

**ComPaRe**

**Trajectoires d'évolution du Covid long**

**Responsable principal du projet**

Nom	Tran
Prénom	Viet Thi
Affiliation(s)	Université Paris Cité

**Comité scientifique**

<b>Nom</b>	<b>Grade/Titre</b>	<b>Affiliation(s)</b>
Servier Clémence	Interne en santé publique	Assistance Publique Hôpitaux de Paris
Raphael Porcher	Prof. Biostatistiques	Université Paris Cité
Philippe Ravaud	Prof. Epidémiologie	Université Paris Cité

## 1. Contexte spécifique

Le COVID long est un terme créé par les patients eux-mêmes pour désigner l'ensemble des symptômes et problèmes de santé survenant après une infection par le SARS-CoV-2 <sup>1</sup>. Selon l'Office National des Statistiques du Royaume Uni, 1.2 millions de personnes rapportent des symptômes persistants après une COVID-19, soit environ 10% des personnes infectées par le SARS-CoV2 et 2% de la population totale <sup>2</sup>.

Le COVID long est une entité hétérogène dont les symptômes sont causés par de multiples mécanismes intriqués : 1) séquelles de la maladie aiguë et/ou de sa prise en charge ; 2) déconditionnement et/ou stress post traumatique ; et 3) syndrome (ou maladie) post COVID-19 (en anglais « *post COVID-19 disease* »), définie par l'OMS comme la persistance des symptômes de la maladie, survenant en général dans les 3 premiers mois et durant au moins 2 mois, et ne pouvant être expliqués par une autre cause <sup>3</sup>. Ceci explique l'hétérogénéité des tableaux cliniques, des lésions organiques et des anomalies biologiques identifiées chez les malades <sup>4-6</sup>.

Aujourd'hui, la plupart des recherches ont décrit 1) les séquelles et conséquences à long terme après une infection par le SARS-CoV2 par le suivi de cohortes de patients initialement hospitalisés <sup>6-8</sup> ; ou 2) les symptômes rapportés par les patients rapportant un COVID long <sup>5,9</sup>. Plusieurs travaux ont également cherché à grouper les patients en fonction des symptômes rapportés en utilisant des méthodes de « *clustering* » et ont, par exemple, opposé les patients rapportant exclusivement de la fatigue, des maux de tête et des symptômes respiratoires et ORL (dyspnée, toux sèche, anosmie) et ceux rapportant davantage de symptômes, incluant de la fièvre ou des symptômes digestifs <sup>10</sup>. Cependant, ces études s'intéressaient au tableau clinique des patients à un moment donné sans inférer sur leurs trajectoires d'évolution au cours du temps.

Dans un précédent travail, nous avons décrit l'évolution de la prévalence de 53 symptômes du COVID au cours du temps dans la cohorte ComPaRe Covid Long. Ces résultats, au niveau populationnel, ont permis d'identifier 1) des symptômes (e.g., toux ou anosmie) dont la prévalence décroissait progressivement au cours du temps (environ 50% des 53 symptômes du COVID long); 2) des symptômes (e.g., troubles de la concentration) dont la prévalence restait stable au cours du temps (30% des 53 symptômes du COVID long); et 3) des symptômes (e.g., perte de cheveux) dont la prévalence augmentait au cours du temps (20% des 53 symptômes du COVID long) <sup>11</sup>. L'évolution différente, au niveau populationnel, des symptômes de la maladie pourrait être expliquée par l'hypothèse qu'elle est causée par de multiples mécanismes intriqués. Ceci suggère la faisabilité d'identifier des trajectoires d'évolution des patients, au niveau individuel, en fonction de leur symptomatologies et caractéristiques initiales.

La connaissance des trajectoires d'évolution des patients permettra d'individualiser la prise en charge de la maladie, pour une médecine de précision. En effet, à ce jour, la prise en charge des patients est la même pour tous (« *one size fits all* ») quelle que soit leurs symptômes

L'identification des différentes trajectoires d'évolution possible des patients n'est utile que s'il est possible de pouvoir prédire, à partir d'information collectées chez un patient, à un moment donné, l'appartenance d'un patient à l'une de ces trajectoires.

