

# Développement et évaluation d'un outil d'aide à la décision clinique pour le suivi des patients hospitalisés pour COVID-19 au CHU de Liège

Nathalie MAES<sup>1</sup>, Jakez ROLLAND<sup>2</sup>, Ronan BOUTIN<sup>3</sup>, Marie CODET<sup>4</sup>, Arnaud LECLEVE<sup>5</sup>, Raphaël SCHILS<sup>6</sup>, Marie THYS<sup>7</sup>, Adelin ALBERT<sup>8</sup>, Michel MOUTSCHEN<sup>9</sup>, Philippe KOLH<sup>10</sup>

<sup>1</sup> Appui à la Recherche Clinique et Biostatistique, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, nmaes@chuliege.be

<sup>2</sup> Bio Logbook, Julien Videment 1, 44200 Nantes, France, jakez.rolland@biologbook.fr

<sup>3</sup> Bio Logbook, Julien Videment 1, 44200 Nantes, France, ronan.boutin@biologbook.fr

<sup>4</sup> Bio Logbook, Julien Videment 1, 44200 Nantes, France, marie.codet@biologbook.fr

<sup>5</sup> Bio Logbook, Julien Videment 1, 44200 Nantes, France, arnaud.lecleve@biologbook.fr

<sup>6</sup> Service de Médecine interne et Infectiologie, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, raphael.schils@chuliege.be

<sup>7</sup> Exploitation des données médico-économiques, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, mthys@chuliege.be

<sup>8</sup> Appui à la Recherche Clinique et Biostatistique, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, aalbert@uliege.be

<sup>9</sup> Service de Médecine interne et Infectiologie, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, mmoutschen@chuliege.be

<sup>10</sup> Gestion du Système d'Informations, CHU de Liège, rue de l'Hôpital, 1, 4000 Liège, Belgique, Philippe.Kolh@chuliege.be

## Résumé:

Les vagues de COVID-19 ont engendré une forte pression sur les hôpitaux, particulièrement sur les unités de soins intensifs, et une mortalité importante. Cette étude menée au CHU de Liège vise au développement et à l'évaluation d'un outil d'aide à la décision clinique (DCSS) pour le suivi des patients hospitalisés pour COVID-19. Le DCSS « Bio Logbook » est basé sur une méthode de détection des dérives précoces des paramètres de biologie. Il établit des scores de risque liés à la COVID-19 à partir de tests de laboratoire mesurés à l'admission du patient et identifie parmi ces tests ceux liés à la sévérité de la COVID-19. Des modèles prédictifs des risques liés à la COVID-19 ont été élaborés et validés pour le décès, l'admission aux soins intensifs, la ventilation mécanique et la transfusion. La plus-value clinique du CDSS a été évaluée à l'aide d'une étude rétrospective de cas. L'outil d'aide à la décision clinique développé et validé spécifiquement lors de cette étude montre des résultats encourageants, en particulier pour la détection précoce des patients particulièrement à risque de décès lors de leur hospitalisation. L'outil fournit ainsi au clinicien une aide lui permettant d'optimiser la rapidité et la qualité de la prise en charge des patients à partir de l'analyse des résultats de laboratoire à l'admission ; il permet aussi d'augmenter la capacité du médecin à identifier les patients à haut ou à plus faible risque de décès dans le séjour. L'étude de cas met en évidence l'importance de la synergie médecin-CDSS. Enfin, cette étude a permis de mettre en place une méthode complète de validation d'un CDSS, y compris de sa plus-value clinique, avant son déploiement.

**Mots clés:** Aide à la Décision Clinique, COVID-19, Résultats de laboratoire

## 1. Introduction

L'épidémie de COVID-19 a touché la Belgique en mars 2020. Les vagues se sont ensuite succédées engendrant une forte pression sur les hôpitaux, particulièrement sur les unités de soins intensifs, et une mortalité importante. L'arrivée des traitements et la vaccination ont amélioré la situation mais les variants successifs du virus ont maintenu un taux d'hospitalisation élevé et une surcharge de travail pour les soignants. Au cours du temps, il s'est avéré nécessaire d'optimiser l'organisation des soins médicaux avec notamment pour objectif la détection et la priorisation du patient à haut risque. L'extrême rapidité de propagation et la complexité de la COVID-19 exigent plus que jamais des soins personnalisés avec une compréhension globale et détaillée de l'état de santé du patient pendant son séjour à l'hôpital.

