

UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG

Année 2019 - 2020

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

**Analyse des composés organiques volatils dans l'air exhalé
de patients atteints d'un cancer broncho-pulmonaire
et de témoins sains.**

Présentée et soutenue publiquement le 16 juin 2020 à 16h00
au Pôle Recherche
par **Clémence RÉMY**

JURY

Président :

Monsieur le Professeur Régis MATRAN

Assesseurs :

Monsieur le Docteur Sébastien HULO

Monsieur le Docteur Clément GAUVAIN

Directeur de thèse :

Monsieur le Professeur Arnaud SCHERPEREEL

AVERTISSEMENT

La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses : celles-ci sont propres à leurs auteurs.

RÉSUMÉ

Titre de la thèse : Analyse des composés organiques volatils dans l'air exhalé de patients atteints d'un cancer broncho-pulmonaire et de témoins sains.

Contexte : Le dépistage du cancer broncho-pulmonaire (CBP) par scanner thoracique permet une réduction de mortalité par ce cancer mais présente des limites (irradiation, coût). Les composés organiques volatils (COVs), mesurés dans l'air exhalé, pourraient permettre de mieux identifier la population cible au dépistage par scanner. L'objectif est de décrire et comparer les COVs retrouvés chez des patients atteints d'un CBP et des témoins sains puis de confirmer les résultats des COVs déjà décrits dans la littérature comme associés au CBP.

Méthode : Nous avons mené une étude cas-témoin interventionnelle multicentrique. Les patients atteints d'un CBP et naïfs de traitement ont été recrutés au CHU de Lille et les témoins sains à l'institut Pasteur de Lille. L'air exhalé a été collecté dans des tubes de sorbant via un dispositif de recueil validé puis analysé dans un laboratoire indépendant par une technique de chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC-MS).

Résultats : Quatorze cas et 15 témoins ont été inclus. Nous avons mis en évidence 11 COVs associés au CBP parmi lesquels 5 étaient déjà décrits dans la littérature : 2-propanol ($p < 0,001$), 2,4-diméthylheptane ($p < 0,001$), benzène ($p < 0,001$), éthanol ($p < 0,001$) et toluène ($p = 0,001$). Six COVs inédits étaient également associés au CBP : cyclotétrasiloxane, octaméthyl- ($p = 0,005$), cyclotrisiloxane, hexaméthyl- ($p < 0,001$), disiloxane, 1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)- ($p = 0,008$), disiloxane, hexaméthyl- ($p < 0,001$), triméthylsilanol ($p < 0,001$) et trisiloxane, octaméthyl- ($p < 0,001$).

Conclusion : Nous avons identifié 11 COVs associés au CBP parmi lesquels 6 appartiennent à la famille des siloxanes et n'ont jamais été décrits dans la littérature. Ces résultats et leur implication diagnostique méritent d'être confirmés au sein d'une population plus importante.

SOMMAIRE

1. INTRODUCTION	1
1.1. Cancer broncho-pulmonaire	1
1.2. Dépistage	3
1.2.1. Définitions	3
1.2.2. Dépistage du CBP	5
1.2.3. Imagerie thoracique	6
1.2.4. Biomarqueurs	9
1.3. Composés organiques volatils (COVs)	9
1.3.1. Définitions	9
1.3.2. Recueil et Mesure	10
1.3.3. COVs et CBP	13
1.4. Objectifs	14
2. MÉTHODE	16
2.1. Type d'étude	16
2.2. Population	16
2.2.1. Critères d'inclusion	16
2.2.2. Critères de non-inclusion	17
2.3. Déroulement de l'étude	17
2.4. Données recueillies	18
2.5. Mesures	19
2.5.1. Recueil	19
2.5.2. Analyses	20
2.6. Paramètres étudiés	21
2.7. Analyses statistiques	22
2.8. Cadre réglementaire	23
3. RÉSULTATS	24
3.1. Description de la population à l'inclusion	24
3.1.1. Caractéristiques globales	26
3.1.2. Caractéristiques des cancers broncho-pulmonaires	27
3.2. Description des conditions de recueil	28
3.3. Analyse des COVs	29
3.3.1. Concordance des mesures	29
3.3.2. Description et comparaison des COVs	30
3.4. Modélisation d'un nez électronique	32
3.4.1. Approche par sommation des COVs associés au CBP	32
3.4.2. Approche par régression logistique et score prédictif	32
4. DISCUSSION	34
5. CONCLUSION	44
6. BIBLIOGRAPHIE	45

LISTE DES ABRÉVIATIONS

CBP :	Cancer broncho-pulmonaire
INCa :	Institut national du cancer
CBNPC :	Cancer broncho-pulmonaire non à petites cellules
CPC :	Cancer broncho-pulmonaire à petites cellules
OMS :	Organisation mondiale de la santé
PA :	Paquets-années
COVs :	Composés organiques volatils
CAS :	Chemical abstracts service
GC-MS :	Chromatographie en phase gazeuse couplée à la spectrométrie de masse
CHU :	Centre hospitalier universitaire
BPCO :	Broncho-pneumopathie chronique obstructive
CPES :	Centre de prévention et d'éducation à la santé
IMC :	Indice de masse corporelle
ROC :	<i>Receiver Operating Curve</i>
AUC :	Aire sous la courbe
CNIL :	Commission nationale de l'informatique et des libertés
DIQ :	Distance inter-quartile

1. INTRODUCTION

1.1. Cancer broncho-pulmonaire

Dans le monde, le cancer broncho-pulmonaire (CBP) est le cancer le plus fréquemment diagnostiqué et représente la première cause de décès par cancer (1). En France, d'après l'institut national du cancer (INCa), le CBP est le 2^e cancer solide incident chez l'homme et le 3^e chez la femme. Il touche aussi bien les hommes que les femmes avec des taux d'incidence estimés à 50,5 et 23,3 pour 100 000 personnes-années correspondant à 31 231 et 15 132 nouveaux cas incidents en 2018, respectivement (2). Bien que chez l'homme, le taux d'incidence reste plutôt stable, chez la femme on constate une forte augmentation avec un taux d'incidence multiplié par 4,3 entre 1990 et 2018 (2). Son pronostic est globalement mauvais avec une survie globale tous stades confondus estimée de 15 % à 5 ans (3). Le CBP est au 1^{er} rang de décès par cancer chez l'homme et au 2^e chez la femme (passé au 1^{er} rang aux États-Unis) (1,2). Du fait de son incidence et de son pronostic globalement sombre, le CBP représente donc un véritable enjeu de santé publique.

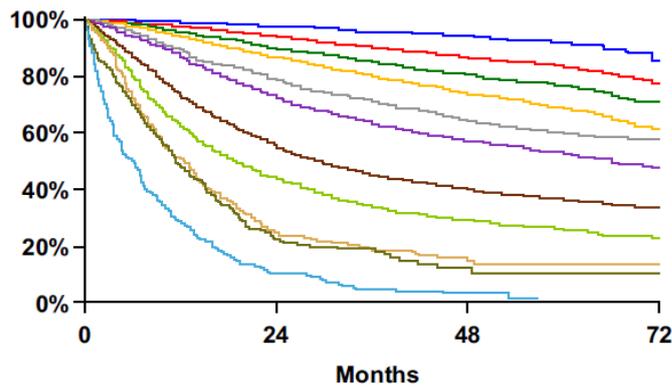
Le diagnostic positif repose sur des prélèvements histologiques (pulmonaires, ganglionnaires, métastatiques). On distingue deux grands types histologiques : le CBP non à petites cellules (CBNPC) qui représente la majorité des cas (environ 85 %) et le CBP à petites cellules (CPC) (environ 15 % des cas) (2).

L'âge médian au diagnostic est de 67 ans chez l'homme et 65 ans chez la femme (2). Le tabac est le principal facteur de risque, responsable d'environ 80 % des CBP (4). Les carcinogènes professionnels représentent le deuxième facteur de risque avec

une fraction attribuable à une étiologie professionnelle d'environ 15 % (4). Enfin, certains facteurs environnementaux sont considérés comme carcinogènes mais leur imputabilité est compliquée à prouver du fait d'une durée d'exposition variable, d'expositions combinées et de seuils de toxicités non connus (5).

Les signes d'appels cliniques sont peu spécifiques et souvent tardifs expliquant le fait que la majorité des CBP soient diagnostiqués à un stade localement avancé ou métastatique (3). Ils peuvent aussi être découverts de manière fortuite et précoce sous la forme de nodules pulmonaires au scanner thoracique. Néanmoins, les étiologies potentielles de ces nodules sont nombreuses (cancer primitif ou métastases, tumeurs bénignes, infections, maladies granulomateuses, etc.) (6). Il existe donc un enjeu diagnostique majeur pour ces nodules car il ne faut pas méconnaître une lésion cancéreuse précoce, peut-être à un stade localisé.

Au terme d'un bilan d'extension, on classe le cancer en stade localisé (I et II), localement avancé (III) ou disséminé (IV) (7). Le stade au diagnostic est le principal facteur pronostique et conditionne la stratégie thérapeutique (3). Aussi, pour les stades localisés, on envisage un traitement à visée curative par chirurgie ou radiothérapie permettant une survie à 5 ans de plus de 70 % (7). La survie à 10 ans pour les CBP de stade I opérés est même estimée à 92 % (8). En revanche, pour les stades avancés, la stratégie reste palliative avec une survie à 5 ans de moins de 10 % (Figure 1) (7) malgré de récents progrès thérapeutiques majeurs (9). Ainsi, les CBP diagnostiqués à un stade localisé sont associés à un meilleur pronostic car éligibles à un traitement à visée curative (3,7,8). Ceci suggère l'intérêt d'un diagnostic précoce chez des patients asymptomatiques et pose donc la question du dépistage (3).



Proposed	Events / N	MST	24 Month	60 Month
IA1	68 / 781	NR	97%	92%
IA2	505 / 3105	NR	94%	83%
IA3	546 / 2417	NR	90%	77%
IB	560 / 1928	NR	87%	68%
IIA	215 / 585	NR	79%	60%
IIB	605 / 1453	66.0	72%	53%
IIIA	2052 / 3200	29.3	55%	36%
IIIB	1551 / 2140	19.0	44%	26%
IIIC	831 / 986	12.6	24%	13%
IVA	336 / 484	11.5	23%	10%
IVB	328 / 398	6.0	10%	0%

Figure 1. Survie globale pour les CBNPC en fonction du stade TNM (8^{ème} édition). D'après Goldstraw et al. (7)

1.2. Dépistage

1.2.1. Définitions

Selon l'Organisation Mondiale de la Santé (OMS), le dépistage « consiste à identifier de manière présomptive à l'aide de tests, appliqués de façon systématique et standardisée, les sujets atteints d'une maladie ou d'une anomalie passée jusque-là inaperçue. Les tests de dépistage doivent permettre de faire le partage entre les personnes apparemment en bonne santé mais qui sont probablement atteintes d'une maladie ou d'une anomalie donnée et celles qui en sont probablement exemptes. Ils n'ont pas pour objet de poser un diagnostic » (10).