## 2. ComPaRe : Communauté de Patients pour la Recherche

Le présent projet sera niché dans la *Communauté de Patients pour la Recherche (ComPaRe)* ([www.compare.aphp.fr](http://www.compare.aphp.fr)). ComPaRe est un projet porté par l'Assistance Publique – Hôpitaux de Paris et l'Université de Paris visant à accélérer la recherche sur les maladies chroniques. ComPaRe s'appuie sur une e-cohorte de patients adultes, souffrant d'une maladie chronique. L'objectif est de recruter 100 000 patients en 10 ans et de les suivre pendant 10 ans.

Dans ComPaRe, les participants sont suivis via des questionnaires en ligne permettant de collecter des critères de jugement rapportés par les patients (en anglais « Patient Reported Outcome Measures » [PROMS]) ou des expériences rapportées par les patients (Patient Reported Experience Measures [PREMS]).

Les données rapportées directement par les patients ont pour objectif d'être enrichies par d'autres sources de données telles que :

- Données provenant de bases médico-administratives (e.g. SNDS)
- Données hospitalières (entrepôts de données hospitaliers)
- Données provenant des outils connectés des patients (mesure de l'activité physique, du sommeil, glycémie)

### 2.1. ComPaRe : une plateforme de recherche

ComPaRe est une infrastructure de recherche permettant à toute équipe de recherche publique (hospitalière, universitaire ou institutionnelle) d'utiliser des données collectées dans ComPaRe ou de demander la collecte de nouvelles données (e.g., en ajoutant de nouveaux questionnaires en ligne) pour répondre à des questions de recherche sur les maladies chroniques.

Le modèle permet d'accélérer la recherche dans les maladies chroniques en:

- *Simplifiant le recrutement des participants dans la recherche clinique.* Le recrutement des patients est réalisé une seule fois. Ensuite, les patients peuvent participer aux études qui les concernent et qui les intéressent.
- *Simplifier la réalisation de recherches nécessitant l'accès à des données externes (entrepôts de données hospitaliers, SNIIRAM, objets connectés etc.).*
- *Améliorant le partage de données.* Toutes les informations recueillies pourront être réutilisées. Ceci allègera le fardeau des participants qui ne devront pas répondre plusieurs fois à la même question, posée par des chercheurs différents ; limitant un « gâchis de la recherche »
- *Réduisant les coûts liés à la recherche avec une plateforme commune.* La plateforme ne sera construite qu'une seule fois et servira de multiples fois. Toutes les améliorations bénéficieront à l'ensemble des recherches qui y sont (et seront) menées.

### 3. Objectifs

Dans ce projet, nous chercherons à différencier des trajectoires d'évolution cliniques de la maladie (définies à partir des symptômes rapportés par les patients)

#### Design

Ce travail comportera deux étapes distinctes nichées au sein de la cohorte ComPaRe COVID long.

- Dans une première étape, nous utiliserons les données de la cohorte ComPaRe Covid long pour identifier des trajectoires d'évolution clinique de la maladie.
- Dans une seconde étape, nous inviterons un sous-ensemble de patients volontaires de la cohorte à réaliser un prélèvement sanguin et chercherons à identifier différentes signatures métabolomiques de la maladie.

#### 3.1. Patients participants

Les critères d'éligibilité sont :

Nous incluons dans cette analyse

- Les patients adultes,
- Inclus dans la cohorte ComPaRe Covid long
- Rapportant une infection par la COVID-19 (confirmée ou non). Nous faisons le choix d'inclure des patients avec une infection suspectée (en plus des patients avec une infection confirmée) car 1) la définition de la maladie post COVID-19 selon l'OMS inclut ces patients <sup>3</sup>; et 2) lors de la création de la cohorte en 2020, nous avons inclus de nombreux patients avec une infection suspectée datant de mars-juin 2020 n'ayant pu être testée en raison du nombre limité de tests PCR à cette période.
- Ayant rapporté, au cours de leur suivi au sein de la cohorte, au moins 1 symptôme parmi les 53 symptômes évalués par le questionnaire long COVID ST au moins 8 semaines après le début des symptômes (en accord avec la définition de la maladie post COVID-19 de l'OMS <sup>12</sup>)

Nous excluons :

- Les patients dont les symptômes étaient résolus avant l'entrée dans la cohorte

### **3.2. Modalités de recrutement**

Les patients seront recrutés de deux manières différentes et complémentaires :

#### **3.2.1. Recrutement via ComPaRe**

Tous les participants participant à ComPaRe et répondant aux critères d'inclusion de la présente étude au moment du lancement du projet seront invités à y participer.

