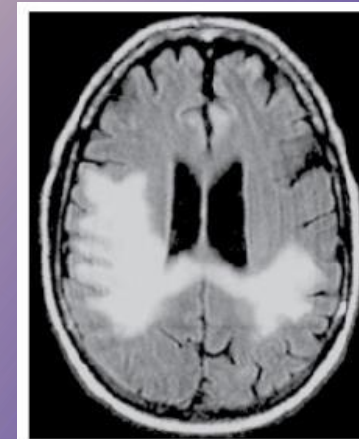




Modélisation et outils d'IA pour évaluer la neurotoxicité radio-induite à partir d'IRM cérébrales collectées chez des patients traités pour un glioblastome

Sophie Ancelet

Autorité de Sûreté Nucléaire et de Radioprotection (ASNR) – PSE-SANTE/SESANE/LEPID
sophie.ancelet@asnr.fr



Rencontres Chercheur·euse·s & Ingénieur·e·s. « Maths et Cerveau »
20 novembre 2025 – Paris

CONTEXTE CLINIQUE, OBJECTIFS ET ORGANISATION DU PROJET



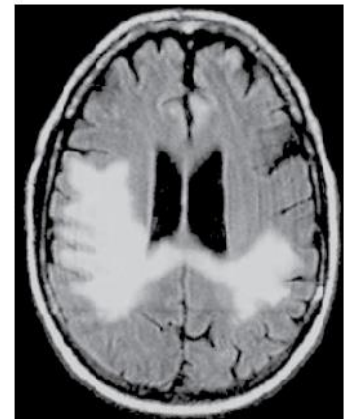
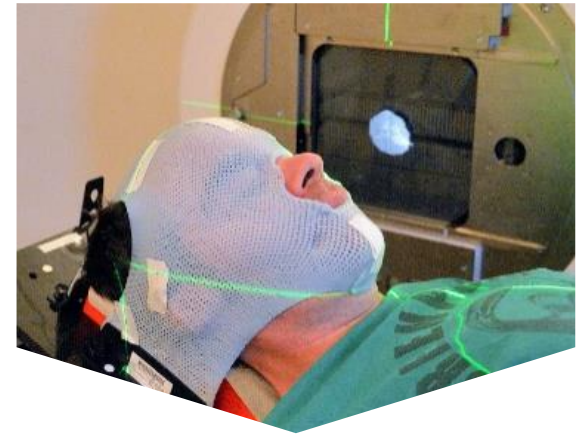
QUELQUES INFORMATIONS GENERALES

- ▶ RADIOAIDE: **RAD**iation-induced neur**O**toxicity assessed by spatio-temporal modelling combined with **Artificial Intelligence** after brain ra**D**ioth**E**rapy
- ▶ Projet en **science des données** cofinancé par **l'agence nationale de la recherche (ANR)**
- ▶ Coordination : ASNR / Sophie Ancelet
- ▶ 8 partenaires des sphères publique et privée dont 1 PME spécialisée en neuroimagerie : Pixyl
- ▶ Durée du projet : Mars 2022- **Mars 2027**



CONTEXTE CLINIQUE (1/2)

- ▶ La **radiothérapie (RT) cérébrale** : un **outil clé** dans le traitement des tumeurs du système nerveux central (SNC)
 - Prise en charge personnalisée et plus en plus précise en termes de zone d'irradiation, afin d'épargner le tissu cérébral sain [Shih et al., 2009]
 - ▶ Néanmoins, la RT cérébrale reste pourvoyeuse d'**effets secondaires**
 - Altération post-RT de différentes fonctions cognitives rapportée dans de nombreuses études [Klein et al. 2002, Scoccianti et al. 2012, Makale et al. 2016; Jacob et al, 2018, Ribeiro 2020, Soussain et al, 2009] et ce parfois même quand la tumeur est contrôlée => Altération significative de la qualité de vie des patients longs survivants
- ▶ **Quels sont les mécanismes intracérébraux neurotoxiques qui seraient à la fois induits par la RT et susceptibles d'expliquer l'altération des fonctions cognitives chez des patients traités par RT cérébrale?**



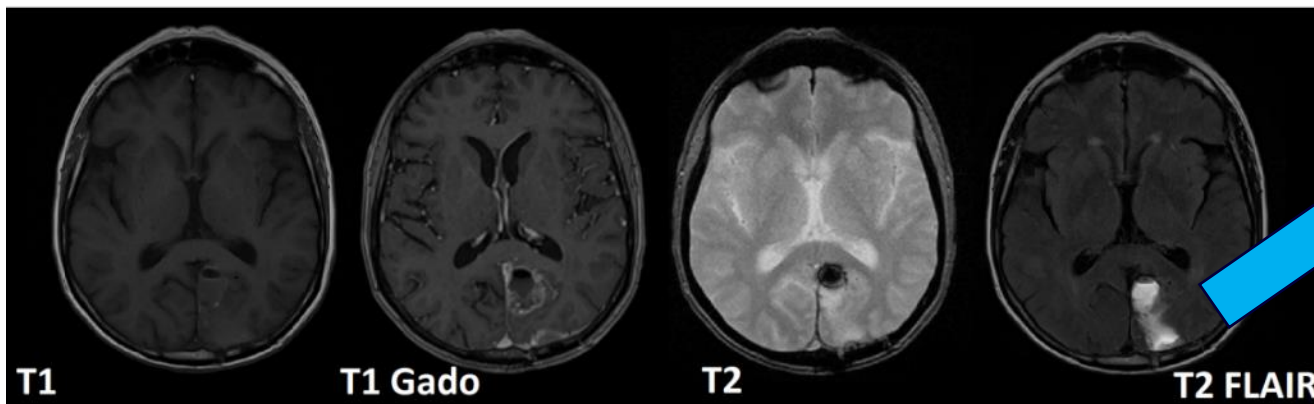
CONTEXTE CLINIQUE (2/2)

► Principales **lésions cérébrales** observées sur IRM après RT du cerveau

- Anomalies irréversibles de la substance blanche
- Atrophie cérébrale
- Œdèmes post-radiques et péritumoraux
- Radionécrose
- Micro-saignements
- Récidive tumorale et/ou cavité tumorale (si exérèse pré-RT)

} **Leuco-encéphalopathie**

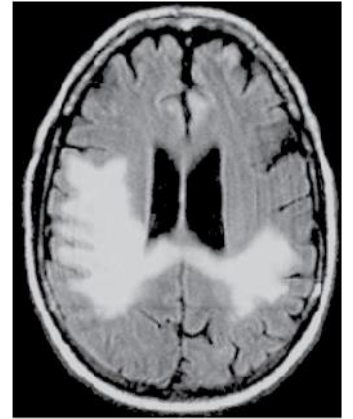
► Différentes **modalités IRMS** permettent de mettre en évidence différents types de lésions cérébrales



Hyperintensités de la substance blanche

- ☐ Anomalies de la substance blanche (**plages de leucoencéphalopathie – PDLs**)
- ☐ Œdèmes

HYPOTHESE ET ENJEU CLINIQUE



Hypothèse : la **leuco-encéphalopathie** serait à la fois induite par la RT et susceptible d'expliquer l'altération des fonctions cognitives chez des patients traités par RT cérébrale

Enjeu: Apporter de **nouvelles connaissances** permettant de valider/invalider cette hypothèse

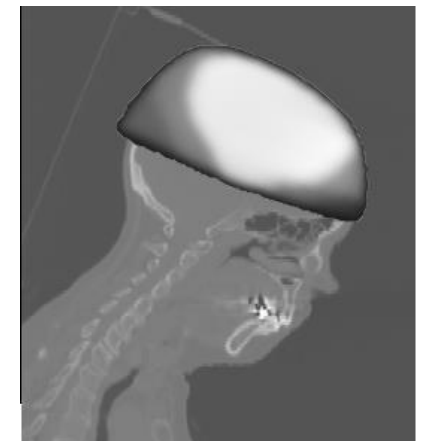
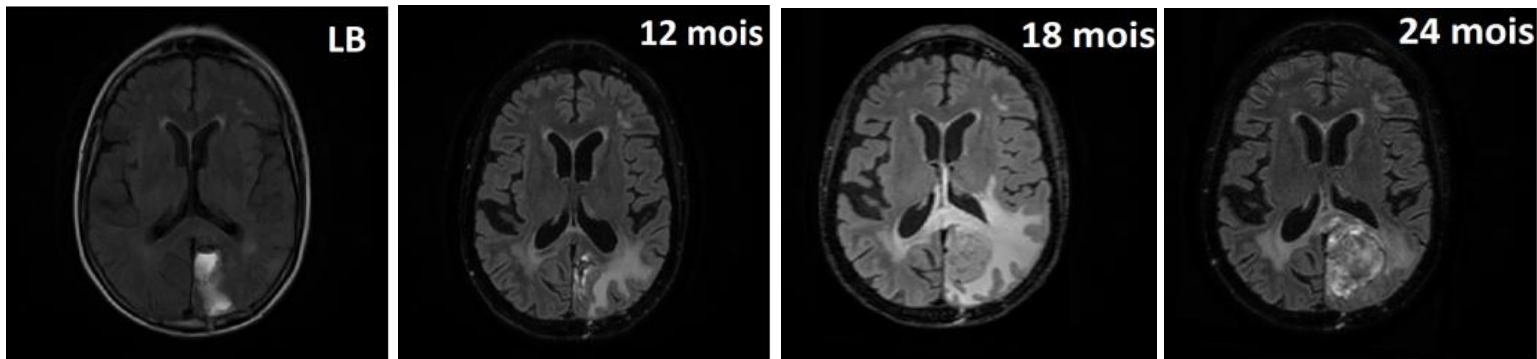
QUESTIONNEMENTS SPECIFIQUES

- ▶ Existe-t-il une **association** entre la dose de RI reçue par certains tissus sains intracérébraux et l'incidence d'une atrophie cérébrale chez des patients traités par RT cérébrale?
- ▶ Existe-t-il une **association** entre la dose de RI reçue par certains tissus sains intracérébraux et la **progression spatio-temporelle de PDLs** au sein de la substance blanche chez des patients traités par RT cérébrale? Si oui :
 - Quelle est la **force** de cette association?
 - Est-elle spatialement homogène ou hétérogène au sein de la substance blanche?
 - Quelle est la dynamique temporelle de cette association?
- ▶ Existe-t-il des **caractéristiques spécifiques des PDLs** et de leur évolution spatio-temporelle post-RT cérébrale ?
- ▶ Existe-t-il une **association** entre la progression spatio-temporelle de PDLs au sein de la substance blanche et l'occurrence de troubles cognitifs chez des patients traités par RT cérébrale?

Radiosensibilité
de la substance blanche
en fonction de la dose
reçue, du temps, de la
région anatomique?

