

# TABLE DES MATIERES

<b>ABRÉVIATIONS .....</b>	<b>12</b>
<b>RÉSUMÉ .....</b>	<b>13</b>
<b>INTRODUCTION.....</b>	<b>14</b>
<b>MATERIELS ET METHODES.....</b>	<b>16</b>
Acquisition d'images et éthique.....	16
Génération de pseudo-étiquettes et segmentation supervisée de lésions pulmonaires .....	16
Validation du modèle à l'aide de comparaisons de volumes avec 4 évaluations manuelles et leur impact sur les données cliniques .....	19
Analyse statistique .....	20
<b>RÉSULTATS.....</b>	<b>21</b>
Segmentation sémantique des lésions pulmonaires par apprentissage profond (COViaX).....	21
La qualité des masques COViaX et la confiance du radiologue dans les volumes calculés .....	22
L'influence des masques COViaX sur les évaluations manuelles des volumes .	22
Conformité des volumes COViaX avec ceux obtenus manuellement et son intérêt en tant que biomarqueurs de diagnostic et de sévérité .....	24
<b>DISCUSSION.....</b>	<b>27</b>
Masque COViaX pour la segmentation des lésions pulmonaires.....	27
COViaX comme outil pour le triage des patients COVID-19. ....	28
COViaX comme outil de diagnostic immédiat du COVID-19.....	29
L'utilisation de COViaX à l'hôpital.....	29
<b>CONCLUSION .....</b>	<b>31</b>
<b>REFERENCES.....</b>	<b>32</b>
<b>TABLES.....</b>	<b>36</b>
Table 1: Caractéristiques des patients .....	36
Table 2: Évaluation qualitative de l'IA selon différentes machines .....	37
Table 3: Performances cliniques de COViaX et comparaisons inter-observateurs pour le diagnostic et la sévérité.....	38

<b>FIGURES .....</b>	<b>39</b>
Figure 1 : Schéma de la méthode de segmentation sémantique .....	39
Figure 2 : Exemple de segmentation et de reconstruction 3D.....	39
Figure 3 : Valeur ajoutée de l'évaluation manuelle du volume à l'aide des masques COViaX.....	39
Figure 4 : Comparaison des volumes de l'IA et de ceux de 4 observateurs : reproductibilité et leur comparaison avec les données cliniques.....	39

# ABRÉVIATIONS

**AUC** : aire sous la courbe

**CNN** : réseau neuronal convolutif

**Cs** : consolidation

**CT** : tomographie informatisée

**DL** : deep-learning (apprentissage profond)

**GGO** : opacités en verre dépoli

**IA** : intelligence artificielle

**ICC** : coefficients de corrélation intraclasse

**Msk** : COViaX Masques

**Obs-Msk** : Observateurs-Masks

**ROC** : Receiver operating characteristic= fonction d'efficacité du récepteur

**RT-PCR** : transcription inverse puis amplification en chaîne par polymérase

**SARS-CoV-2** : severe acute respiratory syndrome coronavirus 2 = syndrome respiratoire aigu sévère à coronavirus 2

**Se** : sensibilité

**Sp** : spécificité

**VR** : la quantification du rapport volume affecté/volume total

**VRall** : rapport volume pulmonaire affecté/total pour toutes les lésions présumées de SARS-CoV2 sur le scanner à faible dose

**VRcs** : rapport volume affecté/volume total pour les lésions de consolidation (condensations)

**VRggo** : rapport volume affecté/volume total pour les opacités en verre dépoli.

# RÉSUMÉ

**Contexte** : La première vague épidémique de COVID-19 s'est propagée depuis la Chine et a touché presque tous les pays du monde. La détection et l'évaluation de la gravité de la pneumopathie du SRAS-CoV-2 à l'aide de la tomodensitométrie à faible dose se sont avérées rapides et sensibles.

**Objectif** : Construire et valider un outil de segmentation automatique par intelligence artificielle (IA) pour automatiser la quantification du rapport volume affecté/volume total (VR) chez les patients suspectés de pneumopathie SARS-CoV2.

**Matériels et Méthodes** : Entre le 9 mars et le 20 avril 2020, 3422 scanners thoraciques à faible dose ont été réalisées dans notre institution pour la gestion du COVID. Un réseau neuronal circonolutif tridimensionnel a été appliqué à un ensemble de 234 scanners d'entraînement pour construire l'outil d'IA. Le logiciel d'IA a été testé sur 60 examens scannographiques indépendants sélectionnés de manière aléatoire à partir de 3 scanners différents. Le rapport des volumes d'atteinte à SARS COV2 sur le volume pulmonaire total (VR) ont été automatiquement segmentés.

Tout d'abord, la tomodensitométrie sans IA a été évaluée en aveugle et de manière indépendante par 4 experts pour déterminer quantitativement les opacités en verre dépoli (VR<sub>ggo</sub>), les lésions de consolidation (VR<sub>cs</sub>) et la somme des deux (VR<sub>all</sub>). Une semaine plus tard, les experts ont à nouveau évalué les différents volumes d'atteinte à l'aide des masques d'IA de manière qualitative. Les volumes donnés par l'IA ont été comparés aux volumes des experts pour vérifier leur concordance. Les VR ont été testés comme biomarqueur de la maladie (prélèvements RT-PCR) et de la sévérité (dépendance à l'oxygène). Une comparaison statistique avec la RT-PCR et la sévérité a été effectuée. L'accord intra- et inter-observateur pour les VR et comparé à ceux des experts.

**Résultats** : Les volumes générés par l'IA ont démontré un haut niveau de précision. Les VR de l'IA étaient comparables à ceux des experts. Les VR des experts étaient plus reproductibles avec les masques de l'IA. Les performances de l'IA pour le diagnostic et la sévérité étaient au moins équivalentes à celles des experts.

**Conclusion** : L'outil de segmentation d'IA et les volumes calculés de notre logiciel COViaX sont fiables et donnent un biomarqueur robuste de la sévérité de la maladie.

# INTRODUCTION

La nouvelle maladie à coronavirus (SARS-CoV-2/COVID-19) signalée en décembre 2019 à Wuhan, en Chine, a été désignée comme pandémie en mars 2020 par l'Organisation mondiale de la santé (OMS). Au 8 mars 2021, la maladie a causé 2 587 225 décès dans le monde et 116 363 935 patients infectés, selon l'OMS (1). Sans mesures de confinement et de distanciation sociale, le taux de reproduction de la maladie COVID-19 ( $R_0$ ) est estimé à environ 2,2 (2). En l'absence de traitement et en raison de la vitesse de propagation, des mesures drastiques d'éloignement social, aussi extrêmes que des fermetures nationales répétées, ont été appliquées dans le monde entier pour contenir l'épidémie. Alors que des vaccins sont actuellement déployés dans le monde entier et que les variants émergents semblent être plus contagieux (3), le suivi de l'épidémie reste plus que jamais crucial et nécessite des outils de diagnostic robustes (4). Une combinaison de signes cliniques (par exemple, toux et fièvre) et une analyse moléculaire obtenue par PCR à partir d'un écouvillon nasopharyngé est utilisée pour un diagnostic généralisé.

Les examens par scanners constituent une approche d'imagerie performante pour cartographier les lésions cliniques d'organes entiers, ici les poumons des patients. Les examens tomodensitométriques des poumons présentent des caractéristiques très suggestives (5–7), notamment des opacifications bilatérales dites "en verre dépoli" (GGO) avec une distribution périphérique et postérieure. Des images de consolidation (Cs) sont ensuite observées au cours de l'évolution de la maladie (8–10). L'aspect de crazy-paving intra-lobulaire (épaississement des septa lobulaire avec remplissage alvéolaire) est moins fréquent. Par conséquent, compte tenu de la sensibilité et de la spécificité du scanner, la pneumonie SARS-CoV2 peut être découverte de manière fortuite. De plus, comme les anomalies de la tomodensitométrie peuvent précéder la positivité de la RT-PCR chez les patients symptomatiques et ceux qui ne présentent pas de symptômes (11). Le dépistage immédiat par tomodensitométrie a été utilisé dans certains centres pour limiter la contagion intra-hospitalière, notamment avant l'entrée dans les salles d'opération (12) dans le contexte de l'urgence et de faible disponibilité des tests RT-PCR.