Pour faire face à ces défis, le Centre Hospitalier Universitaire de Liège (CHU de Liège) a participé au développement et à la validation de plusieurs outils d'Intelligence Artificielle (*Artificial Intelligence*, AI) ou d'Aide à la Décision Clinique (*Clinical Decision Support Systems*, CDSS). Ces outils visaient principalement le diagnostic rapide des patients atteints de la COVID-19 afin d'accélérer la mise en place des procédures d'isolement et de raccourcir les délais de prise en charge [Guiot et al, 2021], la détection du risque d'apparition d'une forme grave, ainsi que le triage à l'admission à l'hôpital [Wu et al, 2020].

Malgré l'importance croissante accordée aux outils AI et CDSS pour améliorer les soins et réduire les coûts, les données probantes à l'appui d'une utilisation généralisée sont limitées [Bright et al, 2012]. Bien qu'utilisés de plus en plus, il est important d'établir des systèmes de contrôle et de validation de ces outils et de définir des recommandations relatives aux procédures d'implémentation [Shahsavarani et al, 2015]. Pour l'hôpital, il est opportun non seulement de pouvoir s'assurer qu'un CDSS fournit des résultats pertinents et valides mais également qu'il apporte une réelle plus-value clinique.

Le logiciel « Bio Logbook » est un outil d'aide à la décision basé sur les données biologiques du patient.

Cette étude avait pour objectif de développer un logiciel spécifiquement dévolu à la prise en charge des patients COVID-19. Elle visait à établir des scores de risque liés à la COVID-19 (décès, admission aux soins intensifs, ventilation mécanique, transfusion) à partir de tests de laboratoire mesurés à l'admission du patient et d'identifier parmi ces tests ceux liés à la sévérité de la COVID-19. L'étude avait également pour objectif de mettre au point un cadre pour l'évaluation de l'impact clinique du logiciel en tant que système d'aide à la décision dans la gestion hospitalière des patients atteints de COVID-19 admis CHU de Liège.

L'étude a été approuvée par le Comité d'Ethique du CHU de Liège (n°2021/240).

## 2. Méthodes

### 2.1 Intervalles de référence et dérive des tests de laboratoire

Bio Logbook a développé une méthode permettant de détecter les dérives précoces des paramètres de biologie. Cette approche nouvelle consiste en l'analyse fine des variations chronologiques d'un paramètre biologique.

- Soit le paramètre varie autour de la moyenne statistique de référence, auquel cas cette variation correspond à celle d'un patient sain ;

- Soit, il ne varie pas autour de la moyenne statistique de la population de référence, auquel cas il est en dérive statistique, et on peut parler “d’état de santé intermédiaire” ;
- Soit, le paramètre est hors des valeurs seuils de référence, il est donc en dérive importante, auquel cas le patient est probablement pathologique (pratique médicale actuelle).

Afin d’affiner la notion de dérive, il est également nécessaire d’homogénéiser les populations de référence des paramètres biologiques. A cet effet, Bio Logbook a développé une méthode innovante de détermination des populations de référence des biomarqueurs. Ces populations de référence optimisées et personnalisées ont été redéfinies et validées à l’aide d’échantillons de données rétrospectives anonymisées du CHU de Liège. Elles sont obtenues par des méthodes algorithmiques, (1) en excluant les patients porteurs d’un paramètre infectieux influant sur le paramètre biologique étudié et ceux ayant des paramètres biologiques influents en dehors des valeurs seuils de référence actuelle ; (2) en personnalisant en fonction du sexe, poids, taille, groupe sanguin du patient.

