

**Projet ANR-21-CE45-0024-01**

## **ReproVIP**

Programme JCJC 2021

A	IDENTIFICATION .....	1
B	PARTICIPANTS A LA REUNION .....	2
C	ORDRE DU JOUR.....	2
D	PRESENTATIONS ET DISCUSSIONS.....	2

### **A IDENTIFICATION**

Acronyme du projet	ReproVIP
Titre du projet	Reproducibility with VIP
Coordinateur du projet (société/organisme)	Sorina POP, CNRS, Laboratoire Creatis
Date de début du projet	01/02/2022
Date de fin du projet	31/07/2024
Site web du projet, le cas échéant	<a href="https://www.creatis.insa-lyon.fr/reprovip/">https://www.creatis.insa-lyon.fr/reprovip/</a>

Rédacteur de ce rapport	
Civilité, prénom, nom	Sorina POP
Téléphone	+33 (0)4 72 43 81 91
Courriel	<a href="mailto:Sorina.pop@creatis.insa-lyon.fr">Sorina.pop@creatis.insa-lyon.fr</a>
Date de réunion	23/11/2023
Réunion faisant l'objet du rapport d'activité	Quatrième réunion plénière du projet ReproVIP

## B PARTICIPANTS A LA REUNION

- En présentiel : Hélène Ratiney, Axel Bonnet, Gaël Vila, Frédéric Cervenansky, Sorina Pop
- A distance (en visio conférence) : Tristan Glatard, Yohan Chatelain, Jérôme Pansanel
- Excusés : Emmanuel Medernach, Alexandre Cornier, Carole Frindel

## C ORDRE DU JOUR

- 14h20 : "Reflexions sur la reproductibilité en spectroscopie quantitative: Apport du projet ReproVIP", ~15+5 minutes, Hélène
- 14h40 : "Impact de la qualité des données sur la reproductibilité en spectroscopie quantitative: retours d'une étude pilote", ~10+5 minutes, Gaël
- 14h55 : "Reproductibilité de l'environnement d'exécution: apports et limites de deux solutions logicielles pour le déploiement d'applications", ~10+5 minutes, Gaël
- 15h10 : "LivingPark - Numerical Variability Analysis of Parkinson Biomarker", ~15+5 minutes, Yohan
- 15h30 : "Client Python pour le lancement d'applications sur VIP: développements et perspectives", ~10+5 minutes, Gaël
- 15h45 : "État des lieux du dashboard repro-vip", ~10+5 minutes, Axel
- 16h00 : "Organisation fin du projet ReproVIP", ~10+5 minutes, Sorina
- 16h15 : Discussions/Divers, tous

## D PRESENTATIONS ET DISCUSSIONS

### D.1.1 PRESENTATION « REFLEXIONS SUR LA REPRODUCTIBILITE EN SPECTROSCOPIE QUANTITATIVE: APPORT DU PROJET REPROVIP » (HELENE)

Il est intéressant de lier les problématiques de reproductibilité sur l'ensemble du pipeline dans une réflexion globale.

Les spectres acquis à l'intérieur du cerveau sont très bruités, nécessitant plusieurs acquisitions et étapes de traitement : pre-processing, moyennage de plusieurs acquisitions et quantification.

Objectif : remonter à une estimation de la concentration des métabolites

- utilisation d'une toolbox d'optimisation en Fortran
- quantification des signaux acquis à temps d'écho court
  - o complexité du spectre

Sources de variabilité : de nombreux algos pour le même calcul : FSL-MRS, LCModel, Tarquin, jMRUI

Problème de minimisation avec fonction objectif complexe

- cQuest ne régularise pas l'optimisation
- il y a plusieurs stratégies d'optimisation étant donné le problème d'estimation

Apport du projet ReproVIP

- mise en place d'un pipeline complet
- capacité à tester, éprouver, comparer l'effet
  - o du paramétrage de l'algo, du preprocessing
  - o des différences entre différents algos
  - o de la graine aléatoire
  - o des versions/compilation

