuropsychologiques dans la cohorte Memento

La cohorte Memento est une étude multicentrique prospective française incluant plus de 2300 participants suivis dans des Centres Mémoire de Ressources et de Recherche (CMRR) et suivis pendant une durée de 5 ans maximum [1]. Son objectif principal est d'étudier l'évolution de signes précliniques précoces de maladie d'Alzheimer et de maladies apparentées et évaluer la valeur pronostique d'une série de marqueurs (neuropsychologiques, de lésions vasculaires, psycho-comportementaux, socio-économiques, génétiques, biologiques, de neuro-imagerie) sur la progression de signes précoces vers la démence clinique puis vers le décès. A leur inclusion dans la cohorte, et ensuite annuellement, les participants réalisent une batterie de tests neuropsychologiques (TNP). Ces TNP mesurent différents domaines cognitifs. Une des limites de ces tests est leur « effet d'apprentissage », impliquant que les scores aux tests sont le reflet des capacités cognitives des sujets, mais également de leur capacité à se souvenir des réalisations précédentes. Différentes stratégies ont été proposées pour en tenir compte dans les analyses statistiques [2]. Cet effet d'apprentissage se traduit le plus souvent par un effet de primo-passation, c'est à dire uniquement une différence entre la première passation et les passations suivantes.

L'hypothèse de ce travail est que l'effet de primo-passation diffère d'un individu à l'autre. L'objectif est d'identifier l'effet individuel de primo-passation pour 3 tests neuropsychologiques (le MMSE, le TMT-A et les fluences verbales Animaux) et d'évaluer les caractéristiques associées à cet effet.

Les analyses seront réalisées à distance en utilisant le portail sécurisé DPUK (et utilisation de R ou SAS).

Bibliographie

1. Dufouil C, Dubois B, Vellas B, Pasquier F, Blanc F, Hugon J, et al. Cognitive and imaging markers in non-demented subjects attending a memory clinic: study design and baseline findings of the MEMENTO cohort. *Alzheimers Res Ther.* 29 août 2017;9(1):67.
2. Vivot A, Power MC, Glymour MM, Mayeda ER, Benitez A, Spiro A, et al. Jump, Hop, or Skip: Modeling Practice Effects in Studies of Determinants of Cognitive Change in Older Adults. *Am J Epidemiol.* 15 févr 2016;183(4):302-14.

Sujet 2: Comparaison de l'évolution des troubles anxieux et dépressifs suite à un diagnostic de cancer du sein ou de mélanome

La cohorte ELCCA (Etude Longitudinale des Changements Comportementaux, psychologiques et sociologiques après un Cancer - voir plus bas) a pour objectif d'étudier la qualité de vie liée à la santé, la situation socioéconomique et les changements psychologiques au cours du temps suite à un diagnostic de cancer du sein ou de mélanome. Les premières études sur la cohorte se sont intéressées aux trajectoires de qualité de vie [1] et de développement post-traumatique des patients [2]. Or, un diagnostic de cancer peut également être lié à des symptômes anxieux ou dépressifs dont les trajectoires pourraient être différentes selon le type de cancer. A même stade, le pronostic de chacun des cancers ainsi que leurs traitements (type et durée) sont différents, et le soutien social perçu peut différer (par exemple liée à une plus forte médiatisation du cancer du sein).

L'objectif de ce travail est d'étudier les évolutions d'anxiété et de dépression dans les deux ans suivant un diagnostic de cancer, et comparer les évolutions suivant le type de cancer.

Description de ELCCA

Les patientes atteintes d'un cancer du sein ont été recrutées à l'Institut de Cancérologie de l'Ouest (ICO) de Nantes dans les services de chirurgie et d'oncologie tandis que les patients atteints d'un mélanome ont été recrutés au CHU de Nantes dans le service d'onco-dermatologie. Les patients étaient éligibles s'ils avaient plus de 18 ans et moins de 65 ans et qu'ils avaient reçu un diagnostic de premier cancer depuis moins d'un mois. Le stade tumoral du cancer devait être de I, II voire III sans métastase à distance. Les patients souffrant de troubles mentaux ou incapables de remplir les questionnaires sans assistance ont été exclus, ainsi que les patients atteints de pathologie lourde autre que le cancer, de pathologie addictive, de syndrome douloureux chronique avec un traitement antalgique moins de 3 mois avant le diagnostic du cancer ou les patients sous antidépresseurs depuis plus de 3 mois. Après avoir fourni un consentement écrit et éclairé, les patients ont été invités à rapporter certaines informations sociodémographiques et cliniques et à remplir une série de questionnaires à 1 mois, 6 mois, 1 an et 2 ans post-diagnostic de cancer. L'anxiété et la symptomatologie dépressive ont été mesurées par le questionnaire d'anxiété et de dépression HADS [3]. Il s'agit d'un questionnaire de 14 items (échelle de Likert à 4 points) structuré en 2 dimensions (anxiété : 7 items et dépression : 7 items). Plus le score d'une dimension est élevé, plus la symptomatologie anxieuse ou dépressive est élevée. Des seuils ont été identifiés [3]: un score inférieur ou égal à 7 signifie qu'il y a absence d'état anxieux ou dépressif, un score entre 8 et 10 correspond à un état anxieux ou dépressif douteux et un score supérieur ou égal à 11 représente un état anxieux ou dépressif certain. Ce questionnaire est très souvent utilisé en cancérologie et une étude a permis de montrer sa validité auprès d'une population française de sujets atteints de cancer [4]. Les informations sociodémographiques et cliniques ont été auto-déclarées. Elles comprenaient l'âge, le sexe, la vie en couple, la présence d'enfants, le niveau de vie perçu, le revenu mensuel du foyer, l'estimation du niveau financier, la situation professionnelle ainsi que les caractéristiques du traitement.