Afin de réaliser un diagnostic complet et rapide, le logiciel Bio Logbook réalise des pronostics personnalisés (modèles statistiques multi paramètres), permettant d’alerter les utilisateurs, des causes ou des conséquences probables des déséquilibres biologiques précoces. Ainsi le programme permet de pronostiquer des centaines de biomarqueurs à partir d’un premier jeu de données de tests de laboratoire. L’enjeu est d’alerter le clinicien sur un risque anormalement élevé que le paramètre biologique soit en dérive ou en dérive importante. Ainsi l’alerte donnera la possibilité au soignant de contrôler les différents paramètres mis en évidence, la mesure complémentaire donnera la valeur vraie des différents biomarqueurs et alors seulement la dérive ou dérive importante aura été détectée de manière factuelle.

Cette méthodologie à l’avantage de détecter des déséquilibres de paramètres auxquels on ne pense pas, oubliés ou encore dont on n’imaginait pas qu’ils soient gravement perturbés.

## 2.2 Modèles prédictifs des risques liés à la COVID-19 et leur validation

Des modèles prédictifs des risques liés à la COVID-19 ont été élaborés pour le décès, l’admission aux soins intensifs, la ventilation mécanique et la transfusion à partir des données rétrospectives des patients COVID-19 âgés de plus de 18 ans hospitalisés au CHU de Liège entre le 14 juillet 2020 et le 9 avril 2021 (deuxième et troisième vagues de l’épidémie). Le premier échantillon de sang maximum 3 jours après la détection de la COVID-19 a été utilisé. Les paramètres avec plus de 20% de données manquantes puis les patients avec des valeurs manquantes ont été supprimés. Pour chaque résultat de paramètre biologique, la sélection des paramètres de laboratoire pertinents et en relation avec celui-ci a été effectuée à l’aide d’une analyse de régression logistique et des *odds ratios* correspondants. Afin de construire le modèle prédictif le plus performant en termes de sensibilité, spécificité et AURC (Area under ROC curve), nous avons testé différentes modélisations. Les modèles de type random forest [Breiman, 2001] se sont montrés plus performant que les régressions linéaires multiparamétriques, les régressions logistiques et les arbres de décisions simples. Nous avons utilisé le package partykit du logiciel R pour créer les modèles de type Random Forest.

La valeur de Shapley a été utilisée pour assigner un coefficient à chaque paramètre biologique afin d’obtenir pour chaque modèle et pour chaque patient la liste des paramètres biologiques aggravants qui augmentent le risque étudié ainsi que les paramètres de protection qui réduisent ce risque.

Un échantillon de 80% des données a été utilisé pour construire les modèles (*Training sample*) et les 20% restants ont été utilisés pour les tester (*Test sample*). Les modèles prédictifs finaux ont ensuite été testés sur un troisième échantillon rétrospectif de patients COVID-19 de plus de 18 ans hospitalisés au CHU de Liège après le 4 mai 2021 et sortis au 3 novembre 2021 (*Validation sample*), soit au cours des troisième et quatrième

vagues de l'épidémie. Les groupes de risques associés à chaque patient et générés par les modèles de Bio Logbook ont été comparés à son séjour hospitalier réel à l'aide d'une analyse d'efficacité (sensibilité/spécificité).

## 2.3 Etude de cas

La plus-value clinique du CDSS a été évaluée à l'aide d'une étude rétrospective de cas réalisée par un infectiologue du CHU de Liège. L'étude a été réalisée en deux étapes :

- Etape 1 : analyse d'un échantillon anonymisé de patients à partir des résultats des tests de laboratoire à l'admission du patient tels que mis à disposition dans le Dossier Patient Informatisé.
- Etape 2 (1 mois plus tard) : analyse du même échantillon de patients à l'aide du logiciel CDSS Bio Logbook.

Le médecin a été formé à l'utilisation du logiciel entre les deux étapes.

A chaque étape, les questions suivantes ont été posées au médecin spécialiste :

- Quels sont les paramètres biologiques pour lesquels vous observez des déviations cliniquement significatives pour le patient ?
- Quels sont selon vous les risques de décès, de passage aux soins intensifs, de ventilation mécanique, de transfusion (risque élevé, ambigu ou faible) pour le patient ?
- Quels sont les paramètres biologiques (déjà mesurés ou non) que vous souhaitez suivre pour ce patient ?