L'OMS a établi des critères stricts pour l'éligibilité d'une pathologie à une politique de dépistage. Il doit s'agir d'une « menace grave pour la santé publique », dont l'histoire naturelle est connue et avec une phase de latence asymptomatique. Cette maladie doit être décelable à un stade précoce, dans une population prédéfinie chez qui un traitement efficace à visée curative peut être administré (10).

L'efficacité d'un programme de dépistage doit être prouvée par des essais randomisés contrôlés et non simplement supposée à partir de l'analyse de données de cohortes analysant la survie selon le stade. Un diagnostic précoce de formes localisées entraînera mécaniquement une survie prolongée (par allongement de la durée de la maladie) sans préjuger d'une réduction de la mortalité. Ceci correspond au biais du temps de devancement (11,12). Il s'y ajoute aussi un phénomène de sélection et d'accumulation des cas de meilleur pronostic dont la survie est prolongée (11). En effet, les formes agressives étant rapidement évolutives et symptomatiques, elles sont statistiquement moins souvent dépistées à la phase préclinique. Ensuite, le biais du surdiagnostic (11,12) se rapporte à la découverte de lésions qui ne se seraient jamais manifestées cliniquement en raison d'une évolution très lente, associées à une survie prolongée et pour lesquelles on propose des prises en charges diagnostiques et thérapeutiques inutiles. Le taux de participation est également un bon indicateur de l'acceptation de l'examen. Il existe par ailleurs un potentiel biais de sélection par l'inclusion de patients nécessairement volontaires et donc attentifs à leur santé, avec possiblement un meilleur état de santé (11). Ces difficultés imposent que l'évaluation de l'efficacité d'un programme de dépistage repose sur des essais randomisés contrôlés avec une méthodologie stricte dont le critère de jugement est la mortalité spécifique et non simplement la survie des cas dépistés.

1.2.2. Dépistage du CBP

Le CBP remplit les critères de l'OMS d'éligibilité à une politique de dépistage sur le plan épidémiologique (incidence, pronostic), clinique (phase de latence asymptomatique), thérapeutique (stratégies curatives pour les stades précoces) et de l'histoire naturelle (symptomatologie tardive et dissémination) (2,7,8,10).

La population à dépister doit être sélectionnée sur des critères prédéfinis. Le tabac représente le principal facteur de risque du CBP (4) et l'incidence du CBP augmente considérablement à partir de 50 ans dans les deux sexes (2). De plus, les sujets doivent être suffisamment robustes pour envisager une chirurgie thoracique (13). Actuellement, les principales études sur le dépistage du cancer du poumon ont retenu comme critères d'inclusion un âge entre 50 et 74 ans et une consommation tabagique d'au moins 15 paquets-années (PA) (14,15). Ainsi, les patients tabagiques âgés de 50 à 75 ans pourraient constituer la population cible.

La mise en œuvre d'une politique de dépistage du CBP peut toutefois être discutée devant la possibilité d'une prévention primaire par éviction du facteur de risque principal : le tabac (16). La lutte anti-tabac s'intensifie en France comme en témoigne le remboursement des substituts nicotiques depuis le 1^{er} janvier 2019. Le sevrage tabagique a en effet montré une diminution du risque de CBP proportionnelle à la durée de l'arrêt et une diminution de la mortalité globale et spécifique par CBP (16). Néanmoins, le risque de CBP persiste plusieurs années après l'arrêt d'où l'importance d'un dépistage associé. Aussi, cette stratégie de dépistage du CBP ne s'entend pas sans proposer systématiquement aux patients un sevrage tabagique (17). D'ailleurs, il a été montré qu'au cours d'un programme de dépistage, les taux de sevrage

tabagique étaient augmentés (17) et que l'arrêt du tabac était associé à une diminution de la mortalité globale par rapport à un fumeur toujours actif (18).

Si le dépistage du CBP est justifiable, sa mise en œuvre nécessite l'existence d'un test qui soit simple, efficace, fiable, reproductible, disponible, acceptable pour la population et peu coûteux (10).

1.2.3. Imagerie thoracique

Évaluée depuis les années 1970, la radiographie thoracique n'a pas montré son efficacité sur la réduction de la mortalité spécifique et globale dans le dépistage du CBP (19), contrairement au scanner thoracique.

En 2011 aux États-Unis, l'étude NLST (14) a évalué un dépistage annuel par scanner thoracique faible dose sur 3 ans versus radiographie de thorax chez 53 454 patients, âgés de 55 à 74 ans et tabagiques actifs ou sevrés depuis moins de 15 ans avec une consommation cumulée minimale à 30 PA. Toute anomalie au scanner de plus de 4 mm de grand axe était considérée comme suspecte et entraînait des investigations. Il a été montré une réduction de la mortalité par CBP de 20 % (IC95% [6,8 - 26,7] ; $p = 0,004$) dans le groupe scanner en comparaison au groupe radiographie de thorax ainsi qu'une réduction de la mortalité globale de 6,7 % (IC95% [1,2 - 13,6] ; $p = 0,02$). Aussi, les recommandations américaines, qui s'appuient notamment sur cette étude, ont généralisé le dépistage annuel du CBP par scanner thoracique pour les patients âgés entre 55 et 80 ans avec un tabagisme actif ou sevré depuis moins de 15 ans et une consommation cumulée d'au moins 30 PA (20).

Néanmoins, il faut noter que l'une des limites principales de l'étude NLST est le taux très élevé de faux positifs (96,4 %) dans le groupe scanner.

Plus récemment, en Europe, l'étude NELSON (15) a inclus 13 195 hommes âgés de 50 à 74 ans, tabagiques actifs ou sevrés depuis moins de 10 ans et ayant une consommation de plus de 15 cigarettes par jour pendant au minimum 25 ans, ou plus de 10 cigarettes par jour pendant au minimum 30 ans. Dans le groupe scanner, l'examen était fait à l'inclusion, puis renouvelé à 1, 3 et 5,5 ans ; le groupe contrôle ne recevait aucune intervention. Les anomalies au scanner étaient mesurées par une approche volumétrique avec l'introduction de la notion de test indéterminé pour des volumes intermédiaires entre 50 et 500 mm³. Dans ces cas-là, le scanner était renouvelé à 3 mois afin d'évaluer le temps de doublement de volume de la lésion. En fonction de l'évolutivité du nodule, le patient était alors classé en dépistage positif ou négatif. Avec cette méthode de classification, 9,2 % des dépistages étaient considérés comme indéterminés et le taux de faux positifs était beaucoup moins élevé (56,5 %) que dans NLST. Cette étude a montré une diminution de la mortalité spécifique par CBP de 25 % dans le groupe scanner à 10 ans de suivi par rapport au groupe contrôle [rapport des taux à 0,76 (IC95% [0,61 - 0,95] ; p = 0,01)].

Ces deux essais randomisés de forte puissance permettent de conclure au bénéfice du dépistage du CBP par scanner thoracique sur la mortalité spécifique par CBP. Au vu des taux de participation [95 % pour NLST (14) et 87,6 % pour NELSON pour au moins 3 examens (15)], le scanner thoracique semble être acceptable pour la population et son implémentation faisable en vie réelle comme démontré par l'étude en France « DK80 » de Leleu et al, impliquant fortement les médecins traitants (21).

Le coût d'un dépistage par scanner associé à une prise en charge chirurgicale (ou radiothérapie) d'un stade localisé permettrait des économies de santé en comparaison aux traitements systémiques des stades avancés (22). Gendarme et al. ont également montré que le dépistage par scanner pourrait être financé par une augmentation du prix du paquet de cigarettes de 0,05 à 0,10 centimes d'euros par paquet en France (23).

En revanche, certaines questions restent à élucider et notamment le rythme et la durée du dépistage. De plus, le scanner est invasif et expose à un risque d'irradiation même à faible dose (12). Grâce à l'approche volumétrique dans l'étude NELSON, le taux de faux positifs est nettement réduit ce qui engendre moins d'explorations invasives inutiles diminuant ainsi le risque de iatrogénie et de complications par rapport à l'étude NLST (14,15). Cependant, les méthodes d'interprétation des scanners doivent être uniformisées et standardisées dans le but de diminuer ce taux de faux positifs tout en conservant l'effet sur la mortalité (12,24).

Ces différents éléments font que l'utilisation du scanner thoracique pour le dépistage du CBP en routine n'est pas validée à ce jour en France (5). Néanmoins, ces recommandations sont susceptibles d'évoluer rapidement selon les résultats des études récentes, notamment l'étude NELSON (15). Aussi, ces limites font rechercher des méthodes alternatives ou complémentaires au scanner thoracique, moins invasives, qui permettraient de mieux cibler la population du dépistage et/ou d'explorer des anomalies pulmonaires mises en évidence au scanner ; par exemple les biomarqueurs (24,25).

1.2.4. Biomarqueurs

Dans le champ du dépistage et du diagnostic du cancer pulmonaire, plusieurs biomarqueurs sont étudiés, parfois en complément de l'imagerie. On leur distingue alors trois objectifs : mieux identifier la population cible en amont du scanner (pré-screening), aider à différencier les nodules bénins et malins dans la population dépistée pour limiter le taux de faux positifs et les explorations inutiles (24,25) ainsi que prédire la survenue d'un cancer avant l'apparition d'anomalies scanographiques (26).

Ils peuvent provenir des divers fluides organiques, tels que le sang, la salive, les expectorations, les urines, le liquide pleural ou l'air exhalé. Il s'agit de protéines tumorales (antigène carcino-embryonnaire, fragment 21-1 de la cytokératine, etc.), de cellules tumorales circulantes, d'ADN tumoral, de microARN ou encore de composés organiques volatils (COVs) (25). Par exemple, la détection de cellules tumorales circulantes dans le sang pourrait avoir une double utilité : prédire la survenue d'un cancer plusieurs mois avant son apparition scanographique et orienter vers une étiologie maligne en cas de nodule pulmonaire (26). De même, les COVs apparaissent comme des biomarqueurs intéressants pour le dépistage du CBP (27) et l'orientation étiologique des nodules (28).

1.3. Composés organiques volatils (COVs)

1.3.1. Définitions

Un « composé organique » est un « composé contenant au moins l'élément carbone et un ou plusieurs des éléments suivants : hydrogène, halogène, oxygène, soufre,

phosphore, silicium ou azote ». Il est dit « volatil » si sa « pression de vapeur est de 0,01 kPa ou plus à une température de 293,15 Kelvin (soit 20 °C), ou s'il a une volatilité correspondante dans des conditions d'utilisation particulières » (29). Les COVs sont classés en famille selon leur composition chimique. On distingue :

- Les hydrocarbures aliphatiques avec notamment les alcanes et les alcènes
- Les hydrocarbures alicycliques incluant les terpènes
- Les hydrocarbures aromatiques
- Les hydrocarbures substitués regroupant les alcools, les cétones, les éthers et esters, les acides, les nitriles, les aldéhydes.