- lien avec la qualité des données

Perspectives

- lien entre incertitude de mesure et reproductibilité : pouvoir positionner le résultat par rapport à une erreur de mesure
- en utilisant les outils dans ReproVIP, continuer le travail commencé avec Gaël et Théotime

- étudier à quel point les bornes de Cramer-Rao sont proches de ce qu'on peut obtenir avec des simulations MC

## Discussions

- article à regarder : « The importance of uncertainty quantification in model reproducibility » (Victoria Volodin and Peter Challenor)
- les incertitudes dans la résolution des problèmes d'optimisation se retrouvent dans plusieurs domaines (recalage, apprentissage, etc)
- dès que qu'on rajoute du bruit, on part dans des solutions différentes
- ce n'est pas un problème uniquement mathématique ou informatique ; les solutions qu'on essaie d'optimiser sont imparfaites (mal définies) ; il y a plusieurs solutions
- une approche intéressante : considérer un ensemble de solutions (et non pas une seule) ; faire de l'augmentation de données
- un résultat seul n'est pas suffisant, même avec son incertitude ; au moins multiplier les observations ; comment inférer la conclusion ? mieux exploiter la variabilité entre les résultats (surtout sur l'aspect quantitatif)
- en neuro il y a beaucoup de discussions sur la variabilité ; souvent on essaie de la masquer ou faire des efforts pour la minimiser, mais il y a peut-être du signal à exploiter
- mieux comprendre le comportement des algos en fonction du bruit
- suggestion d'utiliser verifcarlo
- différence entre mettre du bruit sur les entrées ou dans le code
  - o avec le bruit numérique on se retrouve d'habitude autour d'un min local
  - o avec des initialisations aléatoires on peut se retrouver avec une variabilité différente
- sensibilité aux paramètres : uncertainty quantification

## D.1.2 PRESENTATION « IMPACT DE LA QUALITE DES DONNEES SUR LA REPRODUCTIBILITE EN SPECTROSCOPIE QUANTITATIVE: RETOURS D'UNE ETUDE PILOTE» (GAËL)

### Retour sur le stage avec Théotime

- contexte : la variabilité des sorties (concentration des métabolites) dépend du signal en entrée
- objectif : contrôler l'acquisition en temps réel (savoir à quel moment on atteint la qualité suffisante pour arrêter l'acquisition)

### Plusieurs angles pour aborder le problème :

- sélection d'indices de qualité (attention aux outils)
- étude sur signaux réels, simulés et signal accumulé

### Qualité du signal

- signal sur bruit
- largeur des pics : largeur à mi-hauteur (FWHM)

### Etude sur signal simulé

- SNR lié à la précision
- FWHM lié à la justesse

### Perspectives

- meilleurs indices de qualité
- nouvelles données
- remesurer le FWHM sur des signaux filtrés (avec moins de bruit)

### Variabilité code/données

- les deux sont intéressants en fonction du contexte
- variabilité trop importante : cela nuit à l'algo
- variabilité plus faible : cela peut être utile
- comprendre l'ordre de grandeur et les sources de variabilités

### D.1.3 PRESENTATION « REPRODUCTIBILITE DE L'ENVIRONNEMENT D'EXECUTION: APPORTS ET LIMITES DE DEUX SOLUTIONS LOGICIELLES POUR LE DEPLOIEMENT D'APPLICATIONS» (GAËL)

Focus sur l'environnement d'exécution

Objectif : quantifier l'influence de l'infrastructure de calcul sur la (non-)reproductibilité des résultats d'une application (FSK-FLIRT)

Expérience

- Données IRM publiques
  - o Base de données OASIS-1: 148 images (39 sujets)
- Prétraitement
  - o Réorientation des images sur les coordonnées RAS
- Application
  - o FSL-FLIRT: recalage non-linéaire (sur le template MNI-152)
- Infrastructure de calcul
  - o Grid'5000: 25 machines réparties sur 8 sites, OS / hardware différents

Résultats

- La plupart des sorties ne sont pas reproductibles entre les environnements (GUIX et Docker)
- Avec GUIX, reproductibilité des résultats, mais probablement ignore les optimisations propres à l'architecture
- Avec Docker, reproductibilité des résultats sur les clusters d'une même génération