Les résultats aux deux étapes ont été comparés ainsi qu'aux résultats fournis par le CDSS seul.

Les patients COVID-19 inclus dans cette étude de cas ont été sélectionnés de manière aléatoire et avaient plus de 18 ans, étaient sortis de l'hôpital au moment de l'étude, n'avaient pas été transférés vers ou à partir d'un autre hôpital, avaient une durée de séjour d'au moins 7 jours et avaient au moins un prélèvement sanguin analysé dans les 3 premiers jours de leur hospitalisation. La proportion de patients décédés pendant le séjour hospitalier a volontairement été surreprésentée dans l'échantillon. Les patients ont été traités de manière anonyme mais le médecin disposait de leur classe d'âge, genre, Indice de Masse Corporelle (IMC), et de leurs principales comorbidités cliniques (obésité, diabète, maladie rénale chronique, hypertension, maladie cardiovasculaire, maladie pneumologique chronique, immunosuppression, patient oncologique).

Une enquête a également été soumise au médecin spécialiste afin d'évaluer sa satisfaction par rapport au logiciel et par rapport à l'appui à la décision clinique fourni par cet outil.

## 3. Résultats

### 3.1 Modèles prédictifs des risques liés à la COVID-19 et leur validation

Au total, 74.850 observations relatives à 1376 patients COVID-19 de plus de 18 ans hospitalisés au CHU de Liège entre le 14 juillet 2020 et le 9 avril 2021 ont été utilisées, dont 80% pour la constitution des modèles (*Training sample*) et 20% pour les tests (*Test sample*).

Un nouvel échantillon de patients 234 patients COVID-19 de plus de 18 ans hospitalisés au CHU de Liège après le 4 mai 2021 et sortis au 3 novembre 2021 a été utilisé pour la validation (*Validation sample*).

Les taux réels de décès à l'hôpital, d'admission en USI, de transfusion et de ventilation mécanique dans les 3 échantillons sont donnés au Tableau 1.

Tableau 1. Description des résultats réels pour les patients pris en compte pour la construction des modèles et leur validation

	Construction des modèles (Training sample)		Construction des modèles (Test sample)		Validation des modèles (Validation sample)	
	N	n (%) événements	N	n (%) événements	N	n (%) événements
Décès à l'hôpital	683	76 (11%)	161	20 (12%)	234	31 (13%)
Admission USI	641	111 (17%)	161	28 (17%)	234	55 (23%)
Transfusion	643	66 (10%)	162	17 (10%)	234	8 (3%)
Ventilation mécanique	651	55 (8%)	163	14 (9%)	234	25 (11%)

A partir des modèles obtenus, pour chaque résultat (risque de décès, d'admission aux soins intensifs, de ventilation mécanique, et de transfusion,) 3 groupes ont été créés : risque faible, ambigu, ou élevé. Le Tableau 2 donne la description des groupes dans les 3 échantillons. Pour chaque résultat et chaque classe de risque, l'estimation de la probabilité de l'événement dans l'échantillon de validation est proche de celle calculée sur l'échantillon test.