Ils ont une double origine, à la fois endogène par les voies du métabolisme et exogène (produits ménagers, parfums, industries, peintures, essences, pesticides, alimentation, tabac, bactéries, moisissures,...) (29,30).

La terminologie de ces composés est vaste avec pour chacun de nombreux synonymes et un numéro d'enregistrement unique auprès du *Chemical abstracts service* (CAS) (31).

1.3.2. Recueil et Mesure

Les COVs peuvent circuler dans les différents compartiments corporels et notamment dans l'air exhalé (ou haleine ou exhalat). Aussi, l'exhalat semble être la source de choix pour recueillir et mesurer facilement les COVs grâce à des techniques non invasives, rapides et non douloureuses (32). Deux méthodes d'analyse sont couramment utilisées : l'une basée sur l'identification individuelle des COVs et l'autre sur la mise en évidence de combinaisons prédéfinies de COVs (figure 2) (30).

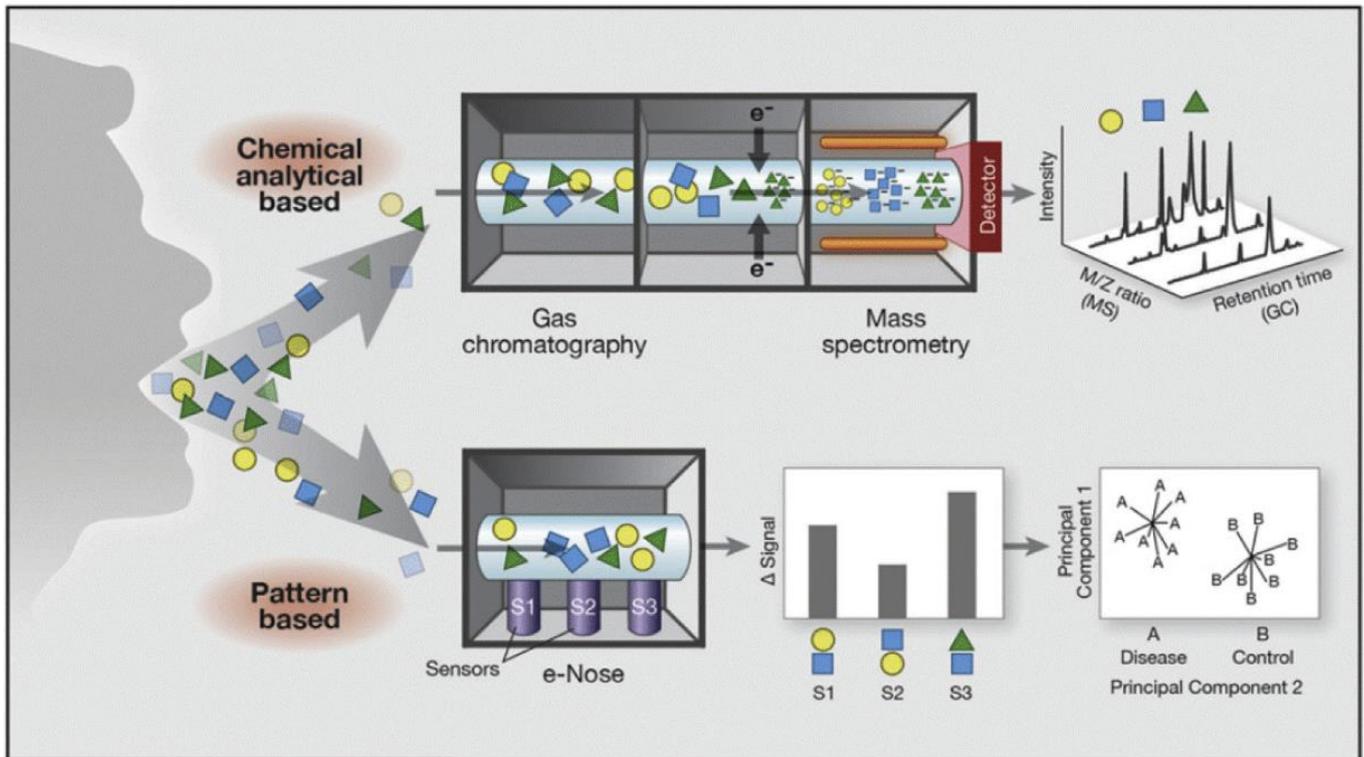


Figure 2. Méthodes d'analyse des COVs. D'après Rocco et al. (30)

La méthode d'identification individuelle des COVs nécessite tout d'abord de collecter un prélèvement d'air exhalé de qualité à l'aide d'un dispositif hermétique pour éviter toute contamination extérieure (30). Le volume suggéré minimum est d'au moins 500 mL (32) permettant de récupérer à la fois l'air des voies aériennes supérieures (150 mL) et l'air contenu dans les alvéoles (350 mL). Le temps de recueil a aussi son importance pour capter un maximum de COVs sanguins circulants dont les échanges avec les voies respiratoires sont continus au niveau de la barrière alvéolo-capillaire (32,33). Le récipient doit pouvoir concentrer un maximum de COVs et leurs propriétés physiques ne doivent pas être modifiées par les conditions de stockage au risque de pertes (33). Parmi les contenants disponibles, ce sont les tubes avec matériaux adsorbants type tenax qui permettent les meilleures conditions de recueil (33). Bien que les techniques de recueil ne soient pas standardisées (33), des dispositifs de recueil actuellement commercialisés ont montré des résultats fiables et

reproductibles (34). Ensuite, le prélèvement est analysé dans des laboratoires spécialisés par une technique de chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC-MS) (30). La GC-MS est considérée comme le « gold standard » pour l'étude des COVs (32,33). Les résultats sont alors rendus sous forme de pics de concentration en fonction d'une limite de détection fixée par les laboratoires (33). Cette méthode est longue, coûteuse et nécessite une expertise mais elle a l'avantage de pouvoir identifier individuellement les COVs (26,30,32).

L'autre technique consiste à reconnaître des groupements préétablis de COVs sur un échantillon d'air exhalé. On parle de nez électroniques. Le recueil peut se faire directement dans le dispositif de mesure ou dans des poches de recueil ensuite analysées par l'appareillage. Les COVs contenus dans l'air exhalé vont être en contact avec des capteurs et ainsi générer un signal en cas de présence. Cette technique est rapide, facile d'accès et plus économique. Néanmoins, elle permet uniquement de détecter des COVs pré-choisis et présents dans l'exhalat sans pouvoir les identifier (26,30,32). L'intérêt des nez électroniques est de pouvoir rapidement rendre un résultat « positif » ou « négatif » en fonction de la combinaison de COV préétablie. En contexte clinique, un résultat positif pourrait permettre par exemple d'identifier un patient à risque de CBP chez qui réaliser un scanner thoracique ou d'argumenter la prise en charge chirurgicale d'un nodule suspect découvert au scanner. Pour établir ces combinaisons, il faut néanmoins identifier les COVs spécifiques du CBP par la première technique d'identification (30). Actuellement, plus de 3000 COVs ont été identifiés dans l'air exhalé humain (30) et de nombreux travaux ont permis de mettre en évidence de COVs spécifiques du cancer du poumon (27,35–37).

1.3.3. COVs et CBP

Pour le CBP, il faut distinguer les COVs endogènes physiologiques mais altérés (modification des concentrations) par les voies métaboliques spécifiques du cancer (altérations protéiques, altérations génétiques, stress oxydatif, etc.) (27,30) et les COVs endogènes produits par les cellules tumorales. Il s'y ajoute les COVs endogènes produits par d'autres conditions pathologiques et les COVs exogènes présents par exemple dans l'air, l'eau, l'alimentation, les traitements ou encore les toxiques (tabac notamment) (32).

Les premières descriptions des COVs de l'air exhalé comme biomarqueurs du CBP datent de 1985. Gordon et al. avaient identifié certains COVs comme étant discriminants du CBP malgré un faible effectif (12 cas et 9 témoins) (35). Depuis, de nombreuses études, pour la plupart cas-témoins, ont été publiées mais leur interprétation est difficile du fait notamment d'une grande diversité des techniques de recueil et de mesure utilisées (27,30,36). Aussi, trois revues de littératures entre 1999 et 2017 permettent de recenser environ une vingtaine de COVs d'intérêt pour le dépistage du CBP (27,36,37).

Comme rapporté par Jia et al. (36), l'étude des COVs pour le dépistage du CBP soulève de nombreuses problématiques méthodologiques liées :

- Aux conditions d'échantillonnage (matériel utilisé, mode ventilatoire du patient, remplissage homogène des collecteurs, recueil des COVs en début – milieu – fin d'expiration, conditions ambiantes de la salle, analyses en laboratoire et seuils de détection)

- Aux risques de contaminations à chaque étape (matériels, recueil, transport, désorption) et aux risques de pertes (volume de recueil, préconcentration, contenant hermétique, transport)
- Aux COVs exogènes de l'environnement (alimentation, boisson, tabac, expositions domestiques et professionnelles, traitements, exercice physique)
- Aux COVs endogènes variant selon l'histologie et le stade du CBP
- Aux COVs endogènes variant avec l'âge, le sexe, les comorbidités autres notamment respiratoires (38).

Un des moyens proposés pour mieux identifier et solutionner ces problèmes serait de réaliser des études sur un plus grand effectif, avec l'utilisation de techniques validées pour la mesure des COVs (36). C'est d'ailleurs l'objectif d'au moins deux études en cours : une menée par l'équipe chinoise de Li et al. (39) et une menée au CHU de Lille par l'équipe d'Hulo et al. intitulée CATOCOV (40).

1.4. Objectifs

L'enjeu d'un dépistage précoce du CBP est majeur car il s'agit là d'un problème de santé publique grave et fréquent (2). De plus, il existe un traitement curatif pour les stades localisés qui permet une amélioration de la survie (3,7,8) ainsi qu'une économie du coût de la santé (22). Le dépistage du CBP par scanner thoracique seul a montré son efficacité pour diminuer la mortalité spécifique (14,15) mais présente des limites en termes d'efficacité et n'est pas dénué de risques (12,24). Aussi les biomarqueurs semblent être des outils complémentaires intéressants afin de mieux cibler la population éligible au dépistage (24,25).