## **4. Données recueillies et utilisées**

L'évolution de la maladie sera évaluée par le score Long COVID ST, mesuré tous les 60 jours.

Le long COVID ST est un outil développé à partir de l'expérience vécue des patients qui évalue 53 symptômes de la maladie. Les 53 symptômes évalués couvrent entièrement les 45 symptômes proposés par le "*Clinical Platform Case Report Form (CRF) for Post COVID condition (Post COVID-19 CRF)*" de l'OMS<sup>13</sup>. Le score Long COVID ST est le compte de symptômes (parmi les 53) à chaque mesure. Ce score a montré une bonne validité de construit et une excellente fiabilité (coefficient de corrélation intraclass (ICC) lors du test-retest 0.83, intervalle de confiance à 95% 0.80 - 0.86).

La rémission des symptômes du COVID long est définie par la déclaration, par le patient, de ne plus avoir de symptômes de la maladie. Le moment de la rémission est défini comme la première déclaration d'une disparition complète des symptômes rapportée par le patient.

Le décès et la date de décès des patients sont obtenus par chaînage entre les données de ComPaRe et le fichier des personnes décédées de l'INSEE.

## **5. Analyses statistiques**

L'analyse des données sera réalisée par

Clémence Servier, Viet Thi Tran, Raphaël Porcher
--

### **5.1. Plan d'analyse**

Pour identifier des trajectoires d'évolution cliniques de la maladie permettant, par exemple, de différencier les patients évoluant vers la rémission (i.e. disparition complète des symptômes) ; une amélioration partielle (i.e. réduction du nombre de symptômes mais sans disparition complète) ; une stabilisation (i.e. sévérité inchangée de la maladie au cours du
--

temps) ; ou une aggravation (i.e. augmentation du nombre de symptômes au cours du temps).

Pour cela, nous réaliserons une modélisation mixte à classe latente (*latent class mixed modeling*) du score long COVID ST au cours du temps <sup>14</sup>. Cette modélisation permet de dériver des trajectoires à partir de mesures répétées avec l'hypothèse qu'il existe plusieurs trajectoires (i.e. classes latentes). Ces méthodes ont été utilisées dans de multiples domaines, tels que la néphrologie ou le handicap <sup>15,16</sup>. Cette méthodologie présente plusieurs avantages : 1) elle ne requiert pas le même nombre de mesures par patient ni les mêmes moments de mesure ; 2) elle est adaptée pour modéliser des scores rapportés par les patients lors de visites répétées (i.e. variables non gaussiennes) ; et 3) permettent d'identifier les trajectoires d'évolution en tenant compte des caractéristiques des participants.

L'identification des trajectoires de la maladie sera réalisée sur l'ensemble des données de la cohorte ComPaRe COVID long ayant complété au moins 2 mesures de symptômes (soit au moins 2000 patients, sans compter l'augmentation potentielle de la taille de la cohorte liée au projet).

Les variables suivantes seront considérées lors du développement du modèle : démographiques (e.g. âge, sexe) ; cliniques (nombre de comorbidités parmi la liste ICPC-2, hospitalisation lors de l'infection aigüe par le COVID-19, hospitalisation en réanimation lors de l'infection aigüe) ; et de prise en charge des patients (vaccination contre la COVID-19).

Les trajectoires des patients seront définies en fonction de leur date de début de symptômes. Nous utiliserons toutes les données disponibles depuis l'entrée dans la cohorte jusqu'à leur dernière mesure connue ou leur décès pour définir les trajectoires. Les données manquantes seront gérées par des méthodes d'imputation multiples par équations chaînées

Le modèle final et le nombre de trajectoires associé sera choisi en fonction des critères statistiques (e.g., critère d'Akaike), la discrimination du modèle, son entropie et son interopérabilité, comme décrit dans d'autres études visant à identifier des trajectoires <sup>15,16</sup>.