LA COHORTE EPIBRAINRAD (1/2)

- ▶ Cohorte de **276 patients** traités par radio-chimiothérapie pour un **glioblastome**
 - 182 patients suivis à l'hôpital de la Pitié-Salpêtrière (Paris), 94 à l'institut de cancérologie (ICANS) de Strasbourg
 - Mise en place en 2015 (coordination: ASNR)
- ▶ **Données individuelles** collectées prospectivement pendant le suivi clinique des patients :
 - Données cliniques : âge, sexe, grade/localisation tumeur, traitements, diabète, tabac,,...
 - Données dosimétriques liées au traitement par RT
 - Résultats issus de tests cognitifs réalisés avant RT puis à 12, 24 et 36 mois
 - **Données d'imagerie : Collection d'IRM cérébrales multimodales (FLAIR, T1-gado, etc.), scanner dosimétrique**



LA COHORTE EPIBRAINRAD (2/2)

► Au 31/10/2025:

- **88% patients décédés**
- **2245 examens IRM** réalisés avant RT, à un mois post-RT puis à un rythme bimestriel si pas d'événement clinique intercurrent
- **5 examens IRM par patient en moyenne**
- 101 patients (39%) avec bilans neuropsychologiques à T0 (avant RT) et T0+12 mois
- 17 patients (7%) avec bilans neuropsychologiques à T0 (avant RT) , T0+12 mois et T0+24 mois

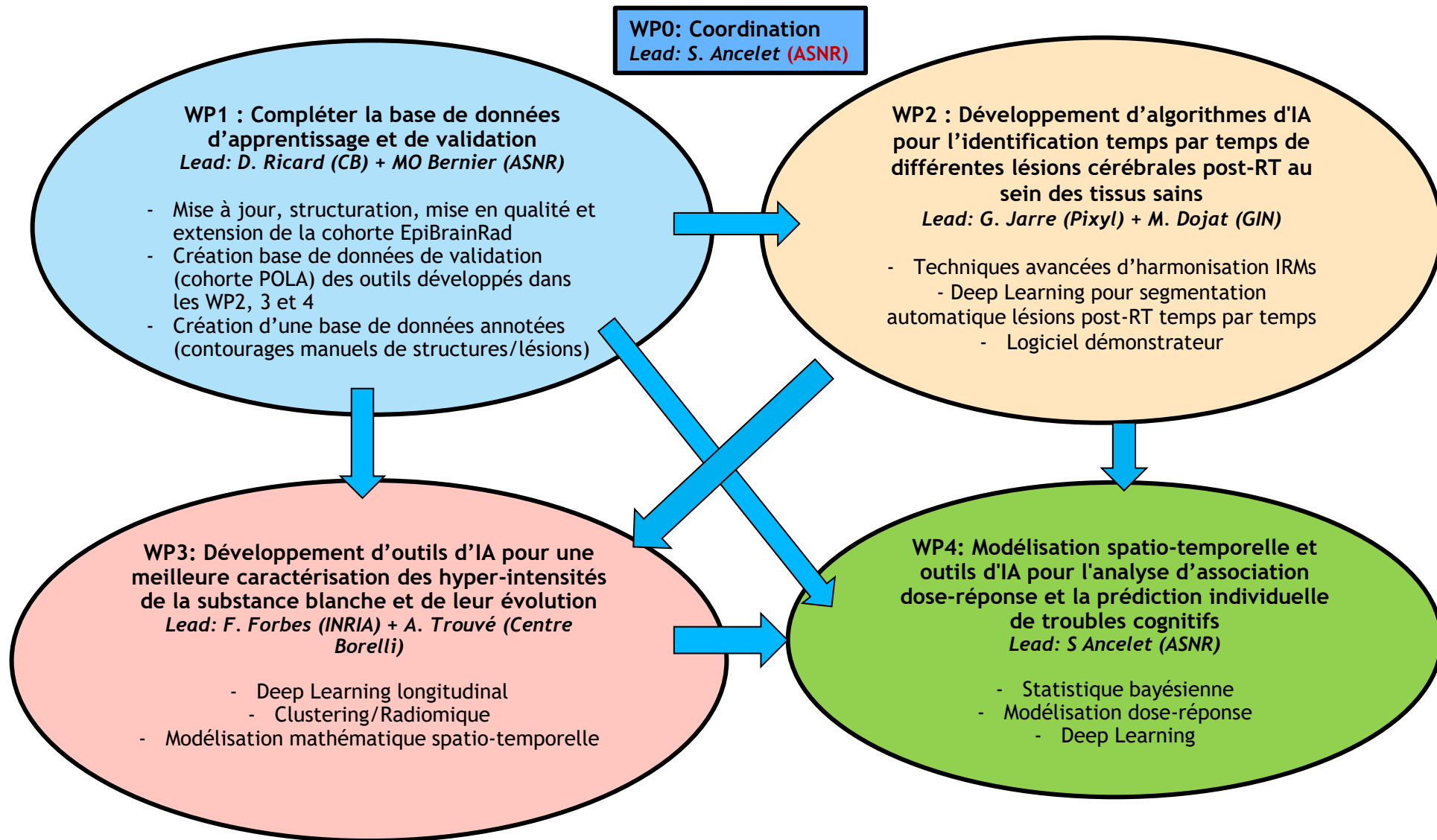
VERROUS TECHNIQUES

- ▶ Aucun logiciel ne permet de segmenter automatiquement des **lésions cérébrales post-radiothérapie** à partir d'IRMs cérébrales **obtenues lors du suivi clinique de patients** atteints de tumeur cérébrale
 - Cotation manuelle des lésions visibles par les radiologues
 - Pas de détection précoce possible à l'œil nu de modifications temporelles « subtiles » des tissus sains
- ▶ Pas de **bases de données IRMS annotées** disponibles
- ▶ Grande **hétérogénéité** des IRMs cérébrales intra et inter patients collectées dans la cohorte EpiBrainRad
 - IRMs réalisées dans différents centres médicaux, sur différentes machines
 - Protocoles radiologiques non harmonisés
 - Cerveaux déformés après exérèse de la tumeur
- ▶ Données d'imagerie longitudinales, hétérogènes (scanner/IRMs) et très volumineuses

OBJECTIFS DU PROJET

- ▶ **Utiliser les mathématiques (et l'informatique)** pour tirer au mieux partie de l'information apportée par les données d'imagerie cérébrales de la cohorte EpiBrainRad et répondre aux questionnements cliniques d'intérêt
- ▶ **Objectifs spécifiques**
- ▶ Développer des **algorithmes de segmentation (IA)** permettant d'identifier **automatiquement** différentes lésions cérébrales post-RT et leur dynamique d'évolution spatio-temporelle
 - **hyper-intensités de la substance blanche**
 - volume substance blanche/grise (atrophie cérébrale)
 - tumeur cérébrale, zone péri-tumorale et lésions post-opératoires
 - ganglions basaux
 - micro-saignements
- ▶ Mettre en place un logiciel démonstrateur pour évaluation par les cliniciens du projet
- ▶ Développer des **modèles mathématiques et algorithmes d'apprentissage** pour :
 - Estimer les associations potentielles entre dose reçue au cerveau, PDLs et troubles cognitifs
 - Identifier d'éventuelles caractéristiques spécifiques des PDLs et de leur évolution spatio-temporelle
 - Prédire individuellement la probabilité d'occurrence de troubles cognitifs à un stade précoce après RT

ORGANISATION DU PROJET

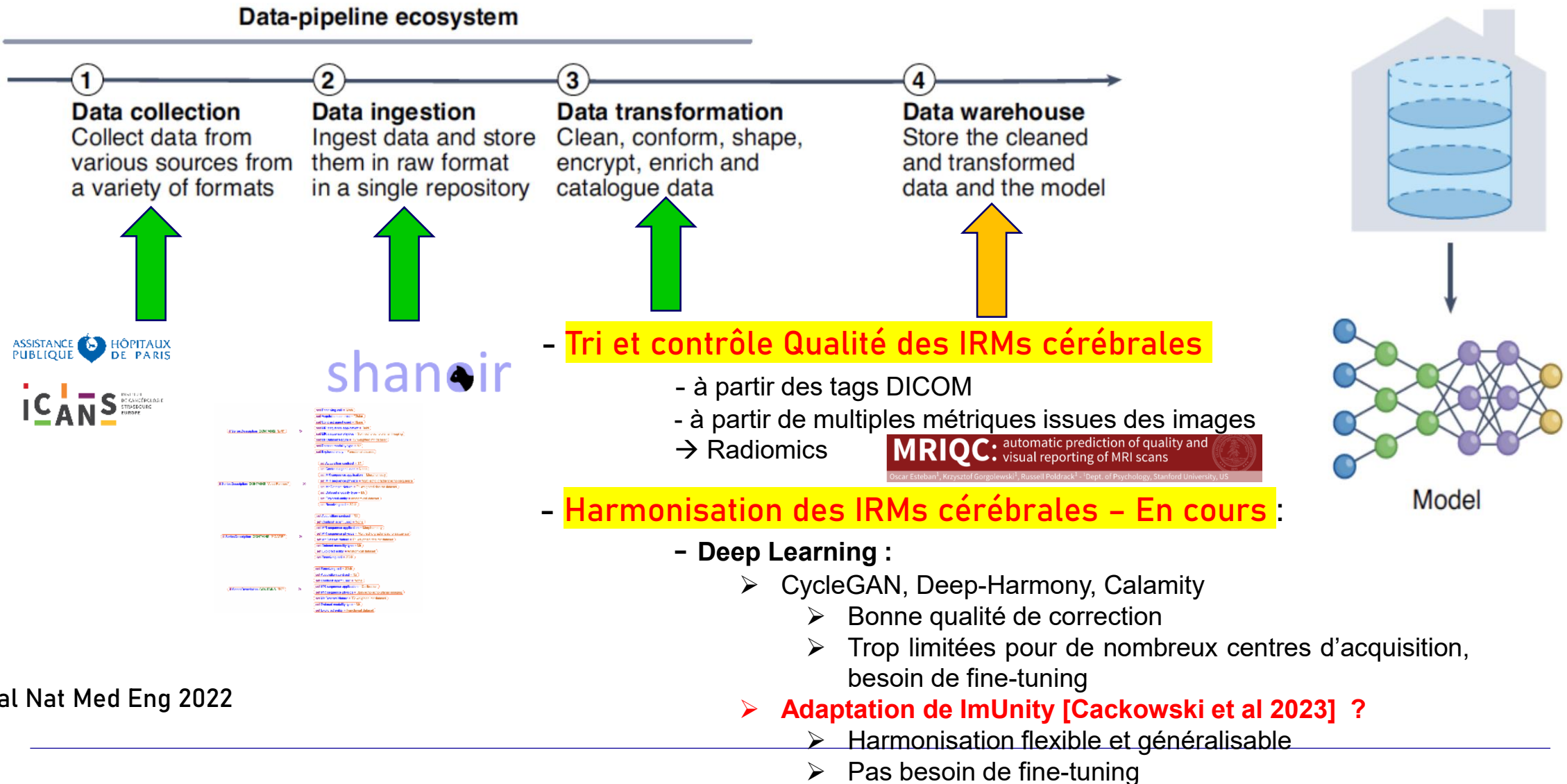


QUELQUES TRAVAUX REALISES



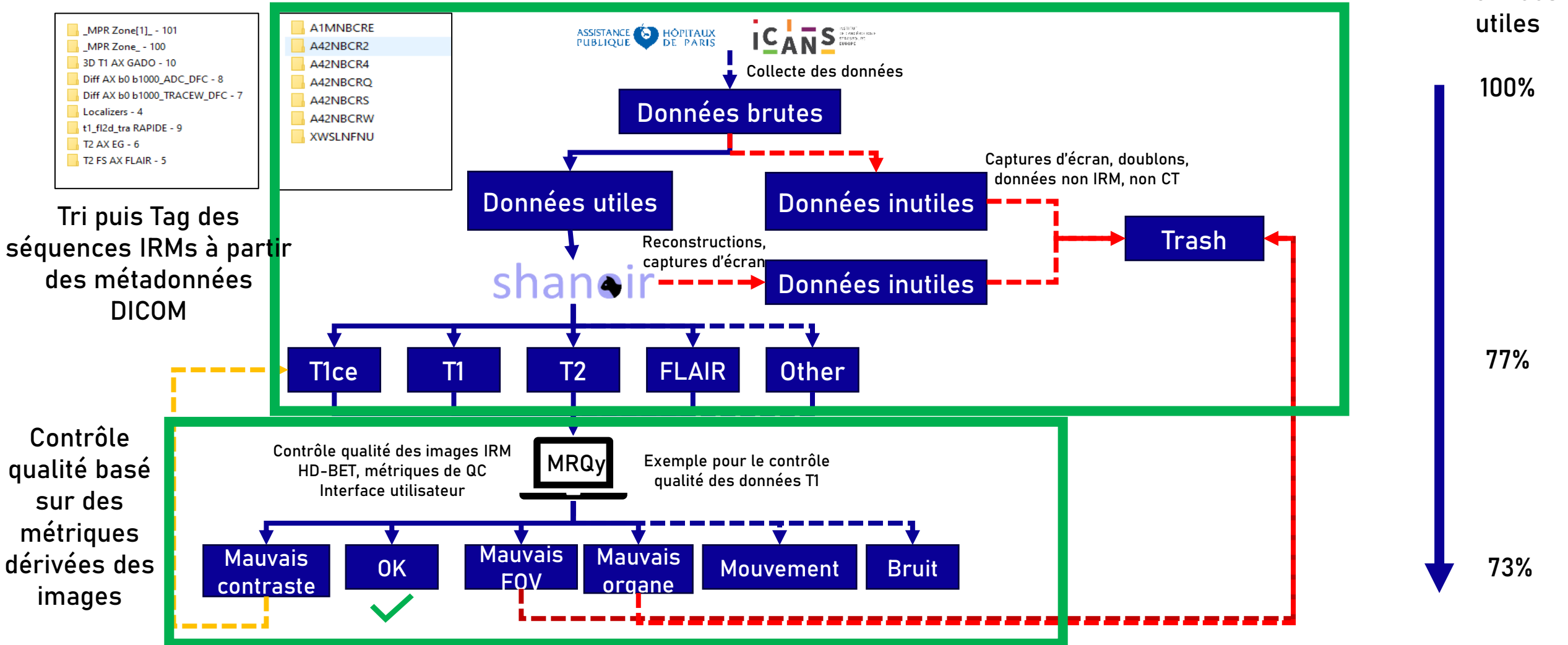
DEUX ETAPES INCONTURNABLES AVANT LA MODÉLISATION...