Tableau 2. Constitution et validation des groupes de risques

Groupe de risque	Seuils	Construction des modèles (Training sample)			Construction des modèles (Test sample)			Validation des modèles (Validation sample)			
		N	n	Probabilité estimée [IC 95%]	N	n	Probabilité estimée [IC 95%]	N	n	Probabilité estimée [IC 95%]	
Décès à l'hôpital	Faible	<0.12	424	0	0.0 [0.0 ; 0.0]	91	3	0.033 [0.0 ; 0.07]	134	7	0.052 [0.015 ; 0.090]
	Ambigu	0.12 – 0.22	99	7	0.071 [0.02 ; 0.12]	33	7	0.21 [0.07 ; 0.28]	74	14	0.19 [0.10 ; 0.28]
	Elevé	>0.22	115	69	0.60 [0.51 ; 0.69]	37	10	0.45 [0.27 ; 0.63]	26	10	0.39 [0.20 ; 0.57]
Admission USI	Faible	<0.17	379	0	0.0 [0.0 ; 0.0]	81	6	0.074 [0.02 ; 0.13]	98	14	0.14 [0.074 ; 0.21]
	Ambigu	0.17 – 0.28	142	18	0.13 [0.07 ; 0.18]	51	9	0.18 [0.07 ; 0.28]	98	24	0.24 [0.16 ; 0.33]
	Elevé	>0.28	120	93	0.78 [0.07 ; 0.85]	29	13	0.45 [0.27 ; 0.63]	38	17	0.45 [0.29 ; 0.61]
Transfusion	Faible	<0.1	471	0	0.0 [0.0 ; 0.0]	105	5	0.048 [0.01 ; 0.09]	128	1	0.010 [0.0 ; 0.023]
	Ambigu	0.1 – 0.16	54	4	0.074 [0.0 ; 0.14]	27	4	0.15 [0.01 ; 0.28]	51	0	-
	Elevé	>0.16	118	62	0.53 [0.44 ; 0.62]	30	8	0.27 [0.11 ; 0.42]	55	7	0.13 [0.039 ; 0.21]
Ventilation mécanique	Faible	<0.08	387	0	0.0 [0.0 ; 0.0]	89	3	0.034 [0.0 ; 0.07]	35	1	0.029 [0.0 ; 0.084]
	Ambigu	0.08 – 0.23	215	17	0.079 [0.04 ; 0.12]	59	6	0.10 [0.02 ; 0.18]	196	23	0.12 [0.072 ; 0.16]
	Elevé	0.23	49	38	0.78 [0.66 ; 0.89]	15	5	0.33 [0.09 ; 0.57]	3	1	0.33 [0.0 ; 0.87]

L'analyse de la sensibilité des modèles dans l'échantillon test montre que ceux-ci sont efficaces pour prédire les patients à haut ou à faible risque (Tableau 3). La validation des modèles montre que la sensibilité n'est pas aussi élevée que dans l'échantillon test mais reste bonne, à l'exception de la prédiction du besoin de ventilation mécanique.

Tableau 3. Sensibilité et spécificité des modèles (classement en 3 sous-groupes, hors ambigus)

	Test sample		Validation sample	
	Sensibilité	Spécificité	Sensibilité	Spécificité
Décès à l'hôpital	0.77	0.77	0.59	0.89
Admission USI	0.68	0.82	0.55	0.80
Transfusion	0.62	0.89	0.88	0.73
Ventilation mécanique	0.63	0.90	0.50	0.94

### 3.2 Etude de cas

L'étude de cas a porté sur 47 patients âgés de  $69 \pm 14$  ans en moyenne, dont une majorité d'hommes (64%) ; ils avaient un IMC moyen de  $27.1 \pm 4.6$ . Au total 57% d'entre eux sont décédés pendant le séjour hospitalier, 61% ont été admis aux soins intensifs, 43% ont nécessité le recours à une ventilation mécanique et 38% à une transfusion.

La Figure 1 donne un aperçu de l'outil CDSS mis à la disposition du médecin spécialiste lors de la seconde étape.