Discussions

- Jérôme et Emmanuel sont intéressés de poursuivre ces travaux ; à noter aussi la disponibilité de serveurs ARM à Strasbourg

### D.1.4 PRESENTATION « LIVINGPARK - NUMERICAL VARIABILITY ANALYSIS OF PARKINSON BIOMARKER", ~15+5 MINUTES» (YOHAN)

Projet Living Park

- reproduire des papiers
- étude plus précise de Freesurfer sur la variabilité (analyse de variabilité numérique)
  - o évaluer quels sont les facteurs qui peuvent expliquer la variabilité

Questions :

- est ce que la variabilité numérique affecte les résultats (d'un point de vue clinique)
- est-ce qu'on observe de la variabilité entre les patients sains (HC) et les patients Parkinson (PD)

Utilisation de fuzzy-libm

- version de FS 7.3.1 dans un Docker avec fuzzy
- ajout du bruit sur le résultat des opérations (ex exponentielle)
- 13 répétitions

Cohorte

- 315 sujets
- 2 niveaux d'analyse : subject and group (PD vs HC)

Cortical thickness

- nombre de chiffres significatifs de 1.5 en moyenne (très bas)

Subject level analysis

- low significance/high uncertainty
- Même analyse avec des versions de FS différentes
- Comparaison avec variabilité issue de l'init avec graine aléatoire : globalement du même ordre
- Ici variabilité issue uniquement de la lib math

Group level analysis

- pas de différence statistique de variabilité entre PD et HC

Comment la variabilité impacte des mesures/indicateurs cliniques ?

- comparer comment les variations longitudinales sont corrélées avec l'évolution de la maladie
- entre les versions de FS, on trouve des choses très différentes

## D.1.5 PRESENTATION « CLIENT PYTHON POUR LE LANCEMENT D'APPLICATIONS SUR VIP: DEVELOPPEMENTS ET PERSPECTIVES » (GAËL)

Présentation des évolutions et de la nouvelle structure du client

Caractéristiques principales

- Disponible sur Pypi
- Classe « utilisateur » VipSession
- Sauvegarde/chargement de métadonnées

Perspectives d'amélioration

- Plus de modularité (VipClient, etc.)
- Stratégies de lancement avancées (VipLauncher + VipLoader)
- Plus de parallélisme (upload)
- Jeu de test public (pytest ?)
- Meilleurs logs (logging)

## D.1.6 PRESENTATION « ÉTAT DES LIEUX DU DASHBOARD REPRO-VIP » (AXEL)

Travaux basés sur le stage de Hippolyte Blot, qui seront poursuivis pendant son apprentissage

Objectif principal du Dashboard : interprétation visuelle des résultats

Métriques

- Comment les définir par rapport à chaque use-case
- Actuellement générées manuellement
- Automatisation prévue dans le cadre d'un plugin Girder

Résultats

- Produits dans VIP, mais stockés à différents endroits
- Envisager d'utiliser la provenance pour une recherche efficace

## D.1.7 PRESENTATION « ORGANISATION FIN DU PROJET REPROVIP » (SORINA)

Prolongation du projet de 6 mois, nouvelle fin au 31/07/2024

Livrables

- Bonne progression
- Quelques livrables à finaliser et fournir rapidement

Publications

- Deux articles ISBI 2023
- En cours : chapitre de livre et article sur l'architecture ReProVIP

Perspectives

- Mieux comprendre les sources de variabilité (dans les modèles de recalage, cquest et autres algos)
  - o Ex : le recalage rigide est relativement simple, comprendre pourquoi le problème est instable. Que pourrait-il être fait pour l'éviter (régularisation) ?
- Utiliser la variabilité pour faire de l'augmentation des données
- Mieux comprendre les différences observées dans les expériences G5K
  - o Ex : quels paramètres/options rendent le déploiement GUIX plus reproductible que l'utilisation des conteneurs ?
- Approfondir les deux use-cases

Suite ReProVIP

- Continuer à y réfléchir ensemble