Travaux réalisés par Michel Dojat, Benjamin Lemasson (Institut des Neurosciences de Grenoble) et Marc Saghia (INRIA)



TRI ET CONTRÔLE QUALITÉ DES IRMS DE LA COHORTE : DÉVELOPPEMENT D'UN ALGORITHME SEMI-AUTOMATIQUE

Travaux réalisés par Michel Dojat, Benjamin Lemasson (Institut des Neurosciences de Grenoble) et Marc Saghia (INRIA)



DETECTION D'UN EFFET CENTRE DANS L'ACQUISITION DES IRMS

HARMONISATION INDISPENSABLE DES IRMS EPIBRAINRAD

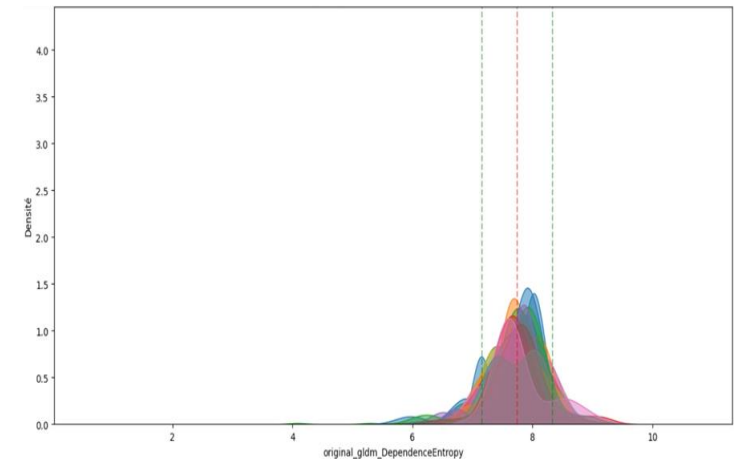
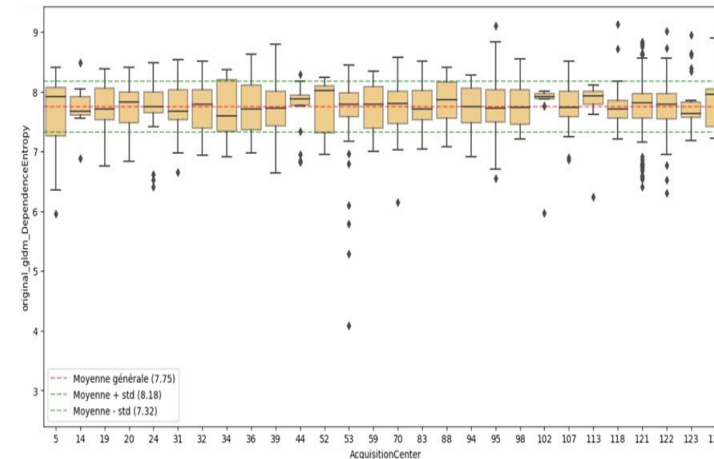
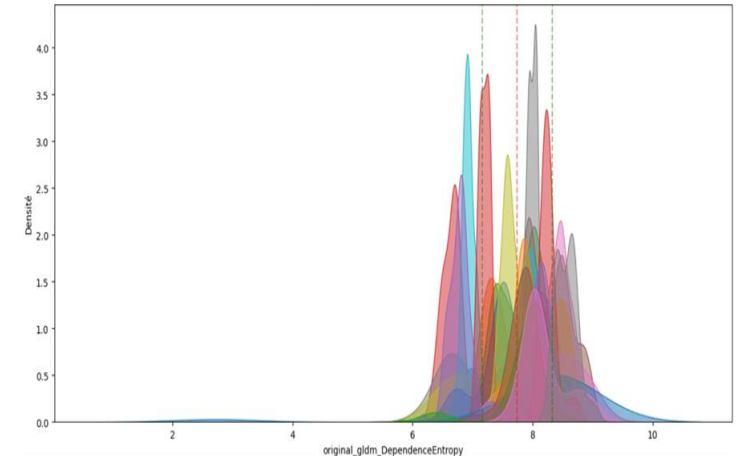
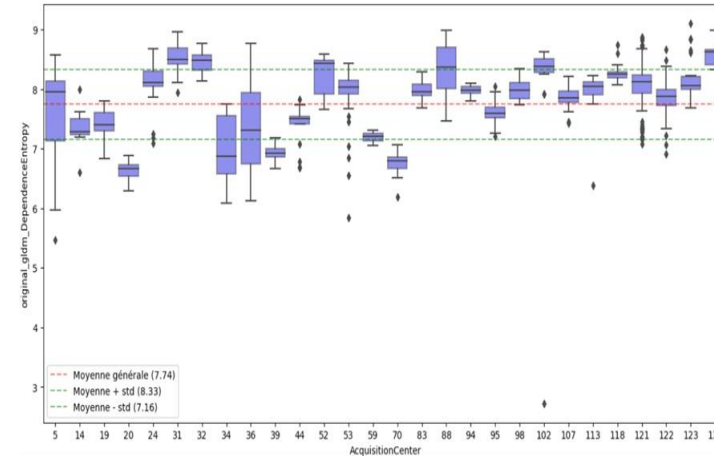
Travaux réalisés par Michel Dojat, Benjamin Lemasson (Institut des Neurosciences de Grenoble) et Marc Saghieh (INRIA)

Pourquoi ?

- ✓ Réduction des biais d'acquisition (humains, machines, ...)
- ✓ Meilleurs résultats pour des algorithmes de classification ou de segmentation automatiques

D'où viennent les variations dans les images ?

- Variabilité due aux différences entre les centres d'acquisition → **effet de centre**
- Différents paramètres techniques, paramètres d'acquisition, variabilité humaine, ...

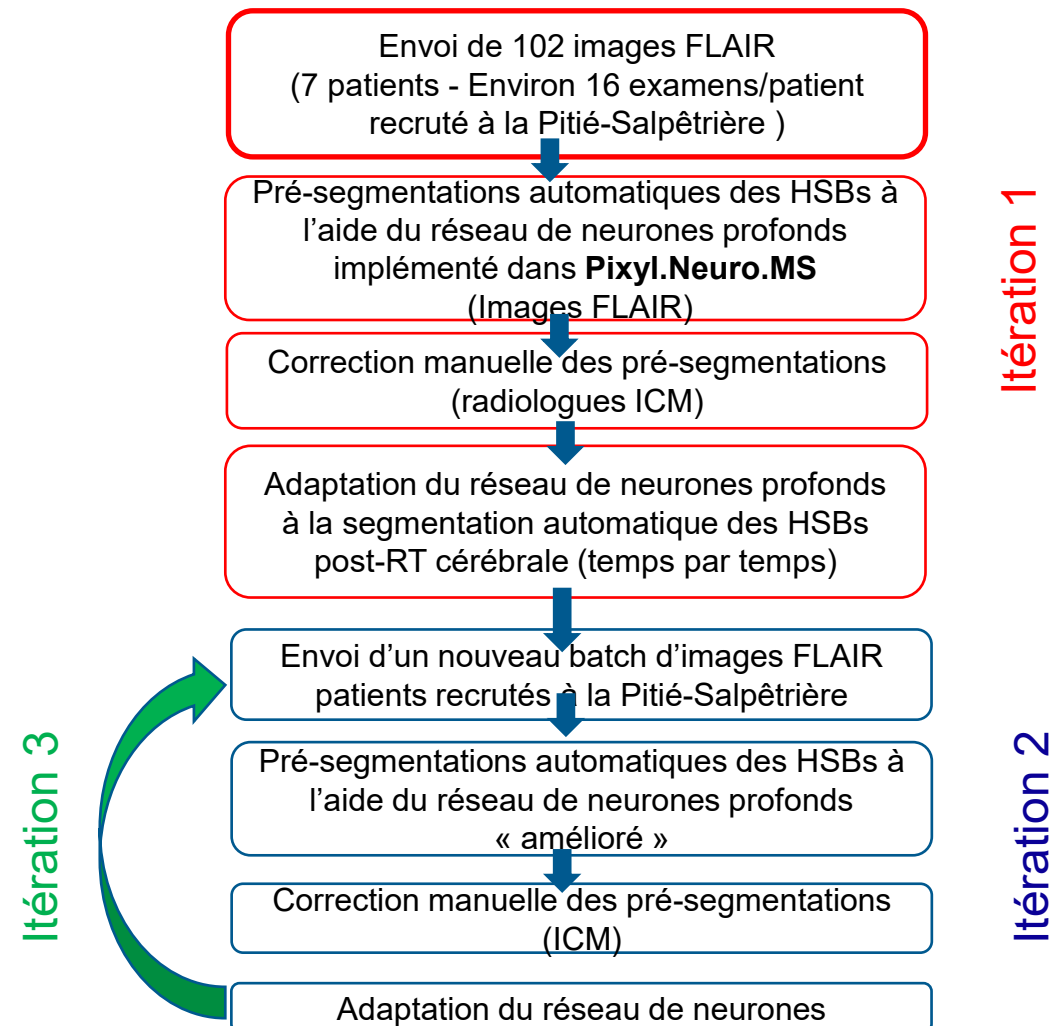
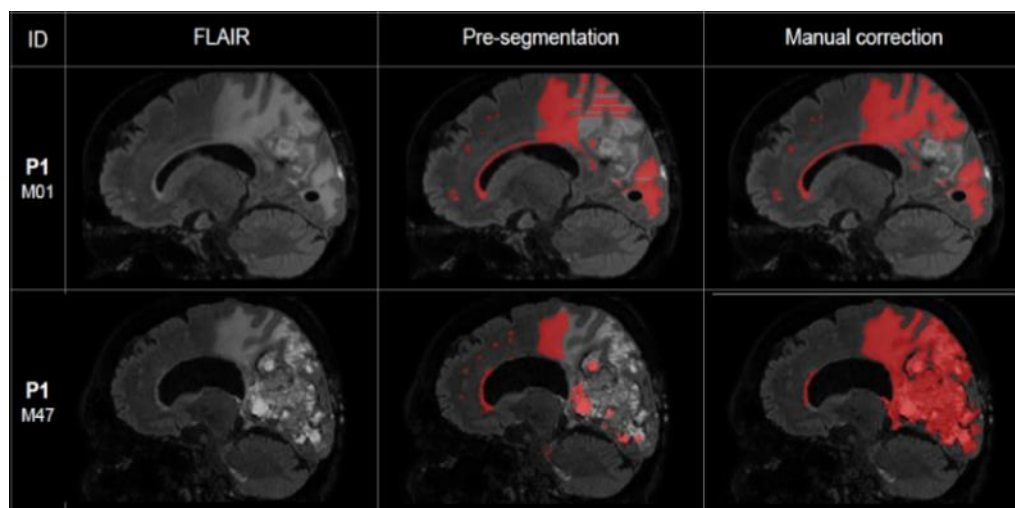


Effet de centre visible, grande variabilité entre les centres. L'harmonisation corrige cet effet.