Les COVs, biomarqueurs dans l'air exhalé, ont déjà fait l'objet de nombreuses études dans le champ du CBP. Ils pourraient être utilisés en complément du scanner pour aider à identifier la population à risque et orienter l'étiologie des nodules indéterminés dépistés au scanner (27,28). Néanmoins, de nombreux facteurs interfèrent dans la recherche de COVs : voies du métabolisme, techniques de recueil et mesure, contaminations, etc. (29,30,32,36). L'enjeu est donc d'identifier ces différentes problématiques et de proposer des solutions afin de permettre la mise en place à terme d'une stratégie de dépistage par COVs qui soit fiable et efficace.

L'**objectif principal** de cette étude est donc de décrire et comparer les COVs retrouvés dans l'air exhalé de patients atteints d'un CBP et de témoins sains puis de confirmer les résultats des COVs déjà décrits comme associés au CBP.

Les **objectifs secondaires** sont :

- Après identification des COVs, d'évaluer plusieurs approches de détermination d'un seuil de positivité pour la détection de combinaisons de COVs, ainsi que l'estimation des performances diagnostiques associées, afin de modéliser les résultats possibles d'un nez électronique
- D'identifier les difficultés potentielles de cette procédure de dépistage de CBP par identification des COVs dans une population de patients atteints d'un CBP et dans une population témoin saine.

2. MÉTHODE

2.1. Type d'étude

Il s'agissait d'une étude transversale de type cas-témoins multicentrique. Les sujets étaient inclus du 29 avril 2019 au 31 janvier 2020. Les patients (cas) étaient inclus au Centre Hospitalier Universitaire (CHU) de Lille dans le service de Pneumologie et Oncologie Thoracique. Les témoins sains étaient inclus à l'Institut Pasteur de Lille.

2.2. Population

2.2.1. *Critères d'inclusion*

Pour être inclus, les **patients (cas)** devaient être âgés de 55 à 70 ans, avoir un diagnostic de CBP prouvé histologiquement, quels que soient le stade et l'histologie, et ne pas encore avoir reçu de traitement anti-cancéreux pour ce cancer.

Les **témoins** devaient avoir entre 55 à 70 ans et être fumeurs actifs ou non-fumeurs. La définition retenue pour le statut non-fumeur était celle des *Centers for Diseases Control and Prevention des États-Unis* (41) correspondant à moins de 100 cigarettes fumées ou équivalent sur toute la vie. Ils devaient être indemnes de toute pathologie respiratoire hormis une potentielle bronchopneumopathie chronique obstructive (BPCO). En cas d'exacerbation de BPCO récente, un intervalle libre d'au moins 1 mois était requis, ceci afin d'éviter tout risque de modification en aiguë des COVs endogènes, par exemple par une infection bactérienne (38) ou par des thérapeutiques spécifiques éventuelles (36).

2.2.2. Critères de non-inclusion

Nous avons exclu les sujets cas ou témoins :

- Aux antécédents de tout autre cancer dans les 5 ans précédant l'inclusion
- Ne pouvant pas recevoir une information éclairée en raison de troubles cognitifs, troubles de compréhension, barrière de la langue
- Présentant des malformations bucco-faciales empêchant le recueil de l'air exhalé
- Ne comprenant pas les manœuvres pour le recueil
- Ayant refusé de participer à l'étude.

Les cas n'étaient pas inclus si un traitement antitumoral avait été initié pour ce cancer avant l'inclusion.

Les témoins n'étaient pas inclus s'ils étaient des anciens fumeurs ou s'ils souffraient d'une pathologie respiratoire en dehors d'une BPCO fréquente sur ce terrain (par exemple asthme, pneumopathie interstitielle,...), ceci afin d'éviter d'identifier des COVs spécifiques d'autres pathologies respiratoires (38).

2.3. Déroulement de l'étude

L'inclusion des cas était faite lors d'une visite médicale de suivi (le plus souvent en consultation d'annonce avec le pneumologue oncologue référent) au CHU de Lille.

Les sujets témoins étaient inclus au décours d'un bilan de santé (examen clinique, prévention, nutrition) effectué sur la base du volontariat au centre de prévention et d'éducation à la santé (CPES) à l'Institut Pasteur de Lille.

Les sujets participant à cette étude étaient reçus au cours d'un unique entretien. Ils répondaient alors à des questionnaires médicaux et d'expositions, puis l'investigateur procédait au recueil de l'air exhalé des sujets.

2.4. Données recueillies

Lors de l'inclusion, à l'aide de l'interrogatoire du sujet et du dossier médical informatisé (pour les patients), les données suivantes étaient collectées pour chaque sujet :

- Le sexe et l'âge
- La profession et les expositions professionnelles éventuelles
- Les antécédents respiratoires, cardiaques, endocriniens
- Les antécédents de cancers solides et hématologiques
- Le statut tabagique : fumeur, actif, sevré, durée d'exposition et consommation cumulée estimée en PA
- L'utilisation de la cigarette électronique
- La prise d'autres toxiques
- Les traitements notamment par corticoïdes (par voie orale ou inhalée)
- Les paramètres anthropométriques [poids, taille, indice de masse corporelle (IMC)]
- Le Performance status selon l'échelle de l'OMS (OMS 0 à 4) (42).

Pour les patients, nous avons également recueilli :

- La date de diagnostic du CBP
- Les circonstances diagnostiques c'est-à-dire les symptômes ayant conduits au diagnostic (symptômes respiratoires, symptômes d'extension loco-régionale, symptômes d'extension métastatique)

- La topographie pulmonaire du CBP, la présence de nécrose au sein de la tumeur et l'envahissement de la lumière des voies aériennes
- Les modalités de réalisation des biopsies
- Les résultats des prélèvements histologiques et la présence de mutations ou réarrangements ou surexpressions
- Les sites d'atteintes extra-respiratoires en cas de stade métastatique
- Le stade TNM.

2.5. Mesures

2.5.1. *Recueil*

Le recueil était effectué par l'investigateur ou par un attaché de recherche clinique formé. Le sujet était installé dans une pièce dédiée dans chaque centre pour limiter les contaminations par des COVs de l'environnement et faciliter l'installation du matériel. L'investigateur plaçait un masque de type ReCIVA® (de l'entreprise OWLSTONE Ltd., basée au Royaume-Uni) (34) sur le visage du sujet. Il s'agissait d'un masque réutilisable sur lequel s'adaptaient par défaut 4 tubes de désorption thermique à base d'adsorbants contenant du TENAX.

Afin de bloquer les COVs exogènes présents dans l'environnement, le sujet respirait un air purifié par une colonne filtrante à base de charbon actif permettant d'adsorber un maximum de COVs ambiants (43). Il devait alors effectuer des mouvements respiratoires calmes. La durée de recueil prévisible était de 10 minutes pour un volume souhaité initialement de 500 mL. L'objectif de volume a secondairement été augmenté à 1000 mL en raison de temps de recueil trop courts. L'échantillonnage était suivi en direct sur un ordinateur via le logiciel fourni par OWLSTONE Ltd. connecté au masque.

Une fois le recueil terminé, le masque était retiré et les différents tubes désassemblés, recapuchonnés à chaque extrémité et anonymisés. Le sens d'échantillonnage de chaque tube était noté. Ils étaient ensuite conservés dans un réfrigérateur dédié avant envoi pour analyse à un laboratoire indépendant.

Certains paramètres liés aux conditions de recueil étaient collectés :

- La consommation d'aliments dans les 3 heures précédant la mesure
- La consommation de boissons dans les 3 heures précédant la mesure
- La prise de cigarette dans les 3 heures précédant le recueil
- Les éventuels soins dentaires en cours
- La réalisation d'un brossage de dents dans les 6 heures précédant le recueil
- Le volume recueilli (normalement fixé à 500 mL puis 1000 mL)
- La durée du recueil en secondes.

2.5.2. Analyses

Le choix du laboratoire d'analyse a nécessité une procédure de marché public. L'un des critères nécessaires était la disponibilité d'une salle blanche pour la réalisation des analyses afin d'éviter tout risque de contamination. C'est l'entreprise TERA Environnement (basée au 628 rue Charles de Gaulle, 38920 Crolles), spécialisée dans les contaminants chimiques de l'air, qui a été choisie. Les tubes étaient envoyés de manière anonymisée. Imposés par le modèle du masque de recueil, 4 tubes par patient étaient collectés. En raison du coût (87 euros par tube), seuls deux tubes étaient conservés et analysés par la technique de GC-MS. Sur les chromatogrammes obtenus, des pics de concentration étaient détectés et les COVs étaient alors identifiés en fonction de leurs poids moléculaires.

Les quatre tubes subissaient enfin une désorption thermique pour volatiliser tous les COVs et étaient renvoyés pour une nouvelle utilisation. Les résultats étaient rendus, pour chaque sujet, dans un délai estimé à environ 3 semaines, sous forme de tableur avec les concentrations en ng/support pour chaque COV identifié et dans les 2 tubes analysés. La limite de quantification fixée par le laboratoire était de 1 ng/support.

2.6. Paramètres étudiés

Le principal paramètre étudié était la distribution des COVs présents dans l'air exhalé des patients cas, comparée à celle des témoins.

Les paramètres d'études secondaires étaient :

- La description des difficultés potentielles de la procédure de dépistage du CBP par analyse de l'air exhalé, à savoir :
 - o La qualité de recueil (durée de recueil, volume de recueil)
 - o La concordance des résultats entre les deux tubes analysés
 - o L'adhésion des patients à la procédure de dépistage
- La détermination d'un seuil de positivité pour la détection d'une combinaison de COVs et l'évaluation des performances diagnostiques d'une telle approche.

2.7. Analyses statistiques

Les analyses statistiques étaient réalisées à l'aide du logiciel R (version 3.5.2). Les variables qualitatives étaient décrites par les effectifs et pourcentages ; les variables quantitatives par la moyenne et l'écart type lorsqu'elles étaient normales, ou par la médiane et l'interquartile (i.e. 25^e et 75^e percentiles) lorsqu'elles étaient de distribution non normale.

La concordance des mesures entre les deux tubes était évaluée par un coefficient de concordance de Kendall pour chaque variable COV après conversion en variable ordinale (divisée en 20 niveaux). Si la concordance était satisfaisante, les concentrations en COVs dans les 2 tubes étaient moyennées pour chaque sujet avant comparaison entre cas et témoins. Les COVs étaient comparées par un test t de Student en cas de distribution normale, ou par un test de Wilcoxon en cas d'écart à la normalité. Tous les tests statistiques étaient bilatéraux avec un risque de première espèce de 5 %.

Deux approches ont été évaluées pour modéliser les résultats d'un nez électronique. La première consistait à sommer les valeurs des COVs identifiés comme plus élevés chez les cas et à déterminer un seuil de positivité pour la valeur de cette somme avec estimation de la sensibilité et la spécificité au moyen d'une courbe *Receiver Operating Curve* (ROC) dont la pertinence était évaluée par l'aire sous la courbe (*Area Under Curve*, AUC). Cette approche modélise un nez électronique calibré pour détecter sans distinction tout COV associé au CBP et dont la positivité serait liée au dépassement d'un seuil de concentration total de ces COVs.