La modélisation des trajectoires des patients sera réalisée dans la population totale et dans le sous-groupe des patients avec une infection confirmée.

## 6. Gestion du projet

Ce projet spécifique sera piloté par un **Comité scientifique spécifique** dont le rôle sera de définir le projet et veiller à sa rigueur scientifique,

- Valider les outils de recueil de données,
- Proposer les éléments d'animation scientifique
- Superviser l'analyse des données
- Rapporter de manière transparente l'ensemble des résultats du projet.

## 7. Aspects éthiques et réglementaires

Le projet ComPaRe a reçu un avis favorable par le Comité de protection des personnes – Ile de France 1 (IRB : 0008367).

Le projet ComPaRe a reçu un avis favorable du Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) sous le numéro 16-395bis.

Le projet a reçu un avis favorable de la CNIL, sous le numéro de dossier n°916397 (DR-2016-459) en date du 25/11/2016.

### 7.1. Gestionnaire de la recherche

L'Assistance Publique – Hôpitaux de Paris sera le gestionnaire de la recherche. Le Département de la Recherche Clinique et du Développement (DRCD) en sera son représentant.

L'Assistance Publique – Hôpitaux de Paris se réserve le droit d'interrompre la recherche à tout moment pour des raisons administratives.

Le responsable scientifique, également investigateur coordonnateur, s'engage à fournir au gestionnaire les informations relatives aux inclusions des participants dans la recherche.

Toute modification du protocole de recherche devra être soumise au gestionnaire.

### 7.2. Encadrement réglementaire de la recherche

Conformément à la loi °78-17 du 6 janvier 1978 modifiée par la loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel et son décret d'application n° 2005-1309 du 20 oct. 2005, ComPaRe a obtenu l'autorisation du CCTIRS N° 16.395 et de la CNIL N° 916397.

### **7.3. Qualification des intervenants**

Les membres de l'équipe porteuses du présent projet sont habilités à mener les tâches qui leur incombent, c'est-à-dire qu'ils disposent des compétences appropriées et documentées.

### **7.4. Protection des données**

La Direction des Services Informatiques (DSI) de l'AP-HP assure le développement, l'hébergement, le maintien en conditions opérationnelles et l'exploitation de la plateforme ainsi que de l'application ComPaRe. Elle met en œuvre tous ses moyens techniques afin de garantir une sécurisation adéquate des données.

### **7.5. Saisie des données**

Les patients saisiront directement leurs questionnaires dans un espace personnel sécurisé, protégé par un identifiant et un mot de passe, sur un serveur accessible, sécurisé par un certificat, et placé derrière un « firewall ».

### **7.6. Confidentialité et sécurisation des données**

Toute information concernant les participants sera confidentielle et l'identité du participant sera strictement préservée.

Toutes les données à caractère personnel recueillies auprès des patients seront sécurisées et encryptées en cas d'échanges de données ; les communications entre les participants ou les chercheurs et la plateforme seront cryptés par SSL (https).

- Un nombre restreint de personnes habilitées aura accès aux données directement nominatives, à des fins exclusivement logistiques et scientifiques. Elles prendront toutes les précautions nécessaires en vue d'assurer la confidentialité de ces données.
- Conservation des données
- Toutes les données seront conservées sur deux serveurs sécurisés, et traitées de façon strictement confidentielle.
- Les données complétées sur le site seront stockées sur des serveurs agréés pour l'hébergement de données médicales.
- Les deux serveurs seront garantis d'avoir un haut niveau de sécurité.

### **7.7. Accès aux données**

Seuls les chercheurs habilités et les intervenants en charge de la coordination de la recherche auront accès aux données, et cela uniquement pendant la durée de celle-ci.

## **7.8. Destruction des données**

Les données des participants seront conservées pendant 10 ans après la fin de la recherche sur les personnes (i.e. fin du suivi du dernier participant recruté). Elles seront ensuite détruites, conformément à la loi Informatique, Fichiers et Libertés.