SEGMENTATION D'HYPERINTENSITES DE LA SUBSTANCE BLANCHE POST-RT DÉVELOPPEMENT ET APPRENTISSAGE DE RÉSEAUX DE NEURONES PROFONDS

Travaux réalisés par la société Pixyl & deux neuroradiologues (Nadya Pyatigorskaya, Lucia Nichelli – ICM)

- ✓ **Comment segmenter des hyper-intensités de la substance (HSBs) blanche post-RT?**
- ✓ **Idée 1 :** Utilisation de **l'outil d'IA développé par Pixyl (Pixyl.Neuro.MS)** pour l'aide au diagnostic et le suivi de la **sclérose en plaques**
- ✓ **Intérêts:**
 - ✓ Partir d'un **réseau de neurones** aux performances éprouvées pour l'identification des HSBs associées à la sclérose en plaques afin de **l'adapter** au contexte spécifique des HSBs post-RT cérébrale
 - ✓ Gain de temps pour les neuroradiologues
⇒ segmentation semi-manuelle
- ✓ **Idée 2:** Processus de construction itératif du modèle de segmentation

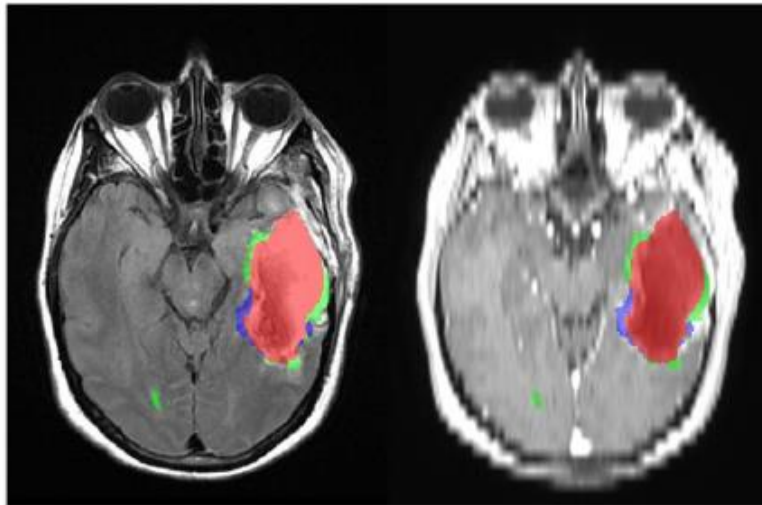


SEGMENTATION DE LA ZONE (PERI)TUMORALE DÉVELOPPEMENT ET APPRENTISSAGE DE RÉSEAUX DE NEURONES PROFONDS

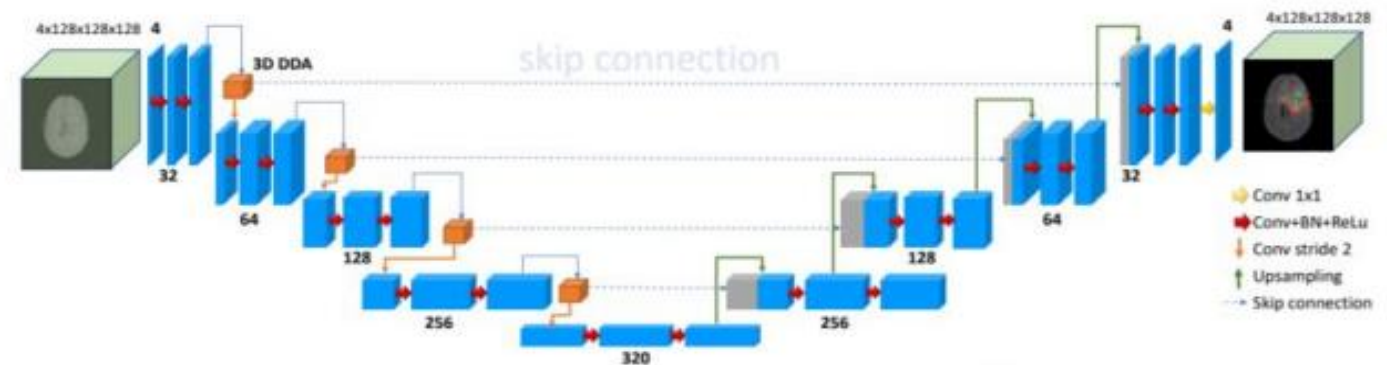
Travaux réalisés par la société Pixyl & deux neuroradiologues (Nadya Pyatigorskaya, Lucia Nichelli – ICM)

Output:

- Contrast-enhanced area
- Necrosis
- Edema/WMH



- ❑ Modèle => Réseau de Neurones DynUNet
- ❑ Données d'apprentissage FLAIR + T1 Gado :
 - ❑ BraTS2023, LUMIERE & Pixyl_MS pour transfert learning
 - ❑ 325 IRMs EpiBrainRAD recrutés à la Pitié-Salpêtrière pour fine tuning



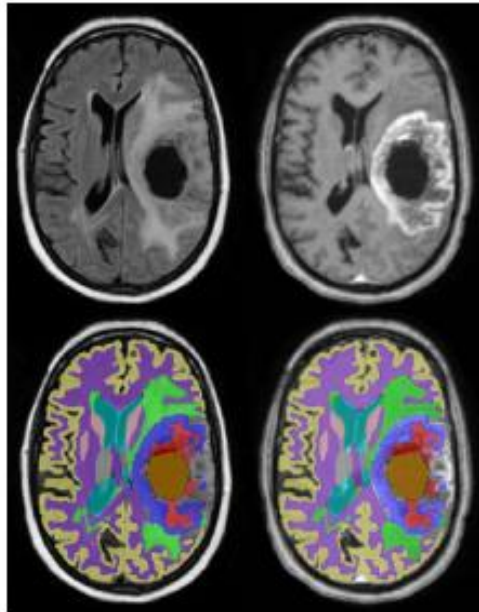
Isensee F. & al. (2019) Automated design of deep learning methods for biomedical image segmentation. arXiv

SEGMENTATION DE TOUTES LES STRUCTURES D'INTERET NÉCESSAIRES POUR WP3 ET WP4

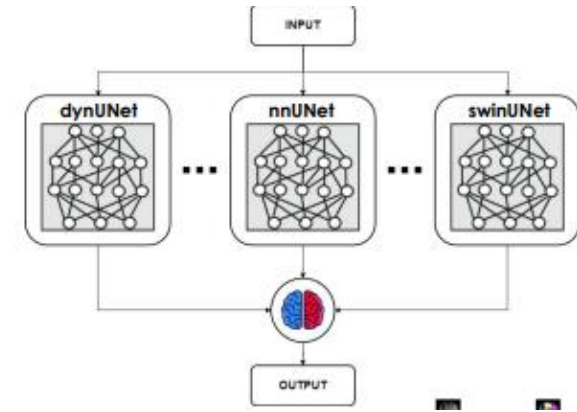
Travaux réalisés par la société Pixyl & deux neuroradiologues (Nadya Pyatigorskaya, Lucia Nichelli – ICM)

Output:

- necrosis
- contrast-enhanced
- hyperintensities
- cavity
- cortical gray matter
- white matter
- ventricles
- basal ganglia
- thalamus



- ❑ Validation de l'algorithme
 - ❑ 40 IRMS EpiBrainRad recrutés à la Pitié-Salpêtrière (avec vérité terrain)
 - ❑ Calcul du coefficient de Dice



N=40 PSL

	dynUNet	nnUNet	swinUNet
contrast_enhanced	0.67	0.64	0.64
wmh_edema	0.85	0.89	0.85
necrosis_cavity	0.65	0.64	0.65
GM	0.92	0.92	0.91
WM	0.94	0.95	0.93
basal_ganglia	0.88	0.92	0.88
thalamus	0.92	0.95	0.92
ventricles	0.95	0.95	0.94
Average	0.86		

- ✓ Difficultés de segmentation de la zone (péri)tumorale
- ✓ Perspectives:
 - ✓ Dernière itération pour amélioration des performances de segmentation
 - ✓ Appliquer à nouveau les algorithmes sur des données mieux harmonisées (WP2)
 - ✓ Validation sur les données EpiBrainRad de patients **recrutés à l'ICANS**
 - ✓ Généralisation sur des données du réseau national POLA (patients avec **oligodendrogliomes de haut-grade**)

Modélisation probabiliste et inférence bayésienne de la relation entre dose reçue au cerveau et anomalies irréversibles de la substance blanche

Travaux de thèse de Théo Silvestre encadré par Sophie Ancelet (ASNR) et Florence Forbes (INRIA)

- ▶ **Question 1:** Existe-t-il une association entre la **dose de RI** reçue par certains tissus sains intracérébraux et **la progression spatio-temporelle de plages de leucoencéphalopathie (PDLs)** au sein de la substance blanche chez des patients traités par RT pour un glioblastome?

▶ Si oui :

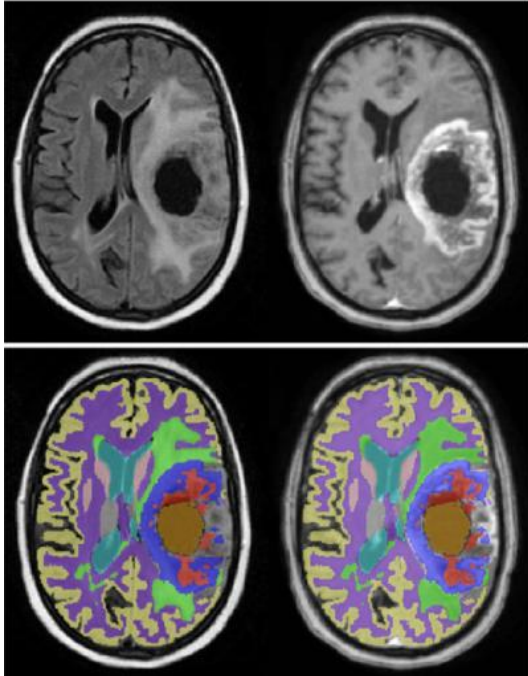
- **Question 2:** Quelle est la force de cette association?
- **Question 3:** Est-elle spatialement homogène ou hétérogène au sein de la substance blanche?
- **Question 4:** Quelle est la dynamique temporelle de cette association?

Radiosensibilité

de la substance blanche en fonction de la dose reçue, du temps, de la région anatomique?

Types de données analysées & objectifs techniques

Pour chaque patient :



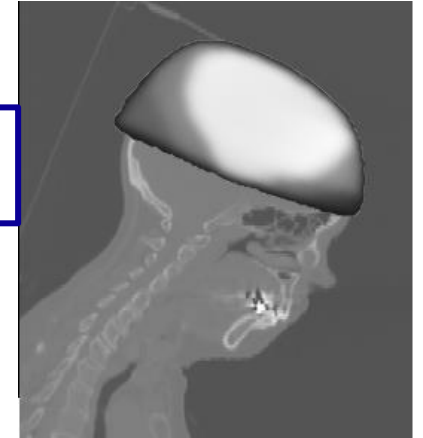
FLAIR + T1 Gd

label 1: nécrose + cavité
label 2: HSBs (œdème + PDLs)
label 3: Tumeur et zone péri-tumorale (tissu rehaussé au gadolinium)
label 4: Substance blanche
label 5: Substance grise
label 6: Ventricules
label 7: Ganglions basaux
label 8: Thalamus

FLAIR + T1 Gd
pre-segmentation

Collection longitudinale d'IRMs cérébrales segmentées par Pixyl
=> *Label de classe associé à chaque voxel*

Variables cliniques (sexe, âge, tabac, diabète, hypertension artérielle, cholestérol, etc)



Sources: travaux thèse Théo Silvestre

Scanner dosimétrique + carte de planification de dose (RTDOSE) associée
=> *Dose reçue en chaque voxel*

- ▶ Développer un **modèle probabiliste** qui permette de décrire **au niveau du voxel** l'association entre l'incidence de PDLs post-RT et la dose de RI reçue en chaque voxel
- ▶ Implémenter un **algorithme d'inférence efficace** – compte tenu du nombre important de données traitées – pour mettre en évidence (si elle existe), estimer et caractériser le plus finement possible l'association dose-réponse d'intérêt

Sélection d'une sous-cohorte de 50 patients EpiBrainRad

Critères d'inclusion retenus:

- ☐ Temps de suivi : **au moins 12 mois post-RT**
- ☐ Centre de suivi : **Hôpital de la Pitié-Salpêtrière**
- ☐ **Données d'imagerie disponibles :**
 - ☐ un CT-scan avec la **carte de planification de dose** (rtdose)
 - ☐ une IRM réalisée avant radiothérapie (à la **ligne de base**) avec les segmentations associées
 - ☐ **au moins 5 IRMs** réalisées à différents temps post-RT et allant au-delà des 12 mois post-RT avec les mêmes segmentations que pour la LB