La deuxième méthode envisagée, plus fine, consistait en l'utilisation d'une régression logistique pour déterminer la participation individuelle de chaque COV dans l'association au CBP au moyen des coefficients de régression, avec estimation d'un score prédictif fonction des valeurs de chaque COV, et dont les performances prédictives auraient également été évaluées au moyen d'une courbe ROC, avec estimation de la sensibilité et spécificité au point seuil ainsi que de l'AUC. Cette approche modélise un nez électronique calibré pour détecter individuellement les COVs associés au CBP préalablement identifiés, avec une meilleure prise en compte de l'hétérogénéité des concentrations de COVs selon les patients.

2.8. Cadre réglementaire

Cette étude a porté sur un échantillon de données collectées dans le cadre de l'essai CATOCOV. Une déclaration spécifique pour ce travail a été faite auprès de la commission nationale de l'informatique et des libertés (CNIL) sous la référence JAN20-466. Par ailleurs, l'étude CATOCOV a reçu l'avis favorable du Comité de Protection des Personnes SUD-EST IV le 13/06/2017. Les cas et les témoins devaient fournir leur consentement écrit pour être inclus.

À l'inclusion, les sujets recevaient une information éclairée orale et écrite (via une note d'information) et signaient un formulaire de consentement en 3 exemplaires dont un leur était remis. Les données étaient anonymisées et chaque sujet se voyait attribuer un numéro d'inclusion.

3. RÉSULTATS

Le recrutement de nos sujets a été freiné par plusieurs problématiques :

- Un retard de livraison du matériel de recueil provenant du Royaume-Uni (OWLSTONE Ltd.)
- Un délai à l'ouverture du centre à l'Institut Pasteur de Lille en raison de problèmes organisationnels et de recrutement de l'attaché de recherche clinique sur place
- Plus récemment, la pandémie de COVID-19 qui a allongé et retardé les délais d'analyses au laboratoire TERA Environnement.

3.1. Description de la population à l'inclusion

Sur la période d'inclusion du 29/04/2019 au 31/01/2020, 94 patients atteints de cancers broncho-pulmonaires ont été recensés pour une éventuelle participation à l'étude, parmi lesquels 68 ont été exclus (causes détaillées dans la Figure 3 ci-dessous), 5 ont refusé de participer et 7 n'ont pas pu être inclus en raison de problèmes organisationnels. Concernant les témoins, 223 sujets ont été identifiés et 160 ont été exclus. Parmi les restants, 13 ont refusé de participer et 4 n'ont pas pu être inclus en raison de problèmes logistiques. Quarante-six témoins ont finalement été inclus mais toutes les données n'étaient pas disponibles au moment de l'analyse. Au total, 14 cas et 15 témoins ont été analysés (Figure 3).

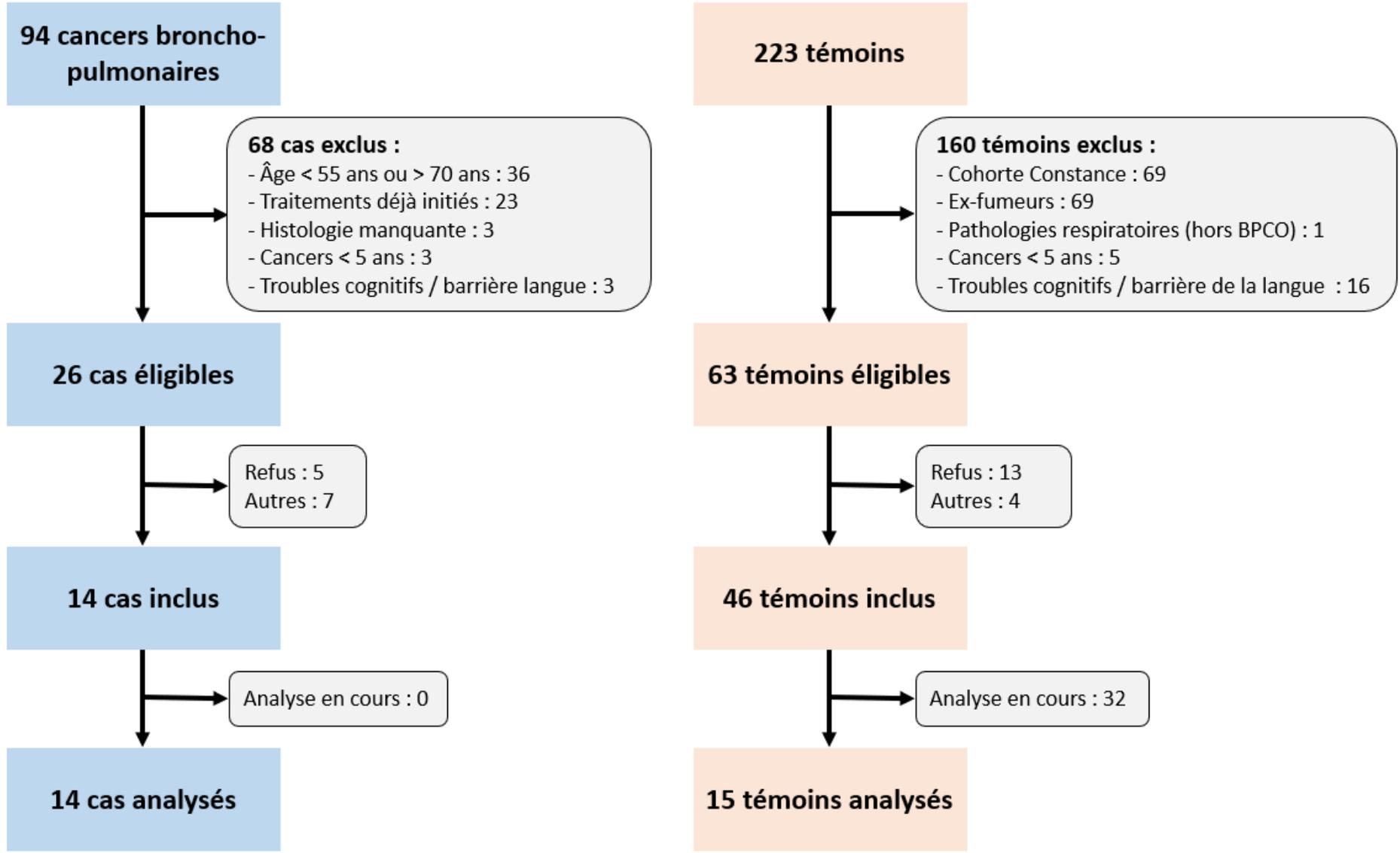


Figure 3. Diagramme de flux.

3.1.1. Caractéristiques globales

L'ensemble des caractéristiques socio-démographiques et cliniques de la population de l'étude sont résumées dans le Tableau 1.

Caractéristiques de la population		Cas (n = 14)	Témoins (n = 15)
Sexe	Homme n (%)	10 (71)	5 (33)
	Femme n (%)	4 (29)	10 (67)
Âge	(années)	61 (57 - 63)	62 (57 - 65)
Tabagisme	n (%)	14 (100)	3 (20)
Consommation cumulée	PA	42 (32 - 49)	31 (16 - 42)
	Sevré n (%)	6 (43)	0 (0)
Cigarette électronique	n (%)	1 (7)	0 (0)
Alcool	n (%)	5 (36)	3 (20)
BPCO	n (%)	3 (21)	0 (0)
Asthme	n (%)	1 (7)	0 (0)
Hypertension artérielle	n (%)	7 (50)	10 (67)
Diabète	n (%)	2 (14)	0 (0)
Dysthyroïdie	n (%)	2 (14)	2 (13)
IMC	(kg/m ²)	23 (19,5 - 29)	26,5 (23 - 32)

Tableau 1. Caractéristiques socio-démographiques et cliniques de la population à l'inclusion. (Les données sont présentées en effectifs et pourcentages ou en médianes et distances interquartiles. PA = paquets-années. BPCO = broncho-pneumopathie chronique obstructive. IMC = indice de masse corporelle.)

Il y avait une majorité d'hommes dans le groupe cas (71 %) et de femmes dans le groupe témoin (67 %). Les âges étaient comparables entre les deux groupes avec une médiane respective de 61 et 62 ans. La totalité des cas étaient des fumeurs parmi lesquels 42 % étaient sevrés alors qu'il y avait 20 % de fumeurs chez les témoins.

3.1.2. Caractéristiques des cancers broncho-pulmonaires

Cancers broncho-pulmonaires (n = 14)		n (%)
OMS	0 - 1	13 (93)
	≥ 2	1 (7)
Circonstances diagnostiques	Symptômes respiratoires	8 (57)
	Symptômes extra-respiratoires	6 (43)
Localisation	Pulmonaire unique	3 (21)
	Pulmonaire multiple	11 (79)
Sites de biopsies	Respiratoires	11 (79)
	Extra-respiratoires	3 (21)
Type histologique	Adénocarcinome	9 (64)
	Carcinome épidermoïde	1 (7)
	Carcinome à petites cellules	0 (0)
	Autres	4 (29)
Contact avec la lumière des voies aériennes		9 (64)
Nécrose intra-tumorale		4 (29)
Stade	I (localisé)	1 (7)
	II (localisé)	0 (0)
	III (localement avancé)	3 (21)
	IV (disséminé)	10 (72)
Ganglions N	Absence d'adénopathie	1 (7)
	Présence d'adénopathies	13 (93)
Métastases M	Absence de métastase	4 (29)
	Métastase unique	3 (21)
	Métastases multiples	7 (50)
Statut PDL1 (n = 12)	PDL1 < 1 %	2 (17)
	PDL1 1 à 49 %	3 (25)
	PDL1 > 50 %	7 (58)
Altérations moléculaires (n = 12)	Aucune	3 (25)
	Mutation EGFR	0 (0)
	Mutation TP53	3 (25)
	Mutation KRAS	4 (33)
	Surexpression MET	2 (17)

Tableau 2. Caractéristiques des patients porteurs d'un CBP et de leur tumeur. (Les données sont présentées en effectifs et pourcentages.)

La grande majorité des cas inclus (93 %) avaient un état général conservé avec un indice de performance status OMS < 2. Tous les patients sauf un (soit 93 %) présentaient un cancer pulmonaire de stade localement avancé (III) ou disséminé (IV). On recensait une majorité d'adénocarcinome (64 %) et un carcinome épidermoïde (7 %). Les 4 restants étaient : un carcinome neuroendocrine à grandes cellules, un CBNPC non épidermoïde et deux CBNPC-NOS (*Not Otherwise Specified*). Des altérations moléculaires étaient retrouvées chez trois-quarts des patients et plus de la moitié (58 %) avaient un statut tumoral PDL1 > 50 % (Tableau 2).