Le scanner thoracique est également un outil efficace pour évaluer l'étendue anatomique de la pneumopathie du SRAS-CoV-2 (13). En effet, la gravité de l'infection à coronavirus est surtout liée à une pneumonie conduisant à un syndrome de détresse respiratoire aiguë (14). Sa survenue est actuellement imprévisible. Une évaluation tomodensitométrique quantitative de l'étendue des lésions liées au SRAS CoV-2 dans les reconstructions de volumes pulmonaires pourrait améliorer de façon décisive la qualité du diagnostic. Cependant, les radiologues procèdent actuellement à cette évaluation manuellement, généralement par inspection visuelle des coupes transversales de la tomodensitométrie, ce qui fait de la reconstruction 3D des volumes des lésions une tâche difficile et peu fiable. Des travaux récents suggèrent que la quantification automatique des lésions basée sur l'apprentissage profond en intelligence artificielle peut résoudre efficacement la segmentation automatique de manière supervisée ou faiblement supervisée (15–17). L'IA sur les tomodensitométries pulmonaires pourrait donc réduire le temps de diagnostic, augmenter la fiabilité de la quantification des volumes lésionnels et ainsi minimiser l'impact de la subjectivité dans les interprétations radiologiques.

Dans ce travail, nous avons développé une méthode de segmentation basée sur l'apprentissage profond (DL), COViaX, conçue pour la segmentation volumétrique des opacités en verre dépoli (GGO) et des consolidations (Cs) dans des séries de tomodensitométries pulmonaires, permettant des reconstructions volumiques 3D précises des lésions. La méthode a d'abord été validée par une analyse d'accord inter-observateur où la qualité de la segmentation automatique a été comparée aux interprétations des radiologues. Ensuite, la performance de COViaX a été évaluée en tant que biomarqueur pour le diagnostic du SRAS-CoV-2 et en tant que prédicteur de la gravité de la maladie.

Nous concluons que la quantification des lésions dans la pneumonie SARS-CoV-2 peut être réalisée avec un cadre d'IA faiblement supervisé permettant une segmentation robuste et qualitative des lésions pulmonaires qui pourrait grandement aider les radiologues.

# MATERIELS ET METHODES

## Acquisition d'images et éthique

Cette étude clinique rétrospective a été approuvée par le comité d'éthique institutionnel (IRB : CRM-2004-069) et le règlement général sur la protection des données selon les directives de la déclaration d'Helsinki III.

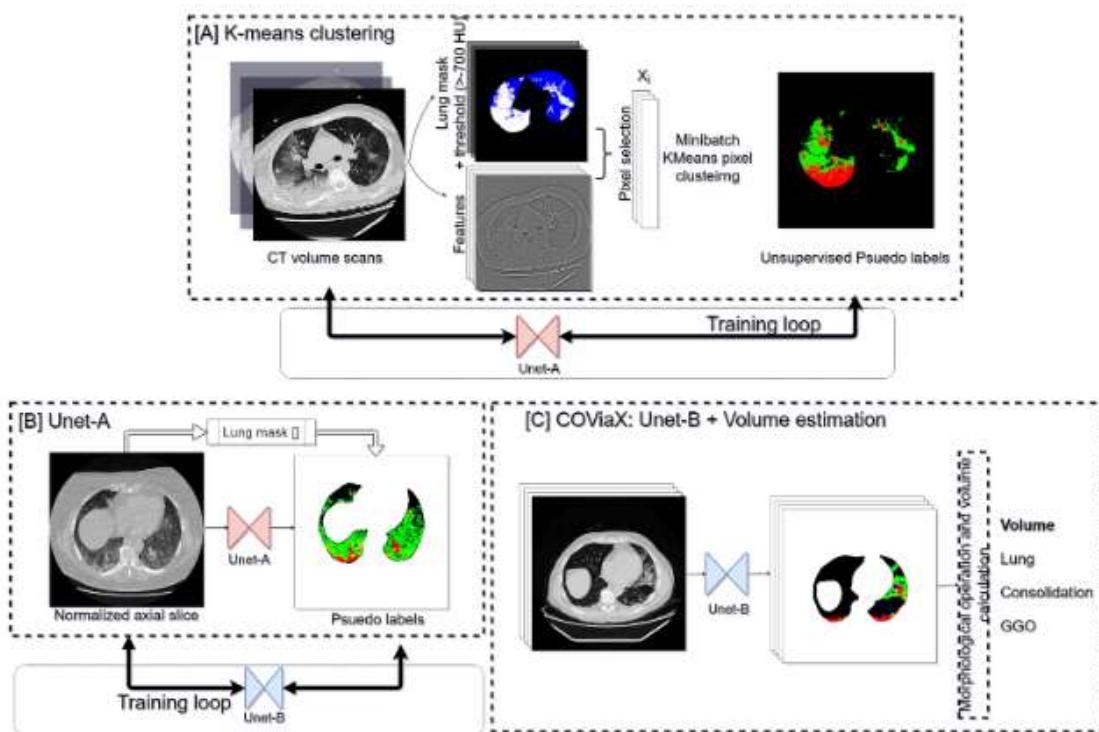
Du 9 mars 2020 au 20 avril 2020, 3422 tomographies thoraciques à faible dose ont été réalisées sur 3 scanners différents.

L'ensemble des examens tomodensitométriques à ultra-faible dose (DLP < 45 mGy.cm) ont été réalisés en acquisition hélicoïdale crano-caudale avec des paramètres fixes selon les recommandations industrielles sur 3 appareils différents : Aquilion Prime (Toshiba Medical Systems, Otawara, Japon), détecteurs  $80 \times 0,5$  mm, collimation 40 mm, temps de rotation 0,5s, pas de 1,388 mm/ rotation, tension du tube 135 kVp, intensité 20mAs, filtre pulmonaire), Revolution EVO (GE Healthcare, Milwaukee, WI, USA), détecteurs  $64 \times 0,625$  mm, collimation 40 mm, temps de rotation 0,35s, pas de 1,375 mm/rotation, tension du tube 120 kVp, intensité de 45mAs, filtre pulmonaire, Discovery CT 750 HD (GE Healthcare), 64 détecteurs de 0,625 mm, collimation 40 mm, temps de rotation 0,4s, pas de 1,375 mm/rotation, tension du tube 120 kVp, intensité de 45mAs, filtre pulmonaire.

## Génération de pseudo-étiquettes et segmentation supervisée de lésions pulmonaires

Les outils de segmentation par apprentissage profond peuvent segmenter de manière fiable des images complexes lorsqu'ils sont entraînés à l'aide d'images bien étiquetées couvrant une variété de cas, une méthode appelée entraînement supervisé. Cependant, le contourage manuel des régions de GGO et de consolidation dans un scanner avec une précision au niveau du pixel est une tâche énorme et un tel ensemble de données qui s'étend sur des cas d'imagerie variés n'est pas disponible. En outre, la délimitation des lésions sans limites claires, telles que les lésions en verre dépoli, peut être très peu reproductible. Pour contourner ce manque de données de base, nous avons d'abord généré des pseudo-contourage de manière non supervisée pour un ensemble limité de données et appliqué des corrections avant d'entraîner un réseau de neurones convolutifs (CNN) final pour la segmentation sémantique.

Tout d'abord, nous avons utilisé un modèle robuste de segmentation de la cavité pulmonaire (18), appelé masque pulmonaire, pour produire un masque de volume pulmonaire pour notre ensemble de données. En utilisant un ensemble de données de 50 patients, nous avons sélectionné les voxels dans la cavité pulmonaire qui ont des valeurs d'unités Hounsfield (HU) supérieures à -700. Ces voxels sont ensuite regroupés en 3 classes à l'aide d'une classification K-Means (par lots) employée sur des caractéristiques de voxel créées à la main, comme le montre la figure 1A. Les caractéristiques ont été obtenues à chaque voxel en utilisant : l'intensité de l'image (en HU), le hessien de la pile d'images dans trois directions, et le filtre de vaisseaux de Frangi à trois échelles différentes. Chaque voxel est ensuite converti en un vecteur à 8 dimensions.