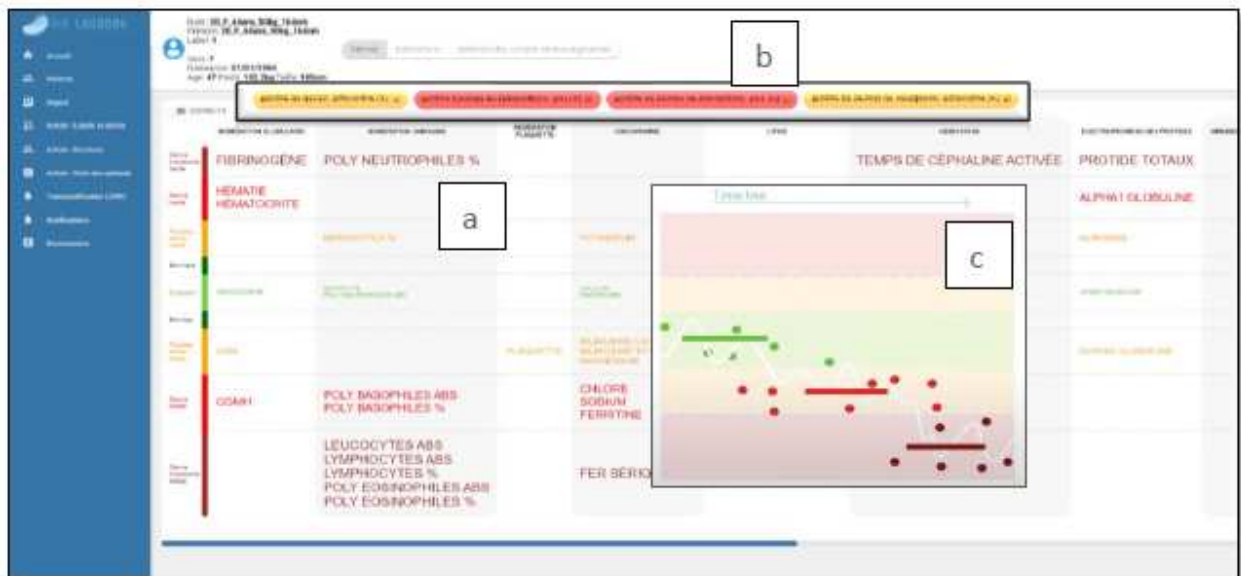


Figure 1. CDSS Bio Logbook mis à disposition du médecin spécialiste, présentant pour chaque patient une table synthétique des paramètres biologiques et de leur dérive (a), les scores de risques de décès, de passage aux soins intensifs, de ventilation mécanique et de transfusion (b), des graphes d'évolution des paramètres accessibles par 'click' sur ceux-ci dans la table (c).

L'utilisation du CDSS par le médecin semble augmenter la sensibilité et la spécificité lors de l'estimation du risque de décès du patient. La sensibilité de 75% pour le médecin sans l'outil augmente à 81% avec l'outil. Elle était de 67% pour l'outil seul (sans le médecin) (Tableau 4). Le risque de transfusion est également mieux identifié lorsque le médecin utilise l'outil. En ce qui concerne le passage aux soins intensifs, l'outil seul donne les meilleurs résultats, avec une sensibilité de 81% et une spécificité de 80%. Le CDSS est moins performant pour la détection des patients à risque de ventilation mécanique car il classe la majorité de ceux-ci dans le groupe à risque ambigu.

Tableau 4. Etude de cas : résultats

	Sensibilité [95%CI]	Spécificité [95%CI]
Décès à l'hôpital		
Médecin sans CDSS	0.75 [0.56 - 0.94]	0.35 [0.14 - 0.56]
Médecin avec CDSS	0.81 [0.64 - 0.98]	0.44 [0.21 - 0.67]
CDSS seul	0.67 [0.43 - 0.91]	0.93 [0.81 - 1.0]
Admission USI		
Médecin sans CDSS	0.77 [0.54 - 1.0]	0.35 [0.13 - 0.58]
Médecin avec CDSS	0.74 [0.56 - 0.92]	0.38 [0.14 - 0.61]
CDSS seul	0.81 [0.62 - 1.0]	0.80 [0.55 - 1.0]
Transfusion		
Médecin sans CDSS	0.50 [0.27 - 0.73]	0.90 [0.79 - 1.0]
Médecin avec CDSS	0.69 [0.44 - 0.94]	0.61 [0.41 - 0.81]
CDSS seul	0.54 [0.27 - 0.81]	0.96 [0.88 - 1.0]
Ventilation mécanique		
Médecin sans CDSS	0.77 [0.54 ; 1.0]	0.32 [0.14 ; 0.50]
Médecin avec CDSS	0.73 [0.51 ; 0.96]	0.33 [0.14 ; 0.52]
CDSS seul	-	-

#### 4. Discussion/Conclusion

L'outil d'aide à la décision clinique développé et validé spécifiquement lors de cette étude pour le suivi des patients COVID-19 hospitalisé montre des résultats encourageants, en particulier pour la détection précoce des patients particulièrement à risque de décès lors de leur hospitalisation. L'outil fournit ainsi au clinicien une aide lui permettant d'optimiser la rapidité et la qualité de la prise en charge des patients à partir de l'analyse des résultats de laboratoire à l'admission ; il permet aussi d'augmenter la capacité du médecin à identifier les patients à haut ou à plus faible risque de décès dans le séjour.