Bibliographie

1. Bourdon M, Blanchin M, Tessier P, Campone M, Quéreux G, Dravet F, et al. Changes in quality of life after a diagnosis of cancer: a 2-year study comparing breast cancer and melanoma patients. *Qual Life Res.* 2016;25(8):1969-79.
2. Bourdon M, Blanchin M, Campone M, Quéreux G, Dravet F, Sébille V, et al. A comparison of posttraumatic growth changes in breast cancer and melanoma. *Health Psychol.* 2019;38(10):878-87.
3. Zigmond A. S., Snaith R. P. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica.* juin 1983;67(6):361-70.
4. Razavi D, Delvaux N, Farvacques C, Robaye E. Validation de la version française du H.A.D.S. dans une population de patients cancéreux hospitalisés. *Revue de psychologie appliquée (Paris).* 1989;39(4):295-307.

Sujet 3: Trajectoires des fonctions motrices et dysautonomiques dans l'atrophie multi-systématisée

L'atrophie Multi-Systématisée (AMS) est une maladie neurodégénérative rare qui se caractérise par la présence variable d'un syndrome parkinsonien, d'un syndrome cérébelleux et d'une dysautonomie. Le diagnostic repose sur des critères cliniques dont la confirmation nécessite une évaluation anatomo-pathologique post-mortem du cerveau. Cette maladie, qui évolue rapidement avec une survie médiane inférieure à 10 ans [1,2], reste très méconnue à ce jour et il n'existe aucun traitement pour apporter un soulagement symptomatique palpable ou pour ralentir la progression inéluctable.

Depuis 2007, les hôpitaux de Bordeaux et de Toulouse sont devenus centre national de référence maladie rare (CRMR) pour l'AMS et ont pu constituer une cohorte unique dans le monde de patients suivis à partir de leur première venue au centre de référence. Cette cohorte a fait l'objet d'une déclaration CNIL, les patients signant un consentement de participation en conséquence. La cohorte AMS française regroupe actuellement plus de 650 patients avec un recul de 11 ans pour les plus longs suivis. Les patients sont suivis annuellement et de manière standardisée dans les deux centres. Au cours de ce suivi annuel, sont recueillies plusieurs échelles dont l'UMSARS [3]. L'UMSARS est conçue pour évaluer les différentes sphères touchées dans le cadre de l'AMS. Elle est constituée de quatre sous domaines : le domaine fonctionnel UMSARS I, la fonction motrice UMSARS II, le handicap UMSARS IV et la tension artérielle mesurée couchée et debout sur 10 minutes UMSARS III.

La fonction motrice et la fonction dysautonomique sont deux marqueurs majeurs de la progression d'AMS [1]. L'objectif de ce travail est de décrire les trajectoires des deux fonctions et établir leurs déterminants, tout en prenant en compte leur corrélation au cours du temps (dans une analyse bivariée).

La fonction motrice sera approchée par la somme des scores à l'UMSARS I et à l'UMSARS II. La fonction dysautonomique sera approchée par la différence de pression artérielle systolique entre la position couchée et la position debout.

Bibliographie

1. Foubert-Samier A, Pavy-Le Traon A, Guillet F, Le-Goff M, Helmer C, Tison F, et al. Disease progression and prognostic factors in multiple system atrophy: A prospective cohort study. *Neurobiol Dis.* 2020;139:104813.
2. Wenning GK, Geser F, Krismer F, Seppi K, Duerr S, Boesch S, et al. The natural history of multiple system atrophy: a prospective European cohort study. *Lancet Neurol.* mars 2013;12(3):264-74.
3. Wenning GK, Tison F, Seppi K, Sampaio C, Diem A, Yekhelef F, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). *Mov Disord.* déc 2004;19(12):1391-402.