**Figure 1 :** Schéma de la méthode de segmentation sémantique. A] Les images non étiquetées ont d'abord été décomposées en caractéristiques, puis, en utilisant le masque pulmonaire (18) et un seuil, les pixels sélectionnés ont été regroupés à l'aide de K-means. B] Unet-A a été entraîné sur le jeu de données pseudo-étiqueté obtenu par le clustering K-means. C] Unet-B a été entraîné sur les images pseudo-étiquetées prédites par Unet-A, et l'analyse volumétrique fournit les volumes de la consolidation, du verre dépoli et de la cavité thoracique.

Ensuite, les k-means ont été utilisés pour regrouper ces voxels en 3 classes qui représentent les vaisseaux, la consolidation et le verre dépoli comme le montre la figure 1[A]. Cette segmentation non supervisée reste imparfaite, par exemple, certains pixels qui appartiennent à la région des vaisseaux sont regroupés dans la consolidation et vice-versa en raison de leur similarité à la fois dans la texture et les intensités. Grâce à un nettoyage morphologique, les petits volumes de consolidation ont été érodés et des pseudo-étiquettes ont été obtenues pour 9000 coupes sélectionnées au hasard. Environ 50 de ces images générées ont été vérifiées au hasard et les erreurs de regroupement entre vaisseaux et consolidation ont été corrigées à la main, en particulier pour les images à faible dose et à fort bruit.

Comme nous l'avons déjà mentionné et comme le montre Hofmanninger (18), pour qu'une méthode supervisée puisse être généralisée à des tomodensitogrammes inconnus, un ensemble de données contenant des images variées est nécessaire. Dans cette optique, nous entraînons un CNN avec l'architecture UNET (19) sur le jeu de données non supervisé et pseudo-étiqueté, avec une augmentation importante de l'image. Les coupes de scanners ont d'abord été normalisées entre [-1024,300] et l'augmentation aléatoire de l'image comprenait un bruit gaussien additif, une mise à l'échelle, une rotation et un flou gaussien. Cette UNET (appelée Unet-A, dans la figure 1[A]) a permis de généraliser la segmentation sémantique à d'autres images. Ensuite, un jeu de données contenant des volumes de tomodensitométrie de 200 patients a été segmenté à l'aide de Unet-A et un autre jeu de données de 20 000 images et de pseudo-étiquettes a été généré. Le jeu de données a été sélectionné pour être un mélange d'images saines et d'images contenant des zones de consolidation et de GGO.

Enfin, une deuxième UNET indépendante (Unet-B sur la figure 1[B]) a été entraînée sur le jeu de données sélectionné, avec la même augmentation importante des données que celle mentionnée ci-dessus.

## **Validation du modèle à l'aide de comparaisons de volumes avec 4 évaluations manuelles et leur impact sur les données cliniques**

Pour tester par la suite la fiabilité de COViaX dans la détection des lésions et son impact sur le diagnostic, nous avons mené une étude sur des lecteurs où quatre radiologues certifiés (entre 7 et 20 ans d'expérience) conformes aux exigences de la société française de radiologie, ont interprété des scanners thoraciques, en aveugle des données cliniques. Les tomodensitométries ont été réalisées sur trois appareils différents à partir d'une cohorte de 60 patients, soit pour la caractérisation d'une pneumopathie à COVID-19, soit pour un dépistage avant chirurgie dans l'attente des résultats de la PCR. Dans cette cohorte, 18 patients étaient négatifs pour le SARS-CoV-2 détecté par PCR ; 32 patients ont été hospitalisés dont 23 étaient oxygénodépendants, tous les patients dépendants de l'oxygène étaient positifs au SARS CoV 2 en test PCR. Les lecteurs ont évalué chaque cas pour le diagnostic du SRAS-CoV-2, le pourcentage du poumon infecté (VRall), du verre dépoli (VRggo) et des lésions de consolidation (VRcs).

Deux semaines après la première session d'interprétation, les radiologues, en aveugle de leur première évaluation, ont repris l'interprétation des 60 scans, à l'aide du masque donné par AI. A nouveau, ils devaient prédire le diagnostic du SRAS-CoV-2 et donner la quantification de l'atteinte pulmonaire de manière globale et selon les deux motifs fréquents : les consolidations et le verre dépoli. Le critère de base était un score de qualité qualitatif sur une échelle de Likert à 4 points sur les masques de segmentation de l'IA comme suit : 1, fortement en désaccord (>10 % de surface de masque inappropriée) ; 2, en désaccord (entre 5 et 10 %), 3 en accord (entre 1 et 5 %) ; 4, fortement en accord (< 1 % d'erreur de segmentation). En utilisant une échelle similaire, on a demandé à chaque radiologue son avis sur la pertinence des volumes globaux calculés comme suit : 1 aucune pertinence, 2 modérée, 3 élevée, 4 excellente.

## **Analyse statistique**

Des coefficients de corrélation intraclasse (ICC) à mesure unique ont été utilisés pour évaluer la reproductibilité inter-juges des mesures de RV. Les ICC ont été calculés à l'aide de l'analyse de variance (ANOVA) aléatoire à deux voies sur les mesures moyennes (les ICC vont de 0,00 à 1,00, les valeurs plus proches de 1,00 représentant une meilleure reproductibilité) (20). L'interprétation de l'ICC a été catégorisée selon Landis et Koch (21) comme suit : <0, aucune reproductibilité ; 0,0-0,20, reproductibilité légère ; 0,21-0,40, reproductibilité moyenne ; 0,41-0,60, reproductibilité modérée ; 0,61-0,80, reproductibilité substantielle ; et 0,81-1,00, reproductibilité presque parfaite.

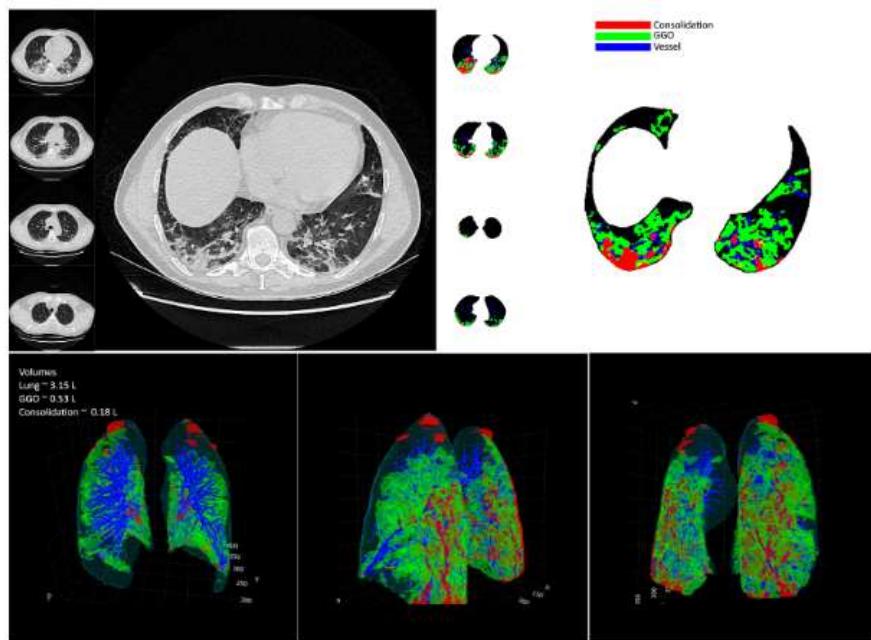
La concordance entre l'IA et les observateurs a été déterminée selon les méthodes de Bland et Altman (22). La différence critique a été définie par la moitié de la différence entre la limite inférieure et la limite supérieure.

La comparaison avec les données cliniques a utilisé les aires sous la courbe (AUC) déterminées sur les courbes caractéristiques de fonction d'efficacité du récepteur (ROC) lissées par ajustement binomial (Jokiel-Rokita et Pulit 2013 (23)). Le test de DeLong (24) a été utilisé pour comparer les AUC des courbes ROC appariées avec le package pROC pour R. Pour les variables quantitatives, la méthode de Youden a été utilisée pour déterminer les seuils des courbes ROC.