Remarques:

- ✓ 8 patients sur 50 ont eu **plusieurs cures de RT**
 - ✓ Dernière IRM de suivi considérée = celle réalisée juste avant la 2eme cure de RT
- ✓ 3 patients sur 50 ont un **protocole de RT hypofractionné** (Dose délivrée au volume cible = 40 Grays, 15 fractions/séance)
- ✓ **Sélection des temps d'examens IRMs** les plus régulièrement espacés possibles (~ 7 mois)

Quelques statistiques descriptives

	sous-cohorte		cohorte	
	AGE	TABAC	AGE	TABAC
Min.	22.0	0.0	20.0	0.0
Median	56.0	0.0	58.0	0.0
Mean	53.7	6.1	56.7	7.5
Max.	75.0	50.0	80.0	100
NA's	0	0	9.0	41.0

	cohorte			
	DIAB	HTA	CHO	TABAC
Oui	15 (5%)	63 (23%)	47 (17%)	90 (32%)
Non	252 (91%)	204 (74%)	220 (79%)	146 (53%)
NA's	10 (4%)	10 (4%)	10 (4%)	41 (15%)

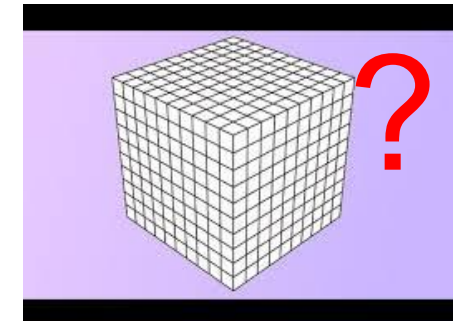
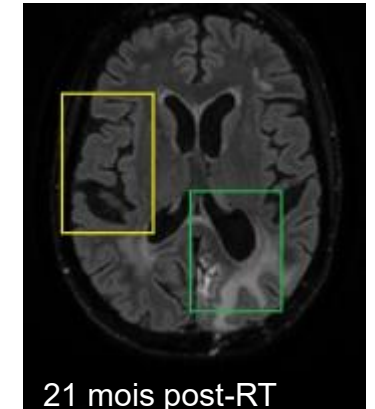
	sous-cohorte			
	DIAB	HTA	CHO	TABAC
Oui	2 (4%)	10 (20%)	8 (16%)	18 (37%)
Non	47 (96%)	39 (80%)	41 (84%)	31 (63%)
NA's	0	0	0	0

	sous-cohorte	cohorte
	SEXE	
Homme	34 (69%)	176 (64%)
Femme	15 (31%)	91 (33%)
NA's	0	9 (3%)
	Chirurgie	
Partielle	13 (27%)	58 (21%)
Complète	28 (57%)	130 (47%)
Biopsie	8 (16%)	76 (27%)
NA's	0	13 (5%)

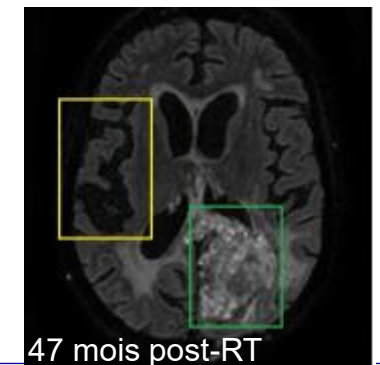
Pré-traitement indispensable des données d'imagerie

Pourquoi?

- Les IRMs segmentées par Pixyl ne sont **pas directement exploitables** pour répondre aux questions d'intérêt de la thèse
- Deux difficultés techniques :**
 - Elles ne sont **pas comparables** d'un patient à l'autre et d'un temps d'examen à l'autre
 - => Scénarii différents d'acquisition des images, déformations locales du cerveau etc...
 - => Pas de grille de voxels commune permettant une étude populationnelle
 - Elles ne permettent pas de distinguer les HSBs associées à de l'oedème de celles associées à des PDLs => Pas d'accès direct à la variable-réponse d'intérêt



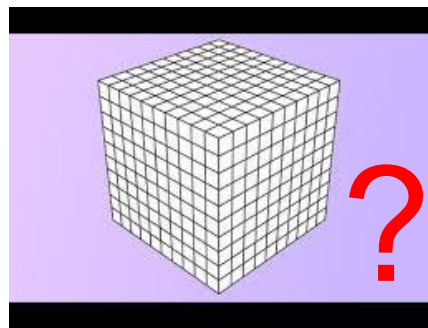
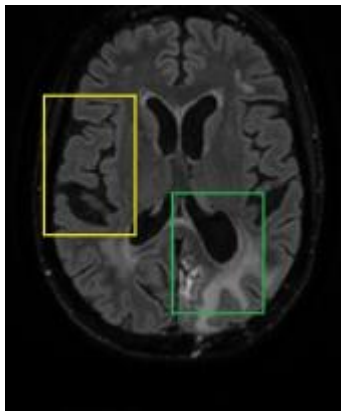
Matrice de voxels commune



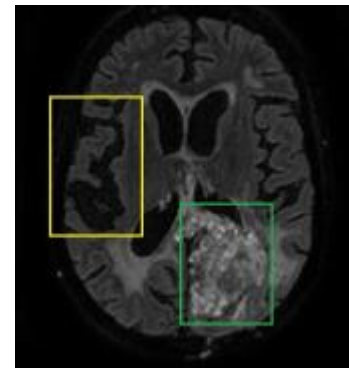
Sources: travaux thèse Théo Silvestre

Pipeline de pré-traitement automatique des données d'imagerie (I)

- **Etape 1: Recalage rigide intra et inter-patients des données d'imagerie**
- Recalage de toutes les IRMs cérébrales d'un même patient sur une IRM de référence (premier temps post-RT) : Utilisation de la routine de **recalage rigide automatisée** proposée dans le cadre du WP3 [Clara Trompette, Anton François, Alain Trouvé]
- Recalage du CT-scan de chaque patient sur son IRM de référence
- Recalage des IRMs de référence de tous les patients sur un cerveau template



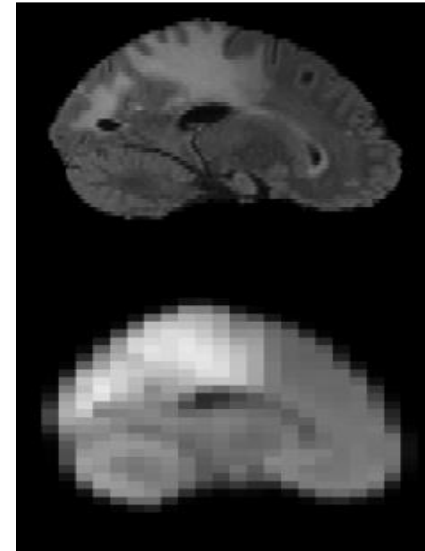
Matrice de voxels commune



Sources: travaux thèse Théo Silvestre

Pipeline de pré-traitement automatique des données d'imagerie (II)

- Parer « au mieux » aux limites du recalage rigide
- Etape 2 : Réduction de la résolution de l'image
 - $512 \times 512 \times 128 \Rightarrow 32 \times 32 \times 32$ (bon compromis temps de calculs/perte d'informations)
 - Conservation des supervoxels formant le masque du cerveau $\rightarrow$ 6763 supervoxels
 - Suppression des supervoxels du cervelet et du tronc cérébral $\rightarrow$ 4789 supervoxels

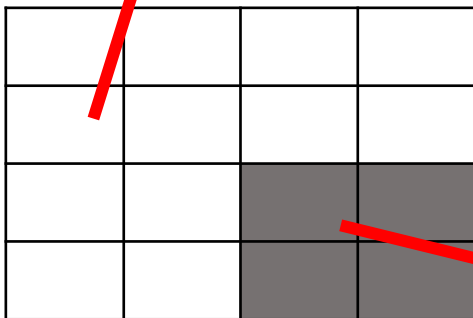


Sources: travaux thèse Théo Silvestre

Unité statistique $\Rightarrow$ Supervoxel ($N=1024$ voxels concaténés)

Voxels **susceptibles de contenir une HSB**
(substance blanche)
Leur nombre est **noté** H_v^t

Illustration
en 2D



Supervoxel v au temps d'examen t

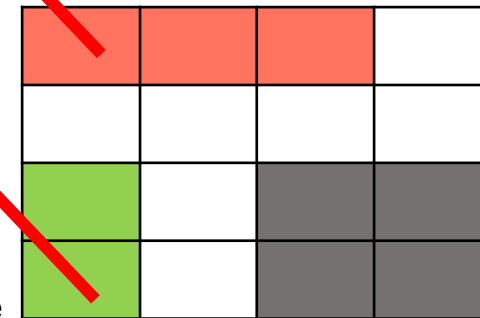
$$N = H_v^t + R_v^t$$

Voxels qui **ne peuvent pas**
contenir d'HSBs
(ganglions basaux,
substance grise,
ventricules, ...)

Leur nombre est **noté** R_v^t et
appelé le reste

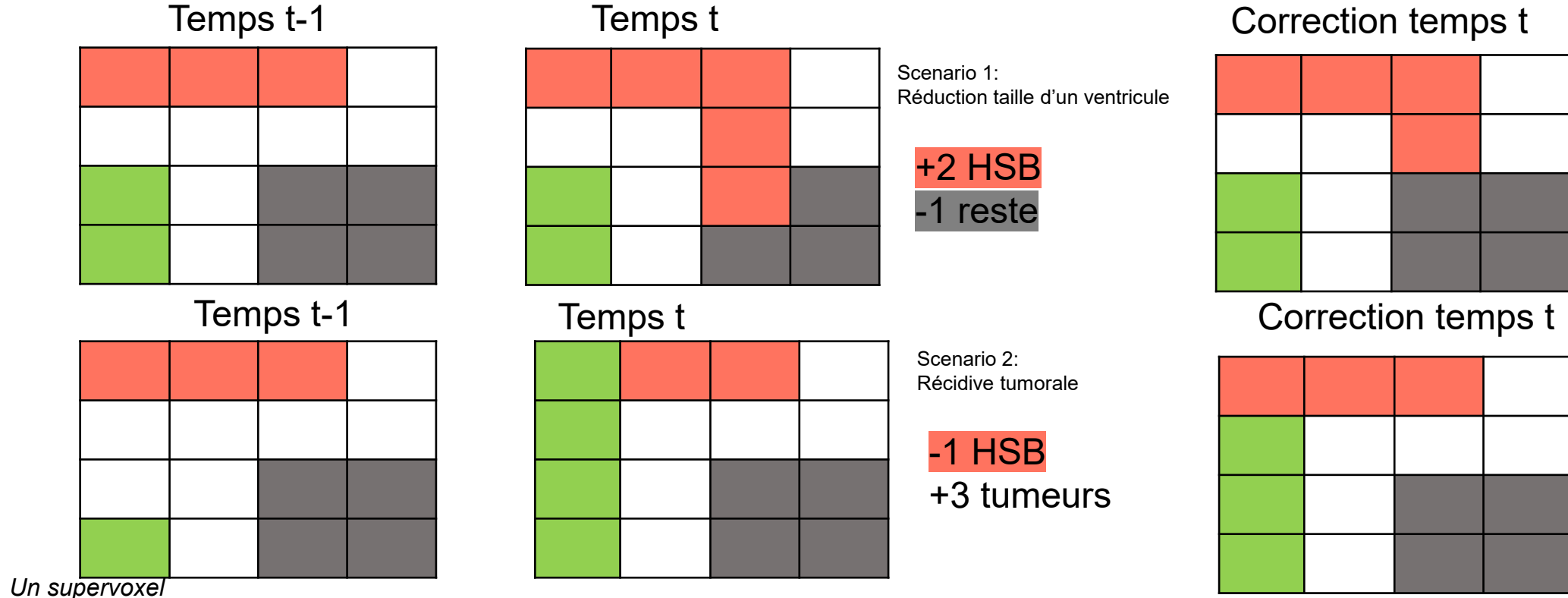
« labellisé » HSB

« labellisé »
Tumeur active
ou Nécrose
ou Cavité
Leur nombre est noté
 A_v^t



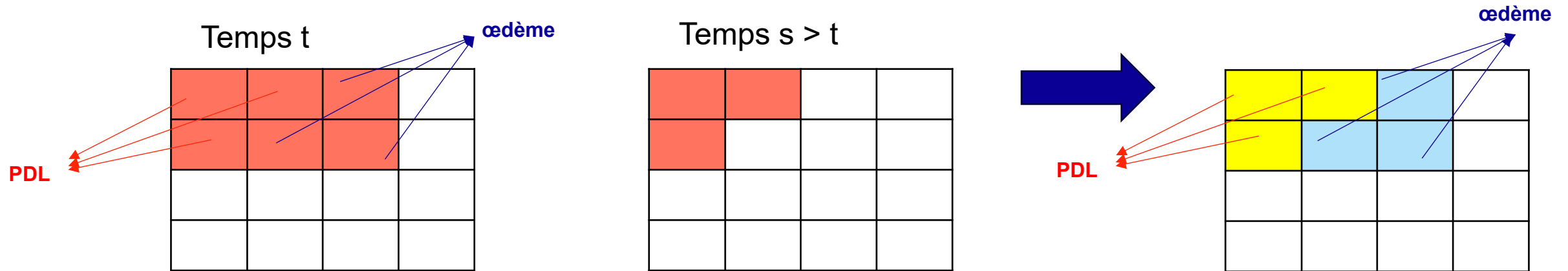
Pipeline de pré-traitement automatique des données d'imagerie (III)