3.2. Description des conditions de recueil

Conditions de recueil		Global (n = 29)	Cas (n = 14)	Témoins (n = 15)
Durée du recueil	s	421 (242 - 445)	242 (227 - 256)	441 (427 - 470)
Volume	500 mL	11 (38)	11 (79)	0 (0)
	1000 mL	18 (62)	3 (21)	15 (100)
Prise d'aliments dans les 3 heures	n (%)	24 (83)	10 (71)	14 (93)
Prise de boissons dans les 3 heures	n (%)	25 (86)	12 (86)	13 (86)
Tabac dans les 3 heures	n (%)	4 (14)	2 (14)	2 (13)
Soins dentaires en cours	n (%)	5 (17)	0 (0)	5 (33)
Brossage de dents dans les 6 heures	n (%)	15 (52)	8 (57)	7 (47)

Tableau 3. Conditions de recueil. (Les données sont présentées en effectifs et pourcentages ou en médianes et distances interquartiles.)

Pour 79 % des cas, le volume recueilli d'air exhalé était fixé à 500 mL avec une durée médiane de recueil de 240 secondes (DIQ [224 - 243]) soit 4 minutes. Ce volume a été secondairement augmenté à 1000 mL afin d'allonger le temps de recueil. Ceci a concerné l'ensemble des témoins et 21 % des cas avec une durée médiane de recueil de 440 secondes (DIQ [419 - 449]) soit 7 minutes et 20 secondes. La majorité des sujets cas et témoins avaient consommé des aliments ou des boissons dans les 3 heures précédant le recueil. Dans chaque groupe, 2 individus avaient fumé dans les 3 heures précédant le recueil (Tableau 3).

3.3. Analyse des COVs

Sur la population globale, on identifiait en médiane 34 COVs (DIQ [29 - 41]) par analyse pour une limite de quantification à 1 ng/support. Pour les sujets ayant eu un volume de recueil de 500 mL, 28 COVs étaient retrouvés (DIQ [22 - 30]), pour ceux avec un volume de 1000 mL, on en identifiait 38 (DIQ [34 - 44]). Compte-tenu de l'importante proportion de cas ayant eu un volume de recueil à 500 mL, ces chiffres étaient superposables à ceux des deux groupes : 29 (DIQ [23 - 34]) chez les cas et 39 (DIQ [24 - 45]) chez les témoins.

Devant le faible effectif de notre population, nous avons choisi de ne pas analyser l'ensemble des COVs identifiés mais d'en sélectionner une partie pour analyse selon deux approches. La première s'est basée sur les données de la littérature. Ainsi, nous avons sélectionné, parmi les COVs retrouvés dans nos analyses, ceux qui étaient déjà identifiés comme potentiellement associés au CBP. La deuxième a consisté en une sélection des COVs qui étaient quantifiés en grande concentration chez les patients (cas). Nous avons donc conservé 21 COVs pour analyse, parmi lesquels 14 étaient déjà décrits dans des revues de littérature et 7 étaient détectés en concentration plus importante chez les cas.

3.3.1. *Concordance des mesures*

Les résultats de la concordance des mesures sont résumés dans le tableau 4. Les mesures étaient concordantes entre les deux tubes avec des coefficients de Kendall $> 0,75$ et significatives ($p < 0,05$) pour la majorité des COVs analysés. Seul l'éthylbenzène avait des mesures discordantes avec un coefficient de Kendall à 0,65 et $p > 0,05$.

Composés organiques volatils	Numéro CAS	Coefficient de Kendall	p
1,3-Pentadiène	504-60-9	0,89	0,007
2-Propanol	67-63-0	0,94	0,003
2,4-Diméthylheptane	2213-23-2	0,98	0,002
Acétone	67-64-1	0,88	0,008
Benzène	71-43-2	0,96	0,002
Butane	106-97-8	0,99	0,001
Cyclotétrasiloxane, octaméthyl-	556-67-2	0,85	0,012
Cyclotrisiloxane, hexaméthyl-	541-05-9	0,85	0,012
Décane	124-18-5	0,95	0,003
Disiloxane, 1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)-	55967-53-8	0,95	0,003
Disiloxane, hexaméthyl- (HMDSO)	107-46-0	0,98	0,002
Éthanol	64-17-5	0,93	0,004
Éthylbenzène	100-41-4	0,65	0,13
Heptanal	111-71-7	1	0,001
Heptane	142-82-5	0,89	0,006
Pentanal	110-62-3	0,75	0,04
Styrène	100-42-5	0,99	0,001
Toluène	108-88-3	0,97	0,002
Triméthylsilanol	1066-40-6	0,98	0,002
Trisiloxane, octaméthyl-	107-51-7	0,98	0,002
Undécane	1120-21-4	0,9	0,005

Tableau 4. Concordance entre les deux mesures de COVs.

3.3.2. Description et comparaison des COVs

Parmi les COVs d'intérêt pour le dépistage du CBP identifiés dans la littérature, on en mettait en évidence 14 dans notre population. On retrouvait une différence de concentration significative entre les témoins et les cas pour 5 d'entre eux : le 2-propanol ($p < 0,001$), le 2,4-diméthylheptane ($p < 0,001$), le benzène ($p < 0,001$), l'éthanol ($p < 0,001$) et le toluène ($p = 0,001$), avec des concentrations plus élevées chez les cas. Les résultats étaient également significatifs ($p < 0,05$) pour le butane, l'heptane et le styrène mais la différence de concentration était très faible et donc

difficile à analyser. L'acétone était retrouvé en grande concentration dans les deux groupes sans différence significative. Les concentrations en décane, éthylbenzène, heptanal, pentanal et undécane étaient minimales dans les deux groupes sans différence statistiquement significative (Tableau 5).

COVs identifiés dans la littérature	Concentrations chez les cas (ng/support)	Concentrations chez les témoins (ng/support)	p
2-Propanol	13,07 (5,51 - 58,41)	0 (0 - 0)	<0,001
2,4-Diméthylheptane	4,15 (1,93 - 5,67)	0 (0 - 0)	<0,001
Acétone	176,1 (109,8 - 287,7)	162,8 (109,7 - 305,7)	0,88
Benzène	17,23 (13,16 - 21,75)	4,2 (3,73 - 5,98)	<0,001
Butane	0 (0 - 0)	0 (0 - 0,42)	0,01
Décane	0 (0 - 3,47)	0 (0 - 1,37)	0,76
Éthanol	21,05 (12,35 - 27,51)	4,65 (3,07 - 6,23)	<0,001
Éthylbenzène	0 (0 - 0)	0 (0 - 0)	1
Heptanal	0 (0 - 0)	0 (0 - 0,45)	0,09
Heptane	0 (0 - 0)	0 (0 - 0,55)	0,04
Pentanal	0 (0 - 0)	0 (0 - 0)	0,1
Styrène	0 (0 - 0)	0 (0 - 0,5)	0,005
Toluène	2,43 (0,11 - 3,66)	0 (0 - 0)	0,001
Undécane	0 (0 - 0)	0 (0 - 0)	0,88

Tableau 5. Comparaison des concentrations en ng/support des COVs d'intérêt identifiés dans la littérature chez les cas et chez les témoins en concentration moyenne. (Les données sont exprimées en médianes et distances interquartiles.)

Sept composés, non identifiés dans la littérature comme biomarqueurs d'intérêt pour le dépistage du CBP, étaient quantifiés à des concentrations importantes chez les cas. On retrouvait une différence de concentration significative ($p < 0,05$) pour 6 d'entre eux : le cyclotétrasiloxane, octaméthyl- ou D4, le cyclotrisiloxane, hexaméthyl- ou D3, le disiloxane, 1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)-, le disiloxane, hexaméthyl- ou L2, le triméthylsilanol ou SiOH et le trisiloxane, octaméthyl- ou L3. Enfin, le 1,3-pentadiène était mesuré à des concentrations élevées et superposables dans les deux groupes sans différence significative (Tableau 6).

COVs non identifiés dans la littérature	Concentrations chez les cas (ng/support)	Concentrations chez les témoins (ng/support)	p
1,3-Pentadiène	92,38 (68,22 - 111,29)	73,8 (59,27 - 133,90)	0,68
Cyclotétrasiloxane,	21 (11,94 - 22,54)	10,1 (7,56 - 11,7)	0,005
Cyclotrisiloxane,hexaméthyl-	61,3 (53,99 - 76,71)	30,65 (19,25 - 42,58)	<0,001
Disiloxane,1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)-	6,87 (2,8 - 10,27)	1,25 (0,85 - 2)	0,008
Disiloxane,hexaméthyl-	939,6 (658,2 - 1218,4)	44 (28,57 - 65,97)	<0,001
Triméthylsilanol	118 (90,12 - 180,39)	7,7 (5,93 - 9,78)	<0,001
Trisiloxane,octaméthyl-	97,17 (70,76 - 118,21)	3 (2,25 - 4,33)	<0,001

Tableau 6. Comparaison des concentrations en ng/support des COVs non identifiés dans la littérature chez les cas et chez les témoins en concentration moyenne. (Les données sont exprimées en médianes et distances interquartiles.)

3.4. Modélisation d'un nez électronique

3.4.1. *Approche par sommation des COVs associés au CBP*

L'ensemble des COVs présents en concentrations plus importantes chez les patients (cas) a été sommé et comparé dans les deux populations avec : dans le groupe cancer un total à 1350 ng/support (DIQ [967,5 - 1587,9]) et dans le groupe témoin un total à 125,95 ng/support (DIQ [94,15 - 140,05]), $p < 0,001$. Les performances diagnostiques de cette valeur sommée sont présentées en Figure 4. Au seuil de 485,1 ng/support, on retrouve une sensibilité de 1 (IC95% [1 - 1]), une spécificité de 0,93 (IC95% [0,8 - 1]) et une AUC à 0,97 (IC95% [0,9 - 1]).

3.4.2. *Approche par régression logistique et score prédictif*

Malheureusement, la faiblesse d'effectif résultant d'une part en une non-normalité des variables d'intérêt et d'autre part en un manque de puissance important, n'a pas permis la réalisation d'une approche par régression logistique et l'établissement d'un score sur notre population.