# RÉSULTATS

## Segmentation sémantique des lésions pulmonaires par apprentissage profond (COViaX)

À partir d'une coupe de scanner thoracique, COViaX classe chaque pixel en 4 classes, à savoir consolidation, verre dépoli, vaisseaux et cavité pulmonaire. Le volume des lésions dans l'image segmentée reconstruite en 3D est calculé et les centroïdes des plus gros blobs sont également déterminés. L'ensemble de l'opération prend environ 1 minute pour un scanner 3D contenant 500 coupes sur une machine avec processeur graphique et processeur central standard. La partie segmentation, qui est calculée sur un processus graphique, prend 20 secondes, tandis que les calculs de volume et de centroïde prennent plus de temps car ils sont traités par un CPU. Les tranches segmentées sont également produites en Dicom, un format qui peut être facilement superposé aux images originales à l'aide d'un logiciel standard. La méthode se généralise bien aux images non vues par le réseau pendant l'apprentissage. La figure 2 présente un exemple de segmentation et de reconstruction 3D.



**Figure 2 :** Segmentation de la coupe unique (A) et sa reconstruction 3D correspondante (B). Avec un espacement de (0,763672, 0,763672, 0,625) mm, le volume de la cavité pulmonaire, le volume du GGO et le volume de consolidation ont été estimés à 3,15 L, 0,53 L et 0,18 L, respectivement.

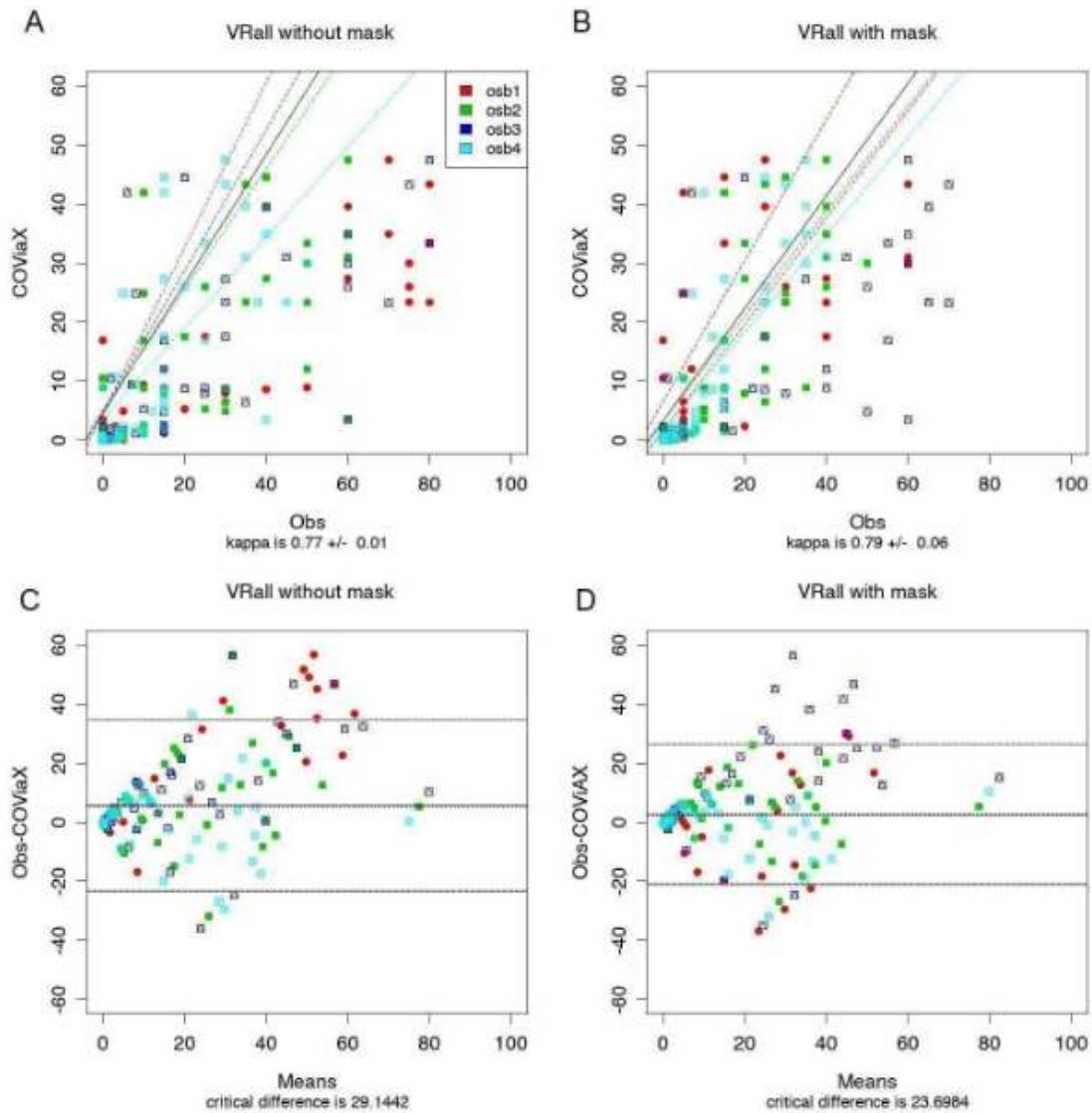
## **La qualité des masques COViaX et la confiance du radiologue dans les volumes calculés**

Tous les masques COViaX (n=28894) ont été analysés manuellement par chaque expert (n=4) pour les lésions segmentées (consolidations et opacités en verre dépoli) et leur qualité globale a été évaluée (comme décrit dans matériels et méthodes) comme "fortement en désaccord", "en désaccord", "d'accord", "fortement d'accord" dans 9%, 22%, 19%, et 50% respectivement. Les observateurs étaient d'accord de manière substantielle ( $k=0,7$ ) pour cette évaluation de la qualité. De même, la confiance dans les volumes de lésions reconstruits a été évaluée comme "aucune confiance", "modérée", "élevée" et "excellente" dans 12%, 15%, 18% et 55% respectivement. Aucune différence statistique n'a été observée entre les 3 appareils de scanner ni entre les observateurs (tableau 2). Au total, une grande majorité a trouvé une grande précision et avait confiance dans les volumes calculés par COViaX.

## **L'influence des masques COViaX sur les évaluations manuelles des volumes**

Dans un deuxième temps pour valider l'intérêt de l'IA, nous avons comparé les volumes obtenus manuellement sans et avec l'aide des masques de l'IA (masques COViaX) à ceux obtenus de manière entièrement informatisée (volumes COViaX) par intégration sur les surfaces des masques sans aucune intervention humaine. La reproductibilité était substantielle entre les prédictions COViaX et les radiologues (Obs,  $ICC=0.79 +/- 0.01$ ). Ces accords ont été encore renforcés avec des indices de confiance montrant une reproductibilité presque parfaite lorsque les masques COViaX ont été comparés aux radiologues faisant des interprétations en utilisant les masques COViaX (Obs-Msk,  $ICC=0.83 +/- 0.06$ ).

Sur les graphiques de Bland-Altman, l'incertitude du volume diminue de 28,3 % à 21,8 %, tout comme la surestimation du volume des lésions (biais) faite par les radiologues (de 5,5 % à 2,8 %). Cela indique que l'estimation du volume assistée par COViaX augmente la concordance entre les radiologues, ce qui entraîne une réduction de la subjectivité.



**Figure 3 :** Valeur ajoutée de l'évaluation manuelle du volume à l'aide des masques COViaX

La valeur ajoutée des masques AI aux évaluations subjectives du rapport de volume est démontrée par une augmentation de la corrélation (A : sans masques, B avec masques) entre les observateurs de 0,79 à 0,81. La différence critique a été définie sur les graphiques de Bland-Altman (C : sans masques, D : avec masques) par la moitié de la différence entre la limite inférieure (biais - 2SD) et la limite supérieure (biais + 2SD). L'utilisation des masques COViaX a diminué sa valeur de 28,3 % à 21,8 % ainsi que le biais de 5,5 % à 2,8 %.