- Parer « au mieux » aux limites du recalage rigide
- Etape 3 : Correction longitudinale des labels des HSBs en procédant de manière récursive sur tous les temps d'examens
- Prise en compte d'artefacts liés au déplacement de la matière dans le temps => Différents scénarii considérés



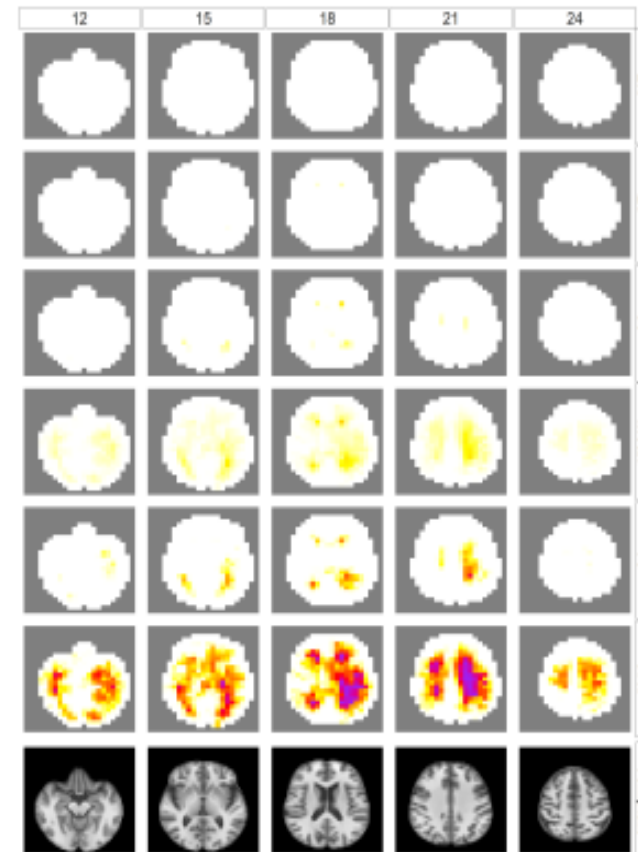
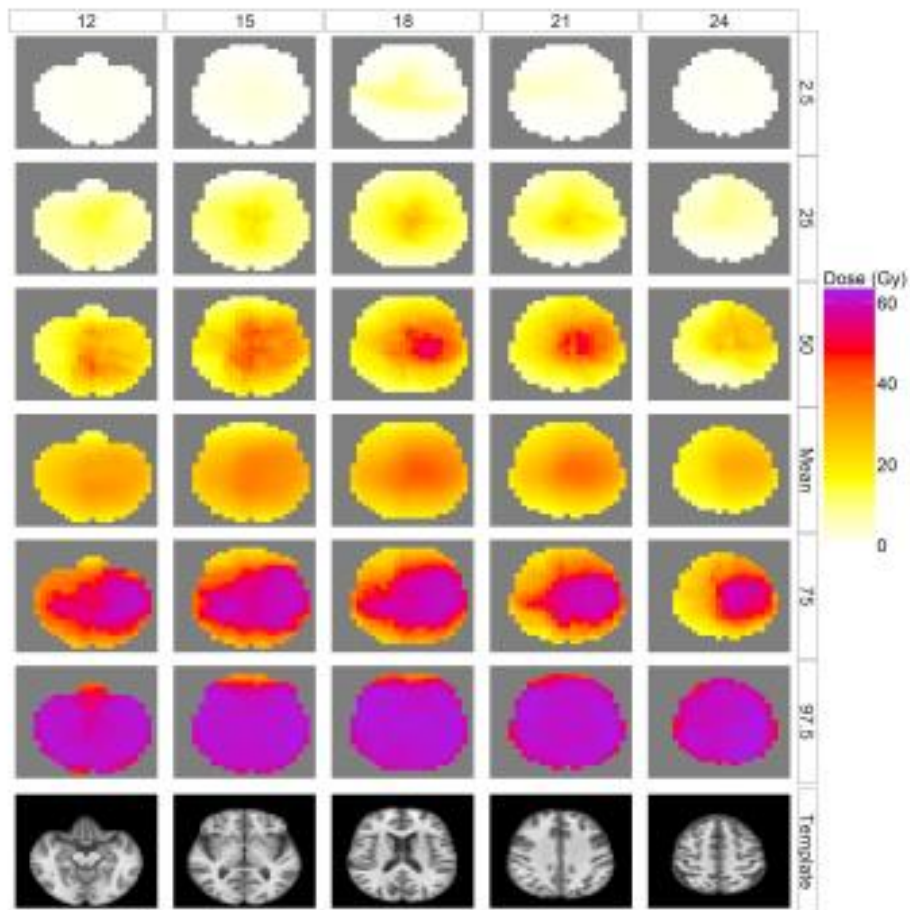
Pipeline de pré-traitement automatique des données d'imagerie (V)

- Parer « au mieux » aux limites du recalage rigide
- Etape 4 : Distinction entre les PDLs des œdèmes au sein des HSBs
- Propriété de croissance et d'irréversibilité des PDLs (alors que les œdèmes finissent par se résorber)

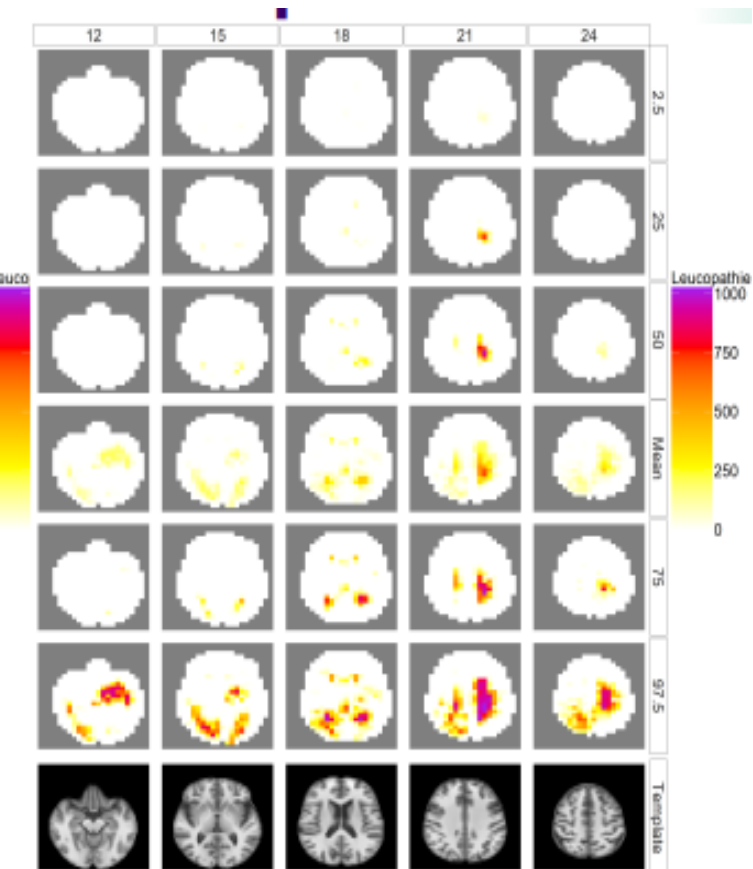


Limite : l'identification d'un œdème nécessite de le voir se résorber dans le temps
-> Suppression des derniers mois de suivi

Distribution de la dose (gauche) et du nombre de leucopathies par supervoxel (droite) dans la cohorte EpiBrainRad



12 mois post-RT



24 mois post-RT

MODÉLISATION DE L'ASSOCIATION ENTRE DYNAMIQUE D'APPARITION DE LEUCOPATHIES POST-RT ET DOSE DELIVREE AU CERVEAU

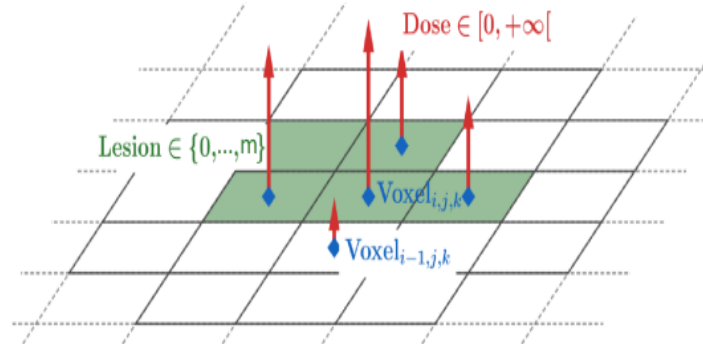
DONNÉES D'ENTREE SUR GRILLE COMMUNE DE VOXELS