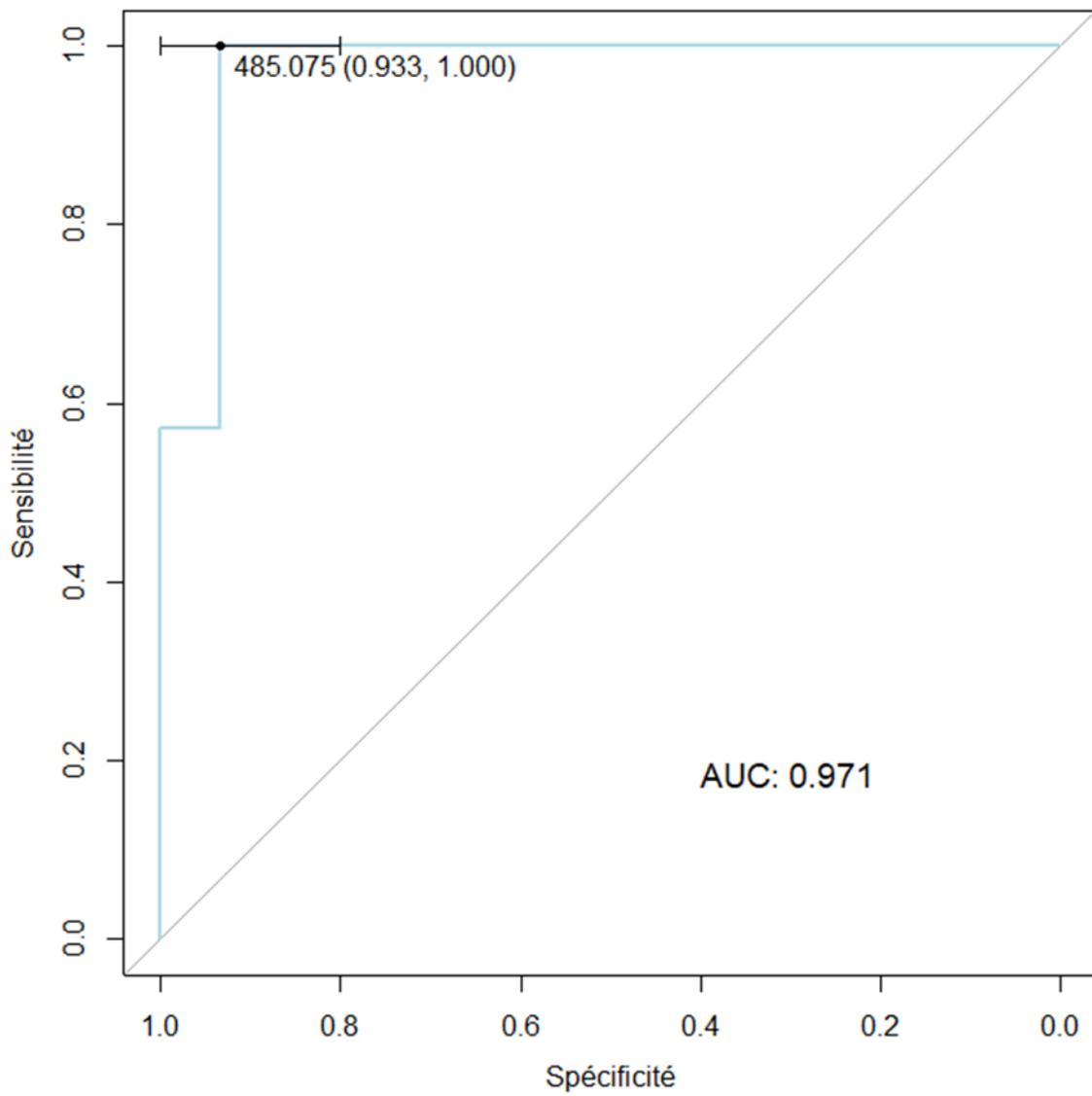


Figure 4. Courbe ROC de la modélisation d'un nez électronique par l'addition des COVs plus élevés dans le groupe cancer.

4. DISCUSSION

Dans cette étude visant à identifier et comparer les COVs dans une population de patients atteints d'un CBP et de témoins sains, nous avons mis en évidence 11 COVs d'intérêt associés à la présence d'un CBP. Parmi ces COVs, 5 ont déjà été identifiés dans la littérature : 2-propanol, 2,4-diméthylheptane, benzène, éthanol et toluène. En revanche, il était également retrouvé une association significative au CBP pour 6 nouveaux COVs : cyclotétrasiloxane, octaméthyl- (D4), cyclotrisiloxane, hexaméthyl- (D3), disiloxane, 1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)-, disiloxane, hexaméthyl- (L2), triméthylsilanol (SiOH) et trisiloxane, octaméthyl- (L3).

Notre étude s'est appuyée sur la littérature pour sélectionner les COVs à analyser. Une des difficultés majeures dans l'interprétation des études portant sur les COVs est la terminologie utilisée : parfois présentés avec leur numéro CAS, ils ont aussi de nombreuses appellations (31).

Contrairement aux données des trois revues de littérature récentes, nous n'avons retrouvé dans nos analyses que 14 COVs sur les 25 **COVs décrits comme potentiels biomarqueurs du CBP** (27,36,37). Parmi ceux-là, seuls 5 étaient quantifiés de manière plus importante chez les cas que chez les témoins : le 2-propanol, le 2,4-diméthylheptane, le benzène, l'éthanol et le toluène. Aussi, 3 COVs (butane, heptane, styrène) n'étaient pas identifiés dans la population cas et étaient mesurés à des concentrations minimales chez les témoins, ce qui explique la différence significative. Ces résultats sont contradictoires avec d'autres publications dans lesquelles le butane, l'heptane et le styrène étaient présents en plus forte concentration chez les sujets atteints de CBP (27,36,37). L'éthylbenzène, également identifié comme potentiel

biomarqueur du CBP dans plusieurs publications (36,37), n'était pas augmenté de manière significative chez les patients atteints de CBP dans notre étude. Cependant, ce résultat doit être analysé avec prudence car il s'agissait du seul COV pour lequel nous n'avions pas une bonne reproductibilité de mesure entre les 2 tubes, suggérant un problème technique lié au recueil ou à l'analyse. L'acétone, qui semble aussi être un biomarqueur du CBP (27,36,37), était quantifié dans notre étude à des concentrations importantes à la fois chez les témoins et chez les cas sans différence significative. Il a été montré que les concentrations en acétone peuvent être influencées par ses propriétés physico-chimiques et par le mode ventilatoire du patient (36,44) ; la méthode de recueil a donc pu influencer ces résultats. Toutes ces discordances par rapport à ce qui a été précédemment décrit peuvent également s'expliquer par le faible effectif de notre population, entraînant un manque de puissance et de probables fluctuations d'échantillonnage importantes, ainsi que par l'utilisation des méthodes de recueil hétérogènes entre les deux groupes. En effet, 79 % des cas avaient eu un volume de recueil de 500 mL et 100 % des témoins un volume de recueil de 1000 mL.

Six **COVs non décrits précédemment dans la littérature** sont associés dans notre étude de manière significative au CBP. Ils appartiennent tous à la classe des siloxanes (SiOH, L2, L3, D3, D4 et disiloxane,1-éthényl-1,1,3,3-tetraméthyl-3-(2-propényl)-). À notre connaissance, les siloxanes n'ont pas été étudiés ni identifiés comme potentiels biomarqueurs du CBP. Notre étude est la première à mettre en évidence une association des siloxanes au CBP. Les siloxanes, composés de silicium, servent à produire notamment du silicone. Ils sont largement utilisés dans l'industrie : fabrication de produits du quotidien (cosmétique, hygiène, alimentation), automobile,

bâtiment par exemple mais aussi dispositifs médicaux (implants mammaires, lubrifiants, pansements, etc.) et se retrouvent dans l'environnement (45). À ce jour considérés comme non dangereux pour la santé, plusieurs études suggèrent néanmoins un effet néfaste de ces composés sur l'organisme avec entre autres des effets inflammatoires au niveau respiratoire (46) et carcinogène au niveau utérin (45) en modèle animal. L'association retrouvée ici pourrait faire discuter leur rôle en tant que biomarqueurs du CBP sous réserve d'une non-contamination par des COVs environnementaux dans le groupe cas. En effet, les prélèvements des cas et des témoins ont été effectués dans des centres et donc des salles différentes. Cependant, nous ne disposons pas de prélèvements d'ambiance ni d'un échantillon suffisamment grand et comparable pour en tirer des conclusions.

La technique d'identification des COVs par GC-MS utilisée dans cette étude est une étape nécessaire pour mettre en évidence les COVs d'intérêt pour le dépistage du CBP mais elle reste coûteuse et longue (26,30,32). Aussi, l'étape suivante est celle de la création d'un nez électronique utilisant une combinaison de COVs définie par cette première phase d'identification (30). Malgré notre faible échantillon, nous avons pu évaluer une approche de **modélisation de nez électronique** par sommation des COVs plus concentrés de manière significative dans le groupe CBP. Cette approche simple semblait présenter des performances diagnostiques intéressantes (AUC 0,97 (IC95% [0,9 - 1]) et sensibilité 1 (IC95% [1 - 1]) - spécificité 0,93 (IC95% [0,8 - 1]) au seuil de 485,1 ng/support) pour dépister un CBP. Néanmoins, les performances diagnostiques du test (sensibilité et spécificité) sont ici très vraisemblablement surévaluées par plusieurs facteurs. Premièrement, le schéma d'étude choisi ici : une étude cas-témoins diagnostique (souvent appelée « two-gates design ») surestime la sensibilité et la spécificité (47). En effet, les résultats sont interprétés non

pas sur la population cible à risque mais sur une double population dont le statut malade ou sain est connu avant la réalisation du test à l'étude. Ensuite, la majorité des cas inclus dans notre étude s'avèrent être des formes avancées de la maladie (93 % de stade III et IV) ce qui peut conduire à un biais de spectre et surestimer la sensibilité en diminuant le nombre de faux négatifs (47). De même, notre groupe contrôle est constitué de volontaires sains probablement en meilleure santé que la population générale (11,47,48), majoritairement non tabagiques (80 %) et sans comorbidités respiratoires. Ceci participe à la non-comparabilité de nos deux groupes et surestime la spécificité en réduisant le nombre de faux positifs (47). De plus, les performances diagnostiques sont ici testées et validées sur la même population qui a servi à développer le modèle et à définir le seuil, ce qui correspond à une validation apparente. Cette validation apparente est artificiellement optimiste et entraîne une surestimation de la sensibilité et de la spécificité (49). Enfin, cette modélisation par sommation ne permet pas d'évaluer l'impact des variables à faibles concentrations qui peuvent pourtant être discriminantes. L'approche idéale nécessiterait l'élaboration d'un score mais cela n'a pas été possible compte tenu du faible effectif.

Les techniques de recueil et d'analyse de COVs dans l'air exhalé ne sont pas standardisées et sont influencées par de nombreux facteurs.