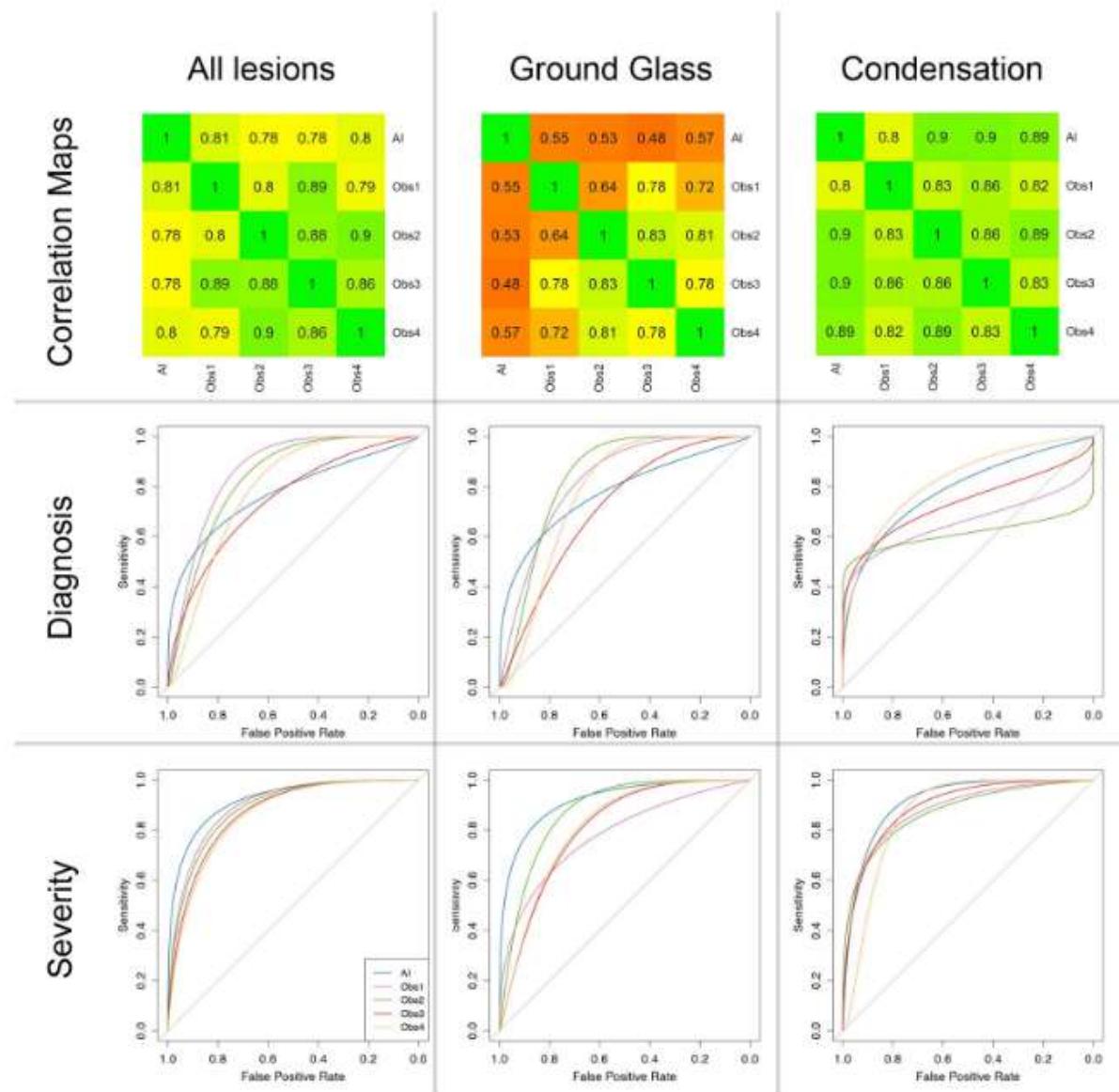
## **Conformité des volumes COViaX avec ceux obtenus manuellement et son intérêt en tant que biomarqueurs de diagnostic et de sévérité**

Des analyses de concordance inter-observateur ont également été réalisées pour tester les performances de COViaX dans la détection du volume des lésions. Les résultats de ces trois analyses sont compilés dans la figure 4.

Dans le cas du volume lésionnel total (VRall), les accords inter-observateurs sur le volume (ICC=0,85 +/- 0,04) étaient presque parfaits. Les accords étaient substantiels entre les prédictions de volume COViaX et les radiologues (Obs, ICC=0,79 +/- 0,01). Des résultats similaires ont été trouvés pour les lésions de condensation où les corrélations étaient presque parfaites entre les observateurs et entre les observateurs et les volumes COViaX (ICC=0,85 +/- 0,03 et 0,87 +/- 0,05 respectivement).

Les lésions en verre dépoli sont le plus souvent diffuses et plus difficiles à définir. En effet, dans ce cas (Fig 4), bien que les corrélations entre observateurs soient importantes (ICC=0,76 +/- 0,07), elles sont modérées entre volumes COViaX et observateurs (ICC=0,53 +/- 0,04).

Au total, l'analyse de la concordance inter-observateur montre que les volumes COViaX étaient fortement corrélés à ceux des observateurs. Les résultats ne semblent être que modérément corrélés pour les opacités en verre dépoli. Ces résultats ont été mis en perspective avec l'intérêt de ces biomarqueurs pour le diagnostic et la sévérité de la pneumopathie.



**Figure 4.** Comparaison des volumes de l'IA et de ceux de 4 observateurs : reproductibilité et leur comparaison avec les données cliniques

L'évaluation manuelle indépendante du rapport des lésions par les observateurs ( $n=4$ ) a été comparée à celle de l'IA calculée de manière entièrement automatique (COViaX) en termes de reproductibilité des 3 volumes (toutes les atteintes, atteinte verre dépoli, atteinte de type condensation = consolidation) sur les cartes de corrélation (ICC) et de leur intérêt (courbes ROC) comme biomarqueurs pour le diagnostic (RT-PCR) et la gravité (dépendance à l'oxygène). Pour le verre dépoli (VRrgo), la reproductibilité des IA et des observateurs n'était que modérée. Il est

intéressant de noter que COViaX a démontré une tendance à la supériorité pour l'évaluation de la gravité, indiquant sa robustesse par rapport à celles des observateurs. Pour les atteintes totales et la condensation, une reproductibilité substantielle à presque parfaite a été observée, avec des résultats similaires sur les performances cliniques.

Pour valider l'intérêt clinique de COViaX, nous avons comparé les volumes (manuel et COViaX) avec la PCR (critère de diagnostic) et la dépendance à l'oxygène (critère de gravité).

Les observateurs ont diagnostiqué le SRAS-CoV-2 sur un scanner thoracique à faible dose avec une sensibilité (se) de 74% (+/- 7%) et une spécificité (sp) et 74% (+/- 3%) respectivement. Ces performances étaient comparables à celles des volumes obtenus manuellement. L'AUC des volumes COViaX était de 0,78 (2ème place), 0,78 (1ère place), 0,77 (1ère place) pour VRall, VRggo, VRcs respectivement, comparable à celle des volumes manuels (AUC de 0,66 à 0,80). COViaX a surpassé obs3 pour VRall (0,78 vs 0,7, p=0,003) et VRggo (0,78 vs 0,66, p=0,001).

Les volumes pulmonaires lésés étaient tous des marqueurs de sévérité. Les AUC de COViaX étaient de 0,92 (1ère place), 0,92 (1ère place), 0,91 (1ère place) pour VRall, VRggo, VRcs respectivement. Les volumes COViaX sont apparus comme des biomarqueurs robustes avec une tendance à la supériorité par rapport à ceux des observateurs (AUC de 0,79 à 0,89). COViaX a surpassé obs1 et obs3 pour VRggo (p<0,05). Ces divergences apparentes pour le verre dépoli entre les accords modérés de COViaX et ses hautes performances cliniques sont discutées plus en détail.