Le profil des patients COVID-19 hospitalisés, leur prise en charge, et leurs traitements ont largement évolué au cours du temps depuis le début de l'épidémie. On observe malgré tout que l'outil reste valide d'une vague à l'autre, à l'exception de la prédiction du besoin de ventilation mécanique.

La comparaison de l'estimation des risques par l'outil seul, le médecin seul, et le médecin avec l'outil montre que les meilleurs résultats sont obtenus par le médecin aidé de l'outil. Ceci met une fois de plus en évidence l'importance de la synergie médecin-CDSS.

L'enquête de satisfaction a révélé un intérêt pour la méthodologie utilisée et pour le potentiel de l'outil d'aide à la décision clinique, en particulier pour le suivi de patients atteints d'autres pathologies, notamment en oncologie et en hématologie clinique.

L'étude présente cependant quelques limites. Idéalement l'étude de cas devrait être réalisée par plusieurs médecins, de spécialités différentes afin de prendre en compte la variabilité des réponses entre médecins et entre spécialités. D'autre part, l'étude a permis de valider les résultats fournis par l'outil et de mettre en avant sa plus-value clinique, mais il faudrait analyser de manière prospective si son utilisation en pratique conduit effectivement à une modification de la prise en charge du patient. Une fois déployé dans l'hôpital il serait important de réaliser une validation prospective et régulière, voire automatique, de l'outil.

D'autres méthodologies pourraient également être explorées pour la construction des algorithmes de prédiction. Par exemple les techniques de modélisation des changements d'états du type « change point detection » offrent une piste intéressante.

## Remerciements

Ce projet a été soutenu financièrement par le programme Européen Horizon 2020 COVID eXponential Programme (COVID-X).

## Références

Bright T.J., Wong A., Dhurjati R., Bristow E., Bastian L., Coeytaux R. R., Smasa G., Hasselblad V., Williams J. W., Musty M. D., Wing L., Kendrick A. S., Sanders G. D., Lobach D. (2012). Effects of Clinical Decision-Support Systems. *Annals of Internal Medicine*, 157, 29-43

Guiot J., Vaidyanathan A., Deprez L., Zerka F., Danthine D., Frix A.-N., Thys M., Henket M., Canivet G., Mathieu S., Eftaxia E., Lambin P., Tsoutzidis N., Miraglio B., Walsh S., Moutschen M., Louis R., Meunier P., Vos W., Leijenaar R.T.H., Lovinfosse P. (2021). Development and Validation of an Automated Radiomic CT Signature for Detecting COVID-19. *Diagnostics*, 11, 41

Shahsavarani A. M., Abadi E. A. M., Kalkhoran M. H., Jafari S., Qaranli S. (2015). Clinical Decision Support Systems (CDSSs) : State of the art Review of Literature. *International Journal of Medical Reviews*, Volume 2, Issue 4, 299-308

Wu G., Yang P., Xie Y., Woodruff H., Rao X., Guiot J., Frix A.-N., Louis R., Moutschen M., Li J., Li J., Yan C., Du D., Zhao S., Ding Y., Liu B., Sun W., Albarello F., D'abramo A., Schininà V., Nicastrì E., Occhipinti M., Barisione G., Barisione E., Halilaj I., Lovinfosse P., Wang X., Wu J., Lambin P. (2020). Development of a clinical decision support system for severity risk prediction at hospital admission: an international multicenter study. *Eur Respir J*, 56:2001104

Breiman, L. Random Forests. *Machine Learning* 45, 5–32 (2001)