## **7.9. Information et consentement des participants**

### ***7.9.1. Information***

Lors de leur inscription en ligne, les participants devront consulter la note d'information expliquant le cadre et les objectifs de ComPaRe. Ils seront informés du respect de la stricte confidentialité de leurs données ainsi que de leurs droits relatifs à la loi Informatique, Fichiers et Libertés. Il leur sera notamment notifié qu'à tout moment ils pourront exercer leur droit d'accès et de rectification des données, conformément à la loi 78-17 du 06 janvier 1978 relative à l'Informatique, aux Fichiers et aux Libertés, modifiée par la loi n°94-548 du 1er juillet 1994, relative au traitement des données nominatives ayant pour fin la recherche dans le domaine de la santé.

### ***7.9.2. Consentement***

Les patients déjà inscrits dans ComPaRe auront préalablement signés le formulaire de consentement général et pourront participer à ce projet. **Le consentement général couvre la participation à ce projet**

## **7.10. Publication et communications**

Les résultats des travaux réalisés à partir des données obtenues sur ComPaRe doivent être rendus publics. Dans le cas où ces travaux donnent lieu à une publication scientifique, un résumé « grand public » sera publié sur le site de ComPaRe

**La mention que la recherche a été réalisée dans le cadre de la cohorte ComPaRe doit être faite dans toute les publications ou communications. Le mot « ComPaRe » doit figurer autant que possible dans le titre de toute publication.**

Le responsable d'un projet de recherche est tenu de communiquer à l'équipe ComPaRe une copie des manuscrits avant soumission pour publication.

Les signataires des papiers utilisant les données ComPaRe doivent être déterminés en accord avec les équipes impliquées l'organisation de la recherche et le recueil des données.

## 8. Bibliographie

1. Callard F, Perego E. How and why patients made Long Covid. *Soc Sci Med.* 2021;268:113426.
2. Office for National Statistics. *Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK: 2 September 2021.* London, UK2021.
3. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect Dis.* 2021.
4. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature.* 2021;594(7862):259-264.
5. Tran VT, Riveros C, Cleprier B, et al. Development and validation of the long covid symptom and impact tools, a set of patient-reported instruments constructed from patients' lived experience. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2021;74(2):278-287.
6. Ayoubkhani D, Khunti K, Nafilyan V, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. *BMJ (Clinical research ed).* 2021;372:n693.
7. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet (London, England).* 2021;398(10302):747-758.
8. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *The lancet Psychiatry.* 2021;8(5):416-427.
9. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine.* 2021;38:101019.
10. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nature medicine.* 2021;27:626-631.
11. Tran V, Porcher R, Pane I, Ravaud P. Course of long COVID symptoms over time in the ComPaRe long COVID prospective e-cohort. *Nature Communications (submitted).* 2021.
12. Soriano J, Allan M, Alsokhn C, et al. *A clinical case definition of post COVID-19 condition by a Delphi consensus.* World Health Organization;2021.
13. World Health Organization. Global COVID-19 Clinical Platform Case Report Form (CRF) for Post COVID condition (Post COVID-19 CRF). 2021; [https://www.who.int/publications/i/item/global-covid-19-clinical-platform-case-report-form-\(crf\)-for-post-covid-conditions-\(post-covid-19-crf-\)](https://www.who.int/publications/i/item/global-covid-19-clinical-platform-case-report-form-(crf)-for-post-covid-conditions-(post-covid-19-crf-)). Accessed 18/02/2021.
14. Proust C, Jacquemin-Gadda H, Taylor JM, Ganiayre J, Commenges D. A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data. *Biometrics.* 2006;62(4):1014-1024.
15. Raynaud M, Aubert O, Reese PP, et al. Trajectories of glomerular filtration rate and progression to end stage kidney disease after kidney transplantation. *Kidney international.* 2021;99(1):186-197.
16. Gill TM, Gahbauer EA, Han L, Allore HG. Trajectories of disability in the last year of life. *The New England journal of medicine.* 2010;362(13):1173-1180.