Individu n

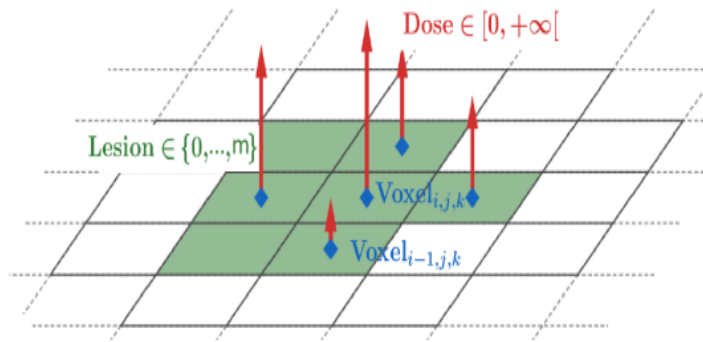
•

•

Individu 2



Individu 1

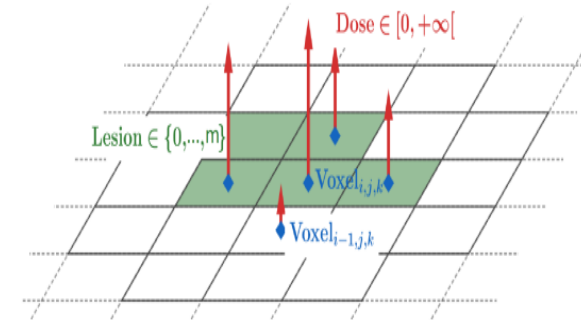


Examen T_i

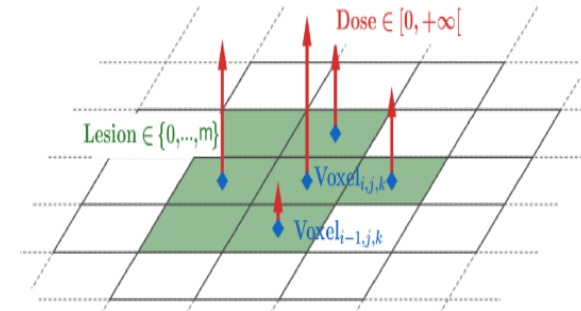
•

•

Examen 2



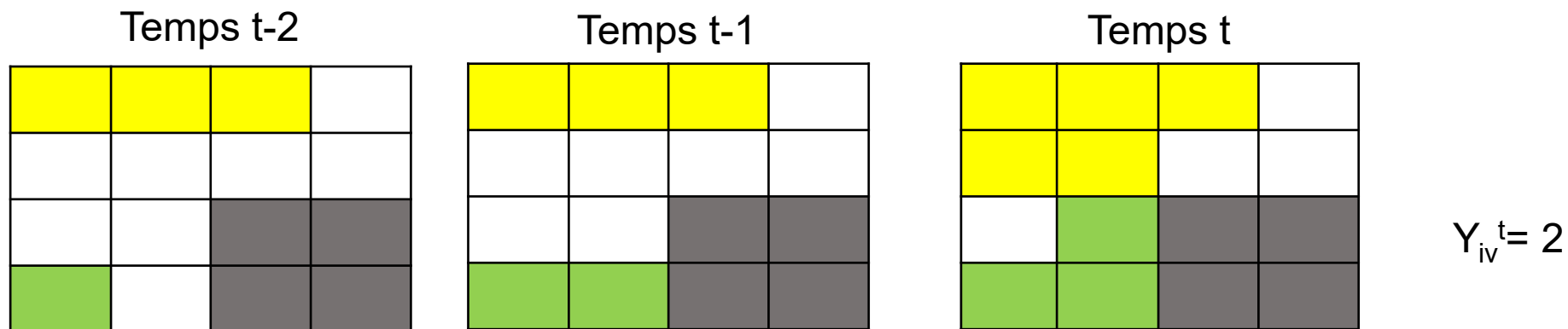
Examen 1



Pour chaque patient

MODÉLISATION DE L'ASSOCIATION ENTRE DYNAMIQUE D'APPARITION DE LEUCOPATHIES POST-RT ET DOSE DELIVREE AU CERVEAU

- Variable-réponse : Accroissement Y_{iv}^t du nombre de PDLs post-RT pour le patient i au supervoxel v pendant un intervalle de temps régulier inter-examens $]t-1, t]$.



- Soit C_v la classe d'appartenance (**inconnue!**) du supervoxel v (supposée identique pour tous les individus et tous les temps)
- Soit M_{iv}^t le nombre maximum de voxels « non lésionnés » au temps $t-1$ du supervoxel v qui peuvent devenir le siège d'une PDL au temps t
 - Dans l'exemple ci-dessus: $M_{iv}^t = 7$
 - $M_{iv}^t = M_{iv}^{t-1} - Y_{iv}^{t-1} - \Delta A_{iv}^{t-1} - \Delta R_{iv}^{t-1}$
 - Dans l'exemple ci-dessus: $M_{iv}^t = 8 - 0 - 1 - 0 = 7$

UN MODÈLE DE MÉLANGE DE RÉGRESSIONS BINOMIALES

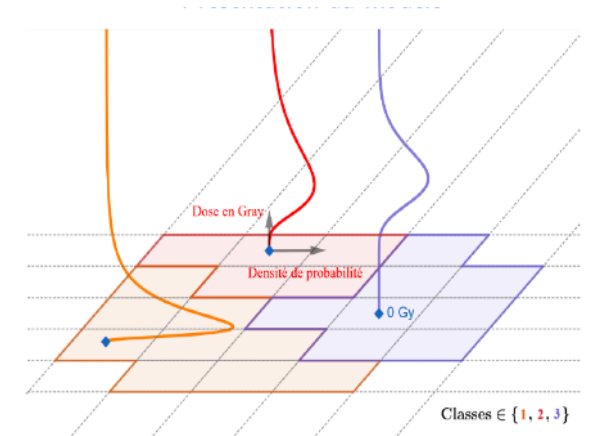
MODÈLE D'OBSERVATIONS

- Conditionnellement à sa classe C_v et à m_{iv}^t , on décrit l'accroissement de PDLs Y_{iv}^t comme la régression binomiale suivante:

$$Y_{iv}^t \mid m_{iv}^t, C_v = k, \alpha, \beta \sim \mathcal{B}(m_{iv}^t, p_{ivk}^t)$$

$$p_{ivk}^t = \text{logit}^{-1}(\eta_{ivk}^t)$$

$$\eta_{ivk}^t = \mathbf{x}_{iv}^{t \dagger} \alpha_k + \mathbf{w}_{iv}^{t \dagger} \beta.$$



- P_{ivk}^t désigne la **probabilité d'apparition** d'un nouveau voxel qui soit le siège d'une leucopathie au sein du supervoxel v contenu dans la classe k , dans l'intervalle de temps $[t-1; t]$ et pour le patient i
- $\mathbf{X}_{iv}^t$ vecteur de p covariables patient/temps spécifique qui peuvent avoir un **effet différent** $\alpha_k \in \mathbb{R}^p$ (inconnu!) sur P_{ivk}^t en fonction de la classe d'appartenance de v
- Typiquement, la dose !
- $\mathbf{W}_{iv}^t$ vecteur de q covariables patient/temps spécifique qui ont un **effet identique** $\beta \in \mathbb{R}^q$ (inconnu!) sur la probabilité P_{ivk}^t quelle que soit la classe d'appartenance de v

UN MODÈLE DE MÉLANGE DE RÉGRESSIONS BINOMIALES COUCHE LATENTE

- Soit $C=\{C_v, v \in V\}$ le vecteur des labels de classe à déterminer
- Hypothèse d'indépendance inter-voxels des labels de classe
- Nombre de classes fixé à K
- Loi de probabilité discrète à estimer :

$$p(c|\pi) = \prod_{v=1}^V p(c_v|\pi), \quad p(c_v = k|\pi) = \pi_k, \quad \forall k = 1 \dots K$$

avec $\pi_k \in [0, 1]$ et $\sum_{l=1}^K \pi_l = 1$, et

CHOIX DES LOIS A PRIORI

- Choix de **lois a priori** vagues sur les paramètres inconnus à estimer

$$\alpha_k \sim \mathcal{N}_p \left(\mu_{\alpha_k} = \mathbf{0}, \Sigma_{\alpha_k} = 10^3 I_p \right)$$

$$\beta \sim \mathcal{N}_q \left(\mu_{\beta} = \mathbf{0}, \Sigma_{\beta} = 10^3 I_q \right) .$$

$$\pi = \{\pi_1, \dots, \pi_K\} \sim \text{sym-Dir}(1/2)$$

INFÉRENCE BAYÉSIENNE DU MODÈLE

► Modèle complexe et coûteux à inférer

Algorithme d'optimisation variationnel

Approximation de la loi a posteriori jointe via une approche champ moyen

$$q(\alpha, \beta, \pi, C, U) = q_{\pi}(\pi_1, \dots, \pi_K) q_{\alpha, \beta}(\alpha, \beta) \prod_{v=1}^V q_{C_v}(C_v) \prod_{v=1}^V \prod_{i=1}^n \prod_{t=0}^{T_i} q_{U_{iv}^t}(U_{iv}^t)$$

Cette méthode est connue pour **sous-estimer la variance a posteriori**.

Mise-à-jour séquentielle :

$$\text{VB-E-}C_v : q_{C_v}^{(r)} \propto \exp \mathbb{E}_{q_{\alpha, \beta}^{(r-1)}, q_{U_{iv}^t}^{(r-1)}, q_{\pi}^{(r-1)}} [\log p(Y, U, C, \alpha, \beta, \pi)]$$

$$\text{VB-E-}U_{iv}^t : q_{U_{iv}^t}^{(r)} \propto \exp \mathbb{E}_{q_{\alpha, \beta}^{(r-1)}, q_{C_v}^{(r-1)}} [\log p(Y, U, C, \alpha, \beta)]$$

$$\text{VB-E-}(\alpha, \beta) : q_{\alpha, \beta}^{(r)} \propto \exp \mathbb{E}_{q_C^{(r-1)}, q_U^{(r-1)}} [\log p(Y, U, C, \alpha, \beta)]$$

$$\text{VB-E-}\pi : q_{\pi}^{(r)} \propto \exp \mathbb{E}_{q_C^{(r-1)}} [\log p(C, \pi)]$$

Algorithme stochastique Monte-Carlo par Chaînes de Markov

Metropolis-Within-Gibbs Adaptatif

- α, β : Random Walk Metropolis Hastings (RWMH [Metropolis et al., 1953, Hastings, 1970]),
- π : Gibbs sampling,
- C_v : Gibbs sampling,

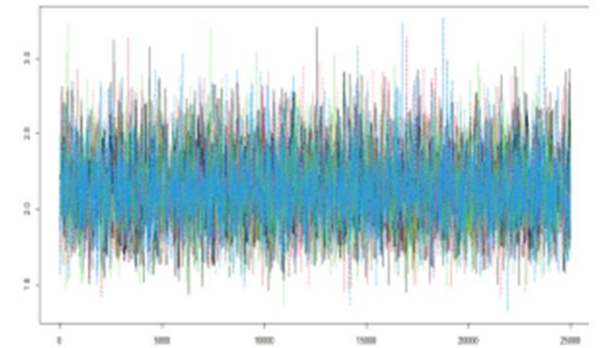
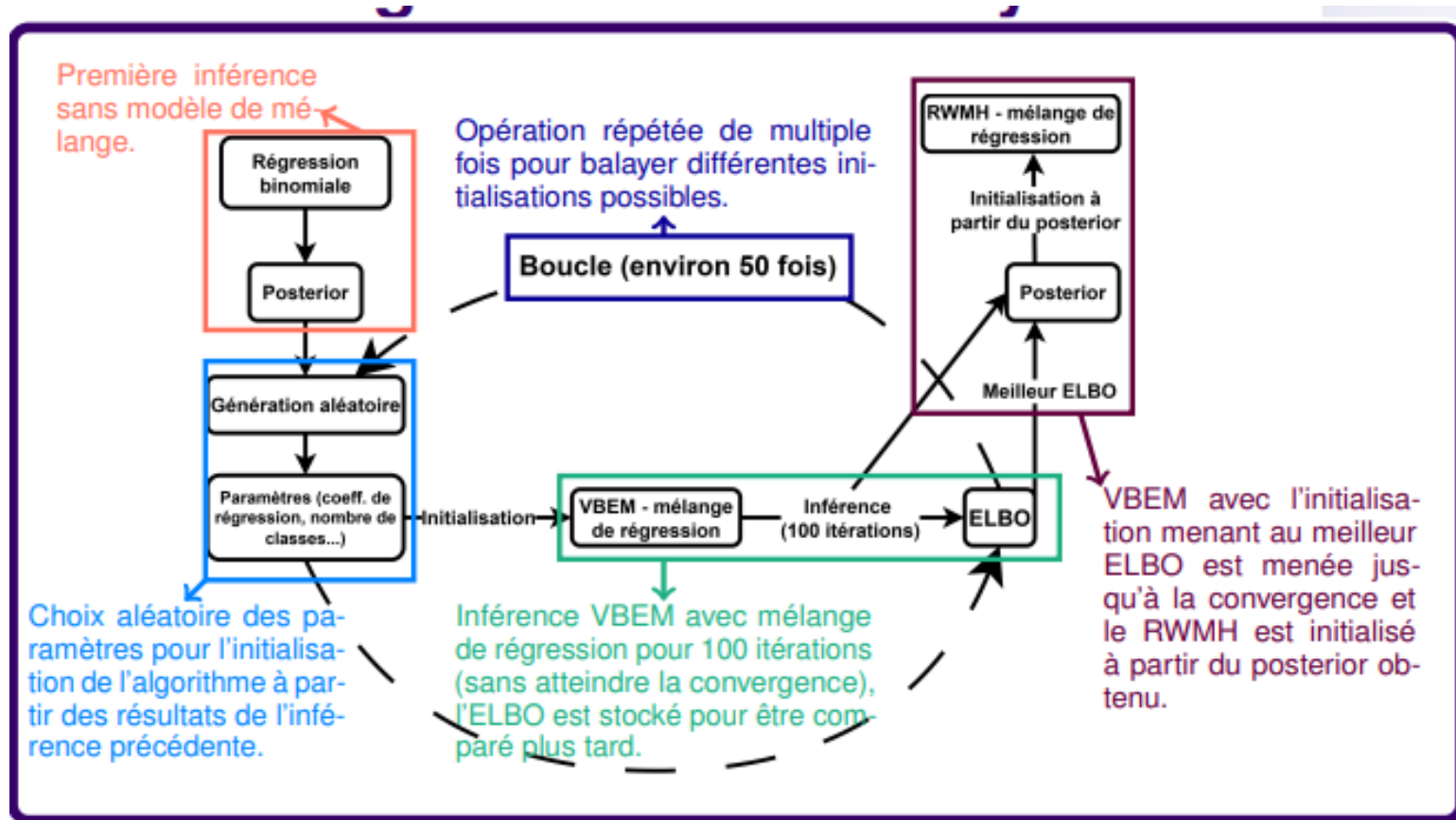


Figure – Représentation d'une chaîne de Markov.