## DISCUSSION

Dans le monde entier, l'épidémie de SRAS-Cov2 a augmenté la charge de travail des professionnels de la santé jusqu'au point de rupture. Dans le cas des spécialistes de l'imagerie médicale, le nombre d'images à analyser a augmenté soudainement et massivement (25). Dans ce contexte, l'un des principaux objectifs d'un outil d'analyse d'images par l'IA tel que COViaX est donc d'aider les radiologues à évaluer les lésions pulmonaires chez les patients infectés par le SRAS-CoV-2 et de soulager leur charge de travail. En raison du besoin urgent d'un tel outil, nous avons développé un logiciel U-NET de segmentation pulmonaire sémantique faiblement supervisé qui apprend à partir d'images partiellement annotées en se basant sur des annotations manuelles assistées par un clustering K-means. Nous montrons ici que cette approche contourne avec succès le besoin d'un ensemble de données entièrement annotées pour la prédiction de masques fiables à utiliser par les radiologues.

### **Masque COViaX pour la segmentation des lésions pulmonaires.**

Quatre observateurs experts indépendants ont évalué la précision du masque COViaX pour la segmentation des lésions pulmonaires. Dans l'ensemble, ils ont déterminé que la segmentation COViaX était très bonne à parfaite dans 69% des masques et que la confiance était élevée à excellente dans 73% des masques, indépendamment de l'appareil utilisé (3 scanners différents). L'analyse de concordance a révélé que les volumes de condensation COViaX étaient fortement corrélés à ceux des observateurs et avec une grande précision. La concordance des opacités de verre dépoli était moins solide entre les observateurs mais surtout entre COViaX et les observateurs. Une analyse plus approfondie a révélé que malgré ces résultats décevants, les volumes de verre dépoli prédits par COViaX étaient les plus précis pour prédire la maladie du SRAS-CoV2 et sa gravité, dépassant les performances de tous les observateurs. Ces résultats démontrent que l'approximation manuelle du volume de verre dépoli est difficile à évaluer avec précision dans la pratique clinique. Ceci semble être préjudiciable car la plupart des lésions de verre dépoli sont liées à une destruction aiguë des cellules épithéliales pulmonaires dans l'espace alvéolaire (26) dont l'étendue semble particulièrement importante dans la pratique de l'urgences. Par conséquent, nous pensons que la mesure COViaX est de meilleure qualité par rapport à l'approximation manuelle.

### **COViaX comme outil de triage des patients COVID-19.**

Il est important de connaître la fraction volumique du poumon infecté (VRall) car elle peut servir de prédicteur de décès précoce ou d'admission en soins intensifs. Pour Ruch (27), l'étendue des lésions sur le scanner initial était indépendamment associée au pronostic (odds ratio=2,35). Dans leur expérience, alors qu'aucun patient avec un scanner thoracique normal n'a développé une maladie sévère, c'était le cas pour 70% des patients avec un VRall >50%. Dans la présente étude, les VRall déterminés étaient clairement corrélés à l'oxygено-dépendance des patients, et étaient donc un facteur prédictif de la gravité de la maladie. Il est important de noter que les performances du COViaX étaient systématiquement supérieures à celles des experts avec une AUC de 92%, 92%, 91% pour VRall, VRgg, VRcs respectivement. Ainsi, COViaX a une application potentielle pour la prédiction de la progression de la maladie et de son issue chez les patients.

Il pourrait être intéressant d'inclure les mesures de COViaX dans une approche multiparamétrique pour des prédictions plus fines et la détermination des comorbidités. Par exemple, Shuo Wang and al (28) a testé l'application de l'IA pour la prédiction du risque sur la base d'une cohorte de validation de 471 patients présentant des caractéristiques d'imagerie déterminées ainsi que des caractéristiques cliniques (âge, sexe et comorbidité). Deux groupes de risque statistiquement distincts ont ainsi été identifiés avec des différences dans la durée d'hospitalisation ( $p<0,0001$ ). De légères corrélations entre le VRall et les biomarqueurs inflammatoires (leucocytes, neutrophiles et IL-2R) pourraient également être indicatives (29). Mei et al (30) dans une récente étude de dépistage ont établi que la combinaison de l'analyse automatisée du scanner avec l'histoire clinique et biologique d'un patient pourrait aider à déterminer un score de sévérité COVID-19 comme pronostic pour trier les patients à risque. Une telle approche, développée de manière automatisée, pourrait soulager significativement la prise en charge des patients, un enjeu majeur dans un contexte de ressources sanitaires dégradées. De manière générale, les outils d'IA sont parfaitement adaptés à cet objectif, en raison de leur faible coût et de leur potentiel d'application à haut débit.

### **COViaX comme outil de diagnostic immédiat du COVID-19.**

Dans cette étude, nous avons observé que le scanner thoracique était modérément performant en tant qu'outil de diagnostic des patients affectés par le COVID-19, la moyenne des Se et Sp du diagnostic du SRAS-CoV2 pour le scanner thoracique était de 74% et 74% respectivement. L'utilisation des masques COViaX n'a pas augmenté ces performances (71% et 78% respectivement). Ces performances sont inférieures à celles décrites précédemment. Dans une grande enquête nationale menée dans 26 hôpitaux (n= 4824 sujets), la sensibilité et la spécificité du CT thoracique pour le diagnostic de la pneumonie COVID-19 étaient respectivement de 90% et 91% (31). D'autres rapports ont montré des résultats dans le même taux de performance : 87 % et 97 % pour (15) (AUC=0,974) ; 97 % et 92 % pour Bo Wang and al (32). Ici, la sensibilité inférieure n'était pas due au logiciel d'IA car il était tout aussi efficace que les 4 observateurs. Il est probable que les taux de performance inférieurs s'expliquent par l'inclusion de patients asymptomatiques aux premiers stades de l'infection, empêchant la détection de lésions pulmonaires sur les scanners. Des taux de sensibilité significativement plus élevés seraient probablement trouvés dans une population de patients symptomatiques. En raison des impacts variables de la maladie sur les patients, il est à nouveau probable que les approches multiparamétriques renforcent également les performances diagnostiques de l'IA ou des radiologues seuls. Par exemple, Mei et al (30) ont observé une augmentation de la spécificité de 75,9 % à 82,8 % lorsque l'IA était associée à d'autres paramètres.

### **L'utilisation de COViaX à l'hôpital.**

Cette étude démontre que la quantification des lésions dans la pneumonie SARS-CoV-2 peut être réalisée avec un cadre d'IA faiblement supervisé permettant une segmentation robuste et qualitative des lésions pulmonaires. COViaX n'est pas sensible à des constructeurs de scanners spécifiques car les examens rapportés ont eu lieu sur trois scanners différents sans différence notable dans l'appréciation qualitative des radiologues. Ainsi, COViaX peut être utilisé en téléradiologie ou dans des essais multicentriques.

Ces résultats montrent que COViaX peut être immédiatement mis en œuvre dans un contexte clinique pour soutenir les analyses de scanners thoraciques avec une grande précision, en particulier dans un contexte épidémique.

À plus long terme, COViaX (ainsi que d'autres approches basées sur l'IA) pourrait être utilisé comme un outil à haut débit pour l'évaluation de l'efficacité de traitements potentiels surveillant les dommages pulmonaires au fil du temps dans de grandes populations de patients. Les approches d'IA pourraient également conduire à la surveillance automatisée de toute l'imagerie thoracique (par exemple via un matériel dédié, un processeur GPU et une extraction automatique) afin de détecter et d'alerter sur les cas positifs quelques secondes après l'acquisition. Un tel workflow pourrait contribuer à réduire la transmission nosocomiale de la maladie, en permettant une décontamination rapide du scanner, de sa table et de sa salle, et une prise en charge adaptée du patient avec analyse et isolement par PCR.

## **CONCLUSION**

L'outil de contourage d'intelligence artificielle et les volumes calculés de notre logiciel COViaX sont fiables et donnent un biomarqueur robuste de la sévérité de la maladie.