Sujet 4: Association entre l'épaisseur des fibres nerveuses rétiniennes et le vieillissement cérébral

L'épaisseur des couches rétiniennes, notamment celle des fibres nerveuses rétiniennes (Retinal nerve Fiber Layer, RNFL), correspondent aux axones des cellules nerveuses rétiniennes qui cheminent dans l'oeil et convergent vers la pupille pour former le nerf optique. Plusieurs travaux antérieurs ont mis en évidence des liens entre l'épaisseur de la RNLF et les fonctions cognitives ou les structures cérébrales [1,2], avec l'hypothèse que ces structures rétiniennes, facilement observables par un examen rétinien par tomographie par cohérence optique (ou Optical Coherence Tomography, OCT), examen simple, rapide et non invasif, pourraient être un reflet des structures cérébrales.

L'objectif de ce travail est d'analyser les liens entre l'épaisseur de la RNFL et le déclin cognitif ultérieur mesuré par des tests cognitifs sur une période de 7 ans. La sortie d'étude possiblement informative pourra être prise en compte en analyse de sensibilité.

Données

L'étude Aliénor est une étude de cohorte ophtalmologique visant à étudier les maladies oculaires liées à l'âge. Les participants de cette étude sont issus de l'étude des 3 Cités (3C) dont l'objectif était d'estimer l'association entre les maladies vasculaires et le risque de déclin cognitif et de démence. En 1999-2001, 2104 participants ont été inclus dans la cohorte 3C à Bordeaux et suivis ensuite régulièrement tous les 2/3 ans jusqu'à 17 ans, avec des évaluations régulières de leurs fonctions cognitives. Au suivi à 7 ans, 963 participants de 3C ont accepté de participer à l'étude Aliénor, qui combine ainsi des données ophtalmologiques et des données cognitives permettant d'analyser les liens entre vieillissement oculaire et cérébral. Au cours de leur suivi, en 2009-2010, les participants Aliénor ont bénéficié d'un examen OCT permettant de mesurer l'épaisseur des fibres nerveuses rétiniennes (RNFL).

Références

1. Méndez-Gómez JL, Pelletier A, Rougier M-B, Korobelnik J-F, Schweitzer C, Delyfer M-N, et al. Association of Retinal Nerve Fiber Layer Thickness With Brain Alterations in the Visual and Limbic Networks in Elderly Adults Without Dementia. JAMA Network Open [Internet]. 9 nov 2018 [cité 9 nov 2021];1(7):e184406. Disponible sur: <https://doi.org/10.1001/jamanetworkopen.2018.4406>
2. Méndez-Gómez JL, Rougier M-B, Tellouck L, Korobelnik J-F, Schweitzer C, Delyfer M-N, et al. Peripapillary Retinal Nerve Fiber Layer Thickness and the Evolution of Cognitive Performance in an Elderly Population. Front Neurol. 2017;8:93.

Sujet 5: Association entre score de santé cardio-vasculaire optimale et évolution de la symptomatologie dépressive chez les personnes âgées.

Un score de Santé Cardio-Vasculaire optimale a été très récemment associé à un risque réduit de démence et un déclin cognitif moindre chez les personnes âgées [1]. L'objectif de ce projet est d'évaluer si ce score de santé Cardio-Vasculaire optimale est aussi associé à l'évolution de la symptomatologie dépressive chez les personnes âgées. Dans un premier temps, il s'agira d'évaluer l'impact de la santé cardio-vasculaire optimale sur la trajectoire de la symptomatologie dépressive telle que mesurée par la CES-D (la CES-D sera considérée comme une variable quantitative). Dans un deuxième temps, on pourra prendre en compte dans l'analyse l'association entre trajectoire de CES-D et la survenue de démence.

Données

Le travail sera réalisé à partir des données de l'étude des 3 Cités (3C), une cohorte populationnelle multicentrique (Bordeaux, Dijon, Montpellier) dont l'objectif était d'étudier l'impact des facteurs vasculaires sur le risque de démence dans la population âgée de 65 ans et plus. Est mis à la disposition du projet un extrait de la cohorte 3C sur les trois villes. La CES-D a été mesurée à l'inclusion puis aux suivis 1,2,4,5 et 6 ayant eu lieu environ 2 ans, 4 ans, 7 ans, 10 ans et 12 ans après l'inclusion. L'analyse pourra être ajustée sur des facteurs de confusion potentiels.

Bibliographie

1. Samieri C, Perier MC, Gaye B, Proust-Lima C, Helmer C, Dartigues JF, et al. Association of Cardiovascular Health Level in Older Age With Cognitive Decline and Incident Dementia. JAMA. 21 août 2018;320(7):657-64.