=> Difficulté: Sensibilité à l'initialisation des deux algorithmes

INFÉRENCE BAYÉSIENNE DU MODÈLE

- Implémentation en C++ d'une stratégie d'inférence bayésienne basée sur **le couplage d'un algorithme variationnel et d'un algorithme Monte-Carlo par Chaînes de Markov**
- Travail important pour la minimisation des temps de calculs



RÉSULTATS

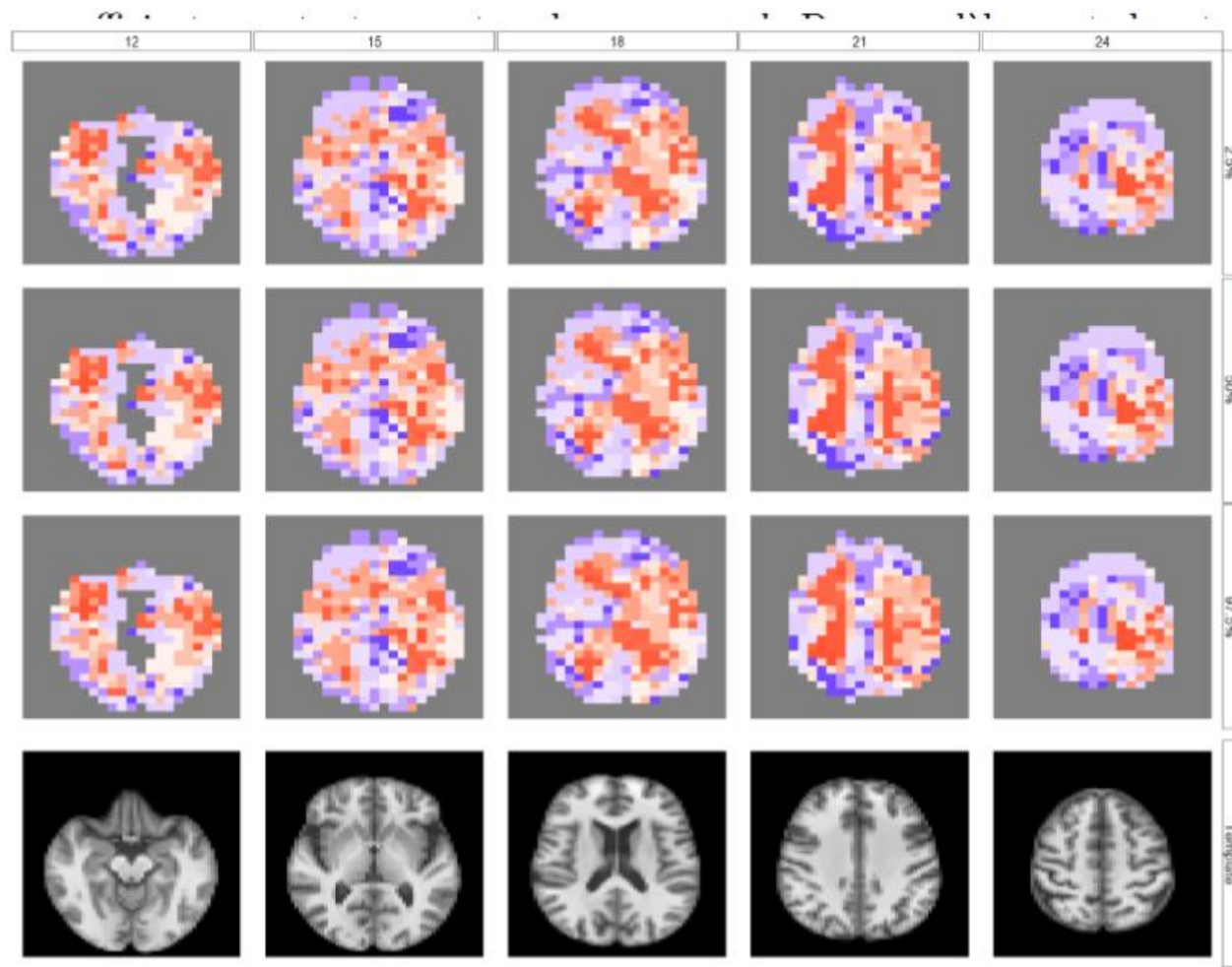
COMPARAISON DE MODÈLES BASES SUR DIFFÉRENTES CONFIGURATIONS DE COVARIABLES

- Un intercept (Int) qui peut varier ou non en fonction de la classe
- Effet linéaire ou linéaire quadratique:
 - De la dose par fraction **d** ajustée par le nombre de fractions **nf**
 - De la dose totale **D** ou de $D+D2/nf = D+Dd$ dans le cas linéaire quadratique [Jones et al., 2001]
- Temps d'examen **t** considéré en interaction ou non avec la dose (.t)
- Covariables individu-spécifiques : **A**=Age , **S**=Sexe, **H**=Hypertension artérielle , **T**=Tabac , **C**=Cholesterol

	Coeff. classes								Coeff. constant						nb. Cl.	WAIC $\times 10^7$		
13	d								Int	nf				24	2.244			
12			D			Dd			Int					29	2.193			
11	d	d2		d.t	d2.t				Int	t	nf			7	2.182			
10	d	d2							Int		nf			19	2.176			
9	d			d.t					Int	t	nf			45	2.169			
8	d							Int			nf			17	2.161			
7			D			Dd		Int						16	2.149			
6			D			Dd	Dd.t	D.t	Int	t	A	S	H	T	C	4	2.148	
5	d			d.t					Int	t	nf			16	2.128			
4	d	d2							Int		nf			13	2.122			
3			D			Dd	Dd.t	D.t	Int	t				36	2.116			
2	d	d2		d.t	d2.t			Int		t	nf			13	2.091			
1	d	d2		d.t	d2.t				Int	t	nf	A	S	H	T	C	18	2.082

- ✓ Grande hétérogénéité du nombre de classes estimées selon les covariables considérées
- ✓ **Modèle 1 : Modèle avec le plus petit critère WAIC (Watanabe Information Criterion)**

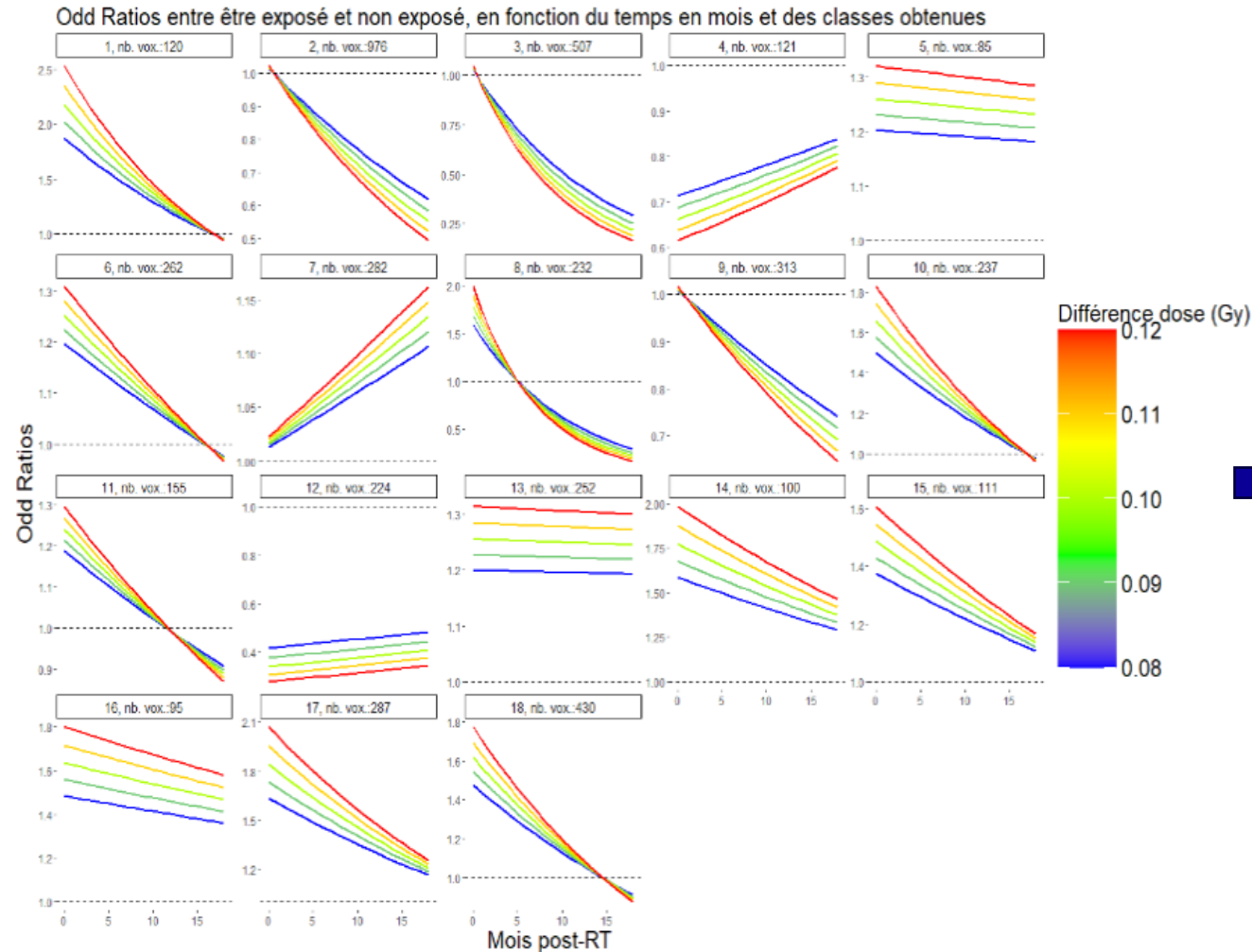
RÉSULTATS



- ✓ Relation dose-réponse strictement positive autour des ventricules notamment
- ✓ Hétérogénéité spatiale de la force de l'association

FIGURE 5.1 – Modèle 1 : répartition au niveau du *supervoxel* d'ORs calculés pour un passage de 0 Gy à 0.1 Gy relatif à l'apparition d'une nouvelle leucopathie dans l'intervalle de temps $]0, 6]$ mois post-RT. Coupes axiales (12,15,18,21,24) en abscisse et quantiles de la distribution des ORs en ordonnée.

RÉSULTATS



- ✓ La dynamique temporelle de l'association dose-réponse d'intérêt dépend de la classe
- ✓ Relation croissante ou décroissante avec le temps selon la classe

Modèle 1: Evolution temporelle des Odd-Ratios (ORs) estimés dans chaque classe sur 18 mois post-RT

PERSPECTIVES AUX TRAVAUX DE THÉO SILVESTRE

- ☐ Considérer des données d'entrée segmentées, pré-harmonisées (travaux en cours WP2) et avec un recalage amélioré (travaux en cours WP3)
- ☐ Appliquer le modèle sur davantage de patients
- ☐ Appliquer le modèle sur des zones spécifiques du cerveau – zone périventriculaire
- ☐ Améliorer l'algorithme d'inférence bayésienne
- ☐ Complexifier le modèle de mélange :
 - ☐ Ajouter la prise en compte de dépendances spatiales entre voxels voisins
 - ☐ Ajouter la prise en compte d'effets aléatoires individus-spécifiques
- ☐ Modélisation de la relation dose, PDLs post-RT et troubles cognitifs

MOT DE LA FIN...

Avant de pouvoir faire de la modélisation mathématique ou d'utiliser/développer des outils d'IA à partir de données d'imagerie cérébrale recueillies lors du suivi clinique de patients, un gros travail de tri, d'harmonisation et de recalage des images est indispensable ...

MERCI DE VOTRE ATTENTION!

RADI  AIDE