## REFERENCES

1. Practice BB. Coronavirus disease 2019. World Heal Organ [Internet]. 2020;2019:2633. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N Engl J Med* [Internet]. 2020 Mar 26 [cited 2020 May 22];382(13):1199–207. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa2001316>
3. Priesemann V, Balling R, Brinkmann MM, Ciesek S, Czypionka T, Eckerle I, et al. An action plan for pan-European defence against new SARS-CoV-2 variants [Internet]. Vol. 397, *The Lancet*. Elsevier B.V.; 2021 [cited 2021 Mar 22]. p. 469–70. Available from: <https://cmmid.github.io/topics/covid19/local->
4. Huang P, Liu T, Huang L, Liu H, Lei M, Xu W, et al. Use of Chest CT in Combination with Negative RT-PCR Assay for the 2019 Novel Coronavirus but High Clinical Suspicion. *Radiology*. 2020 Feb 12;200330.
5. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, et al. CT imaging features of 2019 novel coronavirus (2019-NCoV). *Radiology* [Internet]. 2020 Feb 4 [cited 2020 Nov 27];295(1):202–7. Available from: <https://doi.org/10.1148/radiol.2020200230>
6. M N, EY L, J Y, F Y, X L, H W, et al. Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review. *Radiol Cardiothorac Imaging* [Internet]. 2020 Jun 5 [cited 2020 Nov 27];2(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7233595/?tool=EBI>
7. Shi H, Han X, Jiang N, Cao Y, Alwaid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis* [Internet]. 2020 Apr 1 [cited 2020 Nov 27];20(4):425–34. Available from: [www.thelancet.com/infection](http://www.thelancet.com/infection)
8. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19)

Pneumonia. Radiology. 2020 Feb 13;200370.

9. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, et al. Chest CT findings in coronavirus disease 2019 (COVID-19): Relationship to duration of infection [Internet]. Vol. 295, Radiology. Radiological Society of North America Inc.; 2020 [cited 2020 Nov 27]. p. 685–91. Available from: <https://doi.org/10.1148/radiol.2020200463>
10. Shi H, Han X, Zheng C. Evolution of CT Manifestations in a Patient Recovered from 2019 Novel Coronavirus (2019-nCoV) Pneumonia in Wuhan, China. Radiology [Internet]. 2020 Apr 7 [cited 2020 Nov 27];295(1):20–20. Available from: <http://pubs.rsna.org/doi/10.1148/radiol.2020200269>
11. Chua F, Armstrong-James D, Desai SR, Barnett J, Kouranos V, Kon OM, et al. The role of CT in case ascertainment and management of COVID-19 pneumonia in the UK: insights from high-incidence regions [Internet]. Vol. 8, The Lancet Respiratory Medicine. Lancet Publishing Group; 2020 [cited 2020 Nov 27]. p. 438–40. Available from: <https://pubmed.ncbi.nlm.nih.gov/32220663/>
12. Touati J, Rico G, Habert P, Géralt M, Gaubert JY, Chagnaud C, et al. Prevention of nosocomial transmission of SARS-CoV-2 using pre-operative chest CT: a monocentric study during the outbreak. 2020 [cited 2020 Aug 28]; Available from: <https://doi.org/10.1016/j.jhin.2020.07.005>
13. Med LR. Comment The role of CT in case ascertainment and management of COVID-19 pneumonia in the UK : insights from high-incidence regions. Lancet Respir Med. 2020;2600(20):19–20.
14. Gupta S, Hayek SS, Wang W, Chan L, Mathews KS, Melamed ML, et al. Factors Associated with Death in Critically Ill Patients with Coronavirus Disease 2019 in the US. JAMA Intern Med [Internet]. 2020 [cited 2021 Mar 22];180(11). Available from: <https://pubmed.ncbi.nlm.nih.gov/32667668/>
15. Jin C, Chen W, Cao Y, Xu Z, Tan Z, Zhang X, et al. Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nat Commun [Internet]. 2020 Dec 1 [cited 2020 Nov 27];11(1). Available from: [/pmc/articles/PMC7547659/?report=abstract](https://pmc/articles/PMC7547659/?report=abstract)

16. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, et al. Lung Infection Quantification of COVID-19 in CT Images with Deep Learning Author list.
17. Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* [Internet]. 2020 Aug 1 [cited 2020 Nov 27];56(2). Available from: [/pmc/articles/PMC7243395/?report=abstract](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7243395/?report=abstract)
18. Hofmanninger J, Prayer F, Pan J, Rohrich S, Prosch H, Langs G. Automatic lung segmentation in routine imaging is a data diversity problem, not a methodology problem. 2020 Jan 31 [cited 2020 Apr 14]; Available from: [http://arxiv.org/abs/2001.11767](https://arxiv.org/abs/2001.11767)
19. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lecture Notes in Computer Science* (including subseries *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Bioinformatics*) [Internet]. Springer Verlag; 2015 [cited 2021 Mar 22]. p. 234–41. Available from: <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net>.
20. Donner A, Koval JJ. The Estimation of Intraclass Correlation in the Analysis of Family Data. *Biometrics*. 1980 Mar;36(1):19.
21. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977 Mar;33(1):159.
22. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* [Internet]. 1999 Apr 2 [cited 2021 Mar 22];8(2):135–60. Available from: [http://journals.sagepub.com/doi/10.1177/096228029900800204](https://journals.sagepub.com/doi/10.1177/096228029900800204)
23. Jokiel-Rokita A, Pilit M. Nonparametric estimation of the ROC curve based on smoothed empirical distribution functions. *Stat Comput* [Internet]. 2013 Nov 28 [cited 2021 Mar 22];23(6):703–12. Available from: <https://link.springer.com/article/10.1007/s11222-012-9340-x>
24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988 Sep;44(3):837.

25. Herpe G, Naudin M, Léderlin M, Enikeeva F, Boumendil O, Cassagnes L, et al. COVID-19 impact assessment on the French radiological centers: a nationwide survey. *Eur Radiol* [Internet]. 2020 Jul 3 [cited 2020 Sep 25];1–8. Available from: <https://doi.org/10.1007/s00330-020-07035-w>
26. Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: A review. Vol. 215, *Clinical Immunology*. Academic Press Inc.; 2020. p. 108427.
27. Ruch Y, Kaeuffer C, Ohana M, Labani A, Fabacher T, Bilbault P, et al. CT lung lesions as predictors of early death or ICU admission in COVID-19 patients. *Clin Microbiol Infect* [Internet]. 2020 Oct 1 [cited 2020 Nov 27];26(10):1417.e5–1417.e8. Available from: <https://pubmed.ncbi.nlm.nih.gov/32717417/>
28. Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, et al. A Fully Automatic Deep Learning System for COVID-19 Diagnostic and Prognostic Analysis. *medRxiv* [Internet]. 2020 Mar 26 [cited 2020 Mar 31];2020.03.24.20042317. Available from:  
<http://medrxiv.org/content/early/2020/03/26/2020.03.24.20042317.abstract>
29. Hu Y, Zhan C, Chen C, Ai T, Xia L. Chest CT findings related to mortality of patients with COVID-19: A retrospective case-series study. *PLoS One* [Internet]. 2020 Aug 1 [cited 2020 Nov 27];15(8 August). Available from: [/pmc/articles/PMC7447035/?report=abstract](https://pmc/articles/PMC7447035/?report=abstract)
30. Mei X, Lee HC, Diao K yue, Huang M, Lin B, Liu C, et al. Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. *Nat Med* [Internet]. 2020 Aug 1 [cited 2020 Aug 24];26(8):1224–8. Available from:  
<https://doi.org/10.1038/s41591-020-0931-3>
31. Herpe G, Lederlin M, Naudin M, Ohana M, Chaumoitre K, Gregory J, et al. Efficacy of Chest CT for COVID-19 Pneumonia in France. *Radiology* [Internet]. 2020 Sep 1 [cited 2020 Nov 27];202568. Available from:  
<http://pubs.rsna.org/doi/10.1148/radiol.2020202568>
32. Wang B, Jin S, Yan Q, Xu H, Luo C, Wei L, et al. AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system. *Appl Soft Comput* [Internet]. 2020 Nov 10 [cited 2020 Nov 27];106897.

# TABLES

**Table 1: Caractéristiques des patients**

	All	PCR-	PCR+	p
<b>n</b>	60 (100%)	18 (100%)	42 (100%)	0.023
<b>Age</b>	52+/-17	49+/-13	53+/-19	0.282
<b>Sexe</b>				
F	29 (48%)	11 (61%)	18 (43%)	0.195
M	31 (52%)	7 (39%)	24 (57%)	0.195
Sexe Ratio	0.94	1.57	0.75	0.195
<b>Antécédents</b>				
Tabac	7 (12%)	1 (6%)	6 (14%)	0.334
Obésité	9 (15%)	1 (6%)	8 (19%)	0.18
Hypertension	11 (18%)	1 (6%)	10 (24%)	0.094
Asthme	1 (2%)	1 (6%)	0 (0%)	0.123
Diabète	3 (5%)	0 (0%)	3 (7%)	0.245
Immunodéficience	6 (10%)	2 (11%)	4 (10%)	0.851
Cancer	9 (15%)	2 (11%)	7 (17%)	0.581
<b>Symptômes</b>	50 (83%)	12 (67%)	38 (90%)	0.023
Fièvre	11 (18%)	1 (6%)	10 (24%)	0.094
Diarrhée	10 (17%)	4 (22%)	6 (14%)	0.45
Myalgie	22 (37%)	5 (28%)	17 (40%)	0.35
Anosmie	16 (27%)	3 (17%)	13 (31%)	0.252
Céphalée	17 (28%)	5 (28%)	12 (29%)	0.95
Toux	28 (47%)	5 (28%)	23 (55%)	0.055
Dyspnée	18 (30%)	3 (17%)	15 (36%)	0.14
<b>Soins</b>				
O2	22 (37%)	0 (0%)	22 (52%)	<0.001
Hospitalisation	32 (53%)	3 (17%)	29 (69%)	<0.001
Réanimation	10 (17%)	0 (0%)	10 (24%)	0.023

**Table 2: Évaluation qualitative de l'IA selon différentes machines**

Score	n	ICC	All.obs	Obs1	Obs2	Obs3	Obs4
<b>Qualité des masques</b>							
Tous	60	0.7	3.1 ± 1	2.8 1.2	± 3.5 0.8	± 3.3 0.9	± 2.8 ± 1
Scanner 1	20	0.61	3.2 ± 1	2.7 1.3	± 3.8 0.5	± 3.3 0.9	± 2.9 ± 1
Scanner 2	20	0.3	3.3 0.9	± 3 ± 1.1	3.5 0.7	± 3.5 0.8	± 3.2 ± 0.9
Scanner 3	20	0.84	2.8 1.1	± 2.8 1.4	± 3 ± 1	3.2 1.1	± 2.2 ± 0.8
<b>Confiance au volume de l'IA</b>							
Tous	60	0.7	3.2 1.1	± 3 ± 1.2	3.4 0.8	± 3.2 ± 1	3 ± 1.2
Scanner 1	20	0.59	3.3 1.1	± 3.3	± 3.7	± 3.6	± 2.9 ± 1.3
Scanner 2	20	0.58	3.2 ± 1	2.9 1.2	± 3.5 0.6	± 3.2 0.9	± 3.2 ± 1.1
Scanner 3	20	0.82	2.9 1.1	± 2.9 1.2	± 3 ± 0.9	3 ± 1.3	2.8 1.1

**Table 3: Performances cliniques de COViaX et comparaisons inter-observateurs pour le diagnostic et la sévérité**

Eval	Obs	Mask	AUC	p	Seuil	Se	Sp	Vpp	Vpn
<b>Diagnostic</b>									
Expertise	obs1					0.83	0.71	0.56	0.91
Expertise	obs2					0.72	0.73	0.54	0.86
Expertise	obs3					0.72	0.73	0.54	0.86
Expertise	obs4					0.67	0.78	0.57	0.84
Volume	ai	all	0.78		1.5	0.71	0.78	0.88	0.54
Volume	obs1	all	0.8	0.64	1	0.71	0.83	0.91	0.56
Volume	obs2	all	0.75	0.6	1.5	0.73	0.72	0.86	0.54
Volume	obs3	all	0.7	0.03	4	0.66	0.78	0.87	0.5
Volume	obs4	all	0.74	0.41	4.5	0.68	0.78	0.88	0.52
Volume	ai	gg	0.78		0.8	0.73	0.78	0.88	0.56
Volume	obs1	gg	0.74	0.47	1	0.61	0.83	0.89	0.48
Volume	obs2	gg	0.74	0.38	1	0.73	0.72	0.86	0.54
Volume	obs3	gg	0.66	0.01	1.4	0.63	0.72	0.84	0.46
Volume	obs4	gg	0.72	0.19	0.5	0.76	0.67	0.84	0.55
Volume	ai	cs	0.77		0.4	0.59	0.94	0.96	0.5
Volume	obs1	cs	0.73	0.38	2.2	0.49	1	1	0.46
Volume	obs2	cs	0.74	0.52	2	0.51	1	1	0.47
Volume	obs3	cs	0.74	0.45	2.3	0.54	0.94	0.96	0.47
Volume	obs4	cs	0.76	0.58	2	0.55	0.94	0.96	0.49
<b>Sévérité</b>									
Volume	ai	all	0.92		2.5	0.91	0.78	0.72	0.94
Volume	obs1	all	0.89	0.33	7.5	0.87	0.78	0.71	0.91
Volume	obs2	all	0.86	0.27	22.5	0.7	0.92	0.84	0.83
Volume	obs3	all	0.87	0.12	4	0.91	0.7	0.66	0.93
Volume	obs4	all	0.85	0.11	3.5	0.91	0.62	0.6	0.92
Volume	ai	gg	0.92		2.3	0.91	0.78	0.72	0.94
Volume	obs1	gg	0.77	0.02	4	0.74	0.73	0.63	0.82
Volume	obs2	gg	0.84	0.15	11	0.7	0.89	0.8	0.82
Volume	obs3	gg	0.79	0.01	1.4	0.87	0.68	0.62	0.89
Volume	obs4	gg	0.8	0.05	3	0.87	0.65	0.61	0.89
Volume	ai	cs	0.91		0.4	0.91	0.81	0.75	0.94
Volume	obs1	cs	0.83	0.08	2.2	0.74	0.89	0.81	0.85
Volume	obs2	cs	0.84	0.11	2	0.74	0.86	0.77	0.84
Volume	obs3	cs	0.89	0.62	1.8	0.83	0.81	0.73	0.88
Volume	obs4	cs	0.87	0.35	1.6	0.87	0.81	0.74	0.91

# FIGURES

## **Figure 1 : Schéma de la méthode de segmentation sémantique**

A] Les images non étiquetées ont d'abord été décomposées en caractéristiques, puis, en utilisant le masque pulmonaire (18) et un seuil, les pixels sélectionnés ont été regroupés à l'aide de K-means. B] Unet-A a été entraîné sur le jeu de données pseudo-étiqueté obtenu par le clustering K-means. C] Unet-B a été entraîné sur les images pseudo-étiquetées prédites par Unet-A, et l'analyse volumétrique fournit les volumes de la consolidation, du verre dépoli et de la cavité thoracique.

## **Figure 2 : Exemple de segmentation et de reconstruction 3D**

Segmentation de la coupe unique (A) et sa reconstruction 3D correspondante (B). Avec un espacement de (0,763672, 0,763672, 0,625) mm, le volume de la cavité pulmonaire, le volume du GGO et le volume de consolidation ont été estimés à 3,15 L, 0,53 L et 0,18 L, respectivement.

## **Figure 3 : Valeur ajoutée de l'évaluation manuelle du volume à l'aide des masques COViaX**

## **Figure 4 : Comparaison des volumes de l'IA et de ceux de 4 observateurs : reproductibilité et leur comparaison avec les données cliniques**

# **SERMENT HIPPOCRATE**

Au moment d'être admise à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences.

Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera.

Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admise dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçue à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité.

Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses; que je sois déshonorée et méprisée si j'y manque.