Premièrement, la qualité de recueil dépend des conditions d'échantillonnage (33,36). Le dispositif ReCIVA® utilisé dans cette étude (34) permet de s'affranchir de nombreuses difficultés liées au recueil. Plusieurs paramètres sont ainsi modifiables, tels que le volume, la durée de recueil ainsi que la phase expiratoire à recueillir (fraction alvéolaire, fraction intermédiaire, fraction dans l'espace mort) en monitorant la pression expirée en gaz carbonique (pCO₂). Nous avons fait le choix de prélever la

fraction intermédiaire afin d'obtenir un maximum de COVs à la fois des voies aériennes et des alvéoles. Le mode ventilatoire du patient (hyperventilation, apnée) peut également influencer les niveaux de concentrations en COVs comme expliqué avec l'acétone ci-dessus (36,44). C'est pourquoi le masque ReCIVA® permet de monitorer la fréquence respiratoire et la pCO₂ pendant le recueil pour contrôler le remplissage des tubes en régulant l'ouverture des valves des tubes et ainsi obtenir des résultats fiables et reproductibles (34). Ceci a été vérifié dans notre étude par une bonne concordance des concentrations de COVs entre chaque tube excepté pour l'éthylbenzène.

Deuxièmement, afin d'optimiser les concentrations en COVs et d'en identifier un maximum même à des concentrations minimales, nous avons choisi de recueillir l'air exhalé dans des tubes avec adsorbants pour préconcentrer les COVs (33). Comme décrit dans nos résultats, le volume et la durée de recueil influencent la concentration en COVs (33). Aussi, le volume de recueil a été secondairement augmenté à 1000 mL pour allonger la durée du prélèvement permettant ainsi d'augmenter le nombre de COVs identifiés (28 COVs, DIQ [22 - 30] à 500 mL versus 38 COVs, DIQ [34 - 44] à 1000 mL). Ces chiffres contrastent avec ceux de l'étude pionnière de Gordon et al. (35) qui, en 1985, identifiaient environ 150 pics par patient mais avec une technique de recueil très différente et pour des volumes de recueil estimés à environ 20 litres. La limite de quantification liée à la technique d'analyse (ici GC-MS) et fixée par le laboratoire (ici 1 ng/support) peut aussi impacter la détection des COVs dont les concentrations sont de l'ordre du nmol/l au pmol/l (27,33,38,44). De plus, la GC-MS reste une technique contraignante car non mobile et elle nécessite de nombreuses calibrations (26,27,30,32).

Ensuite, des contaminations extérieures peuvent survenir à chaque étape (30,33,36). Afin de limiter au maximum celles liées au matériel, le masque et les tubes ont été manipulés avec des gants et l'entretien effectué avec les mêmes produits dont les COVs sont connus et peuvent être retranchés des analyses à posteriori (lingettes Anios®). Une salle de prélèvement a été dédiée dans chaque centre. Des prélèvements d'ambiance n'ont pas pu être réalisés au moment de l'étude mais un purificateur d'air à base de charbon actif a été installé en amont du dispositif de recueil. Le charbon activé permet en effet d'adsorber un maximum de COVs ambiants (43). Les tubes ont ensuite été stockés dans un récipient hermétique au réfrigérateur pour limiter la volatilisation de COVs à température ambiante. Les contaminations liées au transport n'ont pas pu être maîtrisées mais les récipients étaient considérés comme clos. Enfin, la désorption a été faite dans une salle d'analyse dédiée dite « blanche » au laboratoire et en respectant le sens d'échantillonnage des tubes afin de faciliter la récupération des COVs selon leur poids moléculaire.

Enfin, de nombreux facteurs de confusion sont à prendre en compte pour l'interprétation des résultats. Malheureusement, l'effectif de notre population n'était pas assez important pour évaluer leur impact sur nos résultats et certaines données étaient manquantes. Dans notre étude, la différence de sexe (71 % d'hommes dans le groupe cas et 67 % de femmes dans le groupe témoin) a pu influencer les résultats (36) de même que le statut fumeur (33,36). Comme montrés par Patassini et al. (50) sur un groupe de 73 sujets sains avec 17 fumeurs et 56 non-fumeurs, le benzène et le toluène sont associés au tabac. Nous avons mis en évidence une association avec le CBP statistiquement significative pour ces 2 COVs ($p < 0,05$) mais il y avait 100 % de fumeurs (actifs ou sevrés) chez les cas et 20 % chez les témoins. Ensuite, la présence de comorbidités respiratoires (38) et l'atteinte de la barrière alvéolo-capillaire (33)

peuvent interférer sur les concentrations en COVs. On comptait 4 sujets (28 %) atteints de pathologies respiratoires (3 BPCO et 1 asthme) chez les cas contre 0 chez les témoins. Pour rappel, les témoins atteints de pathologies respiratoires hors BPCO étaient exclus de l'analyse, contrairement aux cas, ce qui amplifie ce facteur de confusion. Nous n'avons pas pu évaluer l'effet du stade du CBP sur les COVs retrouvés. En effet, la concentration des COVs peut varier selon le stade de la maladie (36). Néanmoins, la prédominance de formes avancées (93 % de stade III et IV) rend l'extrapolation de nos résultats difficiles aux formes précoces qui sont finalement les cibles du dépistage. Concernant les autres facteurs de confusion potentiels comme les pathologies extra-respiratoires (36), l'IMC (51) et la consommation de nourriture avant le recueil (33,36), il ne semblait pas y avoir de différences importantes entre nos deux groupes. Aussi, nous n'avons pas pu décrire les expositions environnementales et professionnelles (36,52) par manque de données.

Malgré les difficultés évoquées rendant difficile l'**extrapolation de nos résultats à la population cible en pratique courante**, certaines données sont encourageantes quant à la mise en place d'un éventuel dépistage du CBP par mesure de COVs dans l'air exhalé. Le taux d'adhésion à l'étude était excellent, d'environ 80 % dans les deux groupes (5 refus sur 26 dans le groupe cas soit 19 % et 13 refus sur 63 dans le groupe témoin soit 20 %). Néanmoins, les témoins tabagiques semblaient moins enclins à participer à l'étude bien que les motifs de refus n'aient pas été clairement demandés. Ces données restent cohérentes avec l'étude de Pinsky et al. (48) qui ont montré que les sujets sains recrutés sur la base du volontariat pour participer à un dépistage étaient en meilleure santé et moins souvent fumeurs que la population générale.

L'estimation du coût par patient pour l'identification des COVs est ici d'environ 280 euros (hors prix du masque ReCIVA® réutilisable de 28 750 euros) alors que le coût unitaire d'un scanner thoracique est estimé à environ 91 euros (entre 61,91 euros et 131,79 euros selon le forfait technique) (23). Ce tarif de 280 euros inclut les consommables à usage unique du masque (20 euros), le transport et l'analyse par GC-MS (87 euros par tube soit 174 euros), le temps pour l'inclusion du patient et le recueil par l'attaché de recherche clinique (42 euros pour environ 1 heure) ainsi que le temps médical pour l'analyse des résultats (42 euros pour environ 30 minutes). Cependant, cette technique n'est ici utilisée que dans le but d'identifier des COVs d'intérêt. Elle n'a pas vocation à être utilisée en pratique courante mais simplement à calibrer un éventuel dispositif de type nez électronique.

Enfin, le choix de la population cible à ce dépistage reste discutable notamment sur les critères d'âge même si ceux établis dans notre étude sont superposables à ceux des grandes études de dépistage (14,15). On comptait en effet 36 patients sur 99 atteints de CBP (36 %) exclus du fait de leur âge : 8 en raison d'un âge inférieur à 55 ans et 28 d'un âge supérieur à 70 ans. L'âge ne devrait pas être le seul critère de sélection et l'état général devrait aussi être pris en compte en vue d'une éventuelle chirurgie (13). Aussi, la fourchette d'âge des patients éligibles à un dépistage devrait être étendue comme c'est d'ailleurs proposé aux États-Unis avec la généralisation du dépistage du CBP par scanner thoracique chez les patients âgés de 55 à 80 ans (20).

La **force de cette étude** réside dans l'utilisation d'un dispositif de recueil performant permettant d'identifier un maximum de COVs avec une bonne reproductibilité. Grâce à cette technique fiable, notre objectif est pertinent pour pouvoir ensuite modéliser un dispositif de type nez électronique qui soit plus simple, facile d'accès et moins coûteux. Aussi, l'analyse était réalisée par un laboratoire centralisé indépendant.

La **principale limite** repose sur la faible population de notre étude conduisant à un manque de puissance. On constate également l'absence de comparabilité entre nos groupes avec d'importantes fluctuations d'échantillonnage tant sur le plan démographique (sexe) que clinique (statut tabagique, cancers évolués). Ceci pose la question de la représentativité de notre faible échantillon par rapport à la population cible du dépistage du CBP et limite l'extrapolation de nos résultats qui ne sont d'ailleurs pas totalement concordants avec la littérature.

Une autre limite concerne la population témoin considérée comme saine par défaut sans examen clinique spécifique ni examen d'imagerie, avec un risque réel de sujet sain en réalité malade. Ceci correspond au biais de vérification qui impacte nos résultats et surestime les performances diagnostiques (53). La sensibilité et la spécificité de notre modèle de nez électronique sont également surévaluées par le schéma de l'étude en « two-gates design » et la validation apparente (47,49).

On rapporte également de nombreux biais de confusion du fait des facteurs influençant la mesure et l'interprétation des COVs qui n'ont pas pu être pris en compte.

Enfin, la multiplicité des tests statistiques par rapport à notre faible effectif doit être soulignée car elle impose d'interpréter nos résultats avec prudence.

Perspectives

Nos résultats ont pour but d'être complétés et renforcés par ceux d'une étude de plus grande ampleur : l'étude CATOCOV (40). Cet essai, dont les inclusions sont en cours, a pour objectif d'inclure 1500 cas et 500 témoins. Avec cet effectif, les résultats permettront d'obtenir une réelle puissance statistique et également de réaliser des ajustements sur les différents facteurs de confusion retrouvés. Les données permettront ensuite la modélisation d'un nez électronique avec projet de validation interne par bootstrapping pour corriger l'optimisme (49) puis l'élaboration d'un score prédictif.

L'utilisation de biomarqueurs pour le dépistage du CBP et, par exemple, l'implémentation du nez électronique seraient faisables à grande échelle comme démontré dans l'étude de Larose et al. sur la population américaine (54). Un biomarqueur permettrait en effet de sélectionner la population éligible à un dépistage du CBP par scanner thoracique avec une rentabilité et une réduction des coûts de santé qui dépendraient des performances diagnostiques du test c'est-à-dire du nombre de cancers détectés précocement et donc du nombre d'années de vie gagnées.

Aussi, les COVs semblent être des biomarqueurs intéressants avec un champ d'application bien plus large que le CBP : pathologies respiratoires (asthme, BPCO, pathologies interstitielles diffuses par exemple), pathologies extra-respiratoires et pas uniquement dans le cancer (38,44). Ils sont d'ailleurs actuellement évalués dans le diagnostic du COVID-19 (55).

5. CONCLUSION

Notre étude a permis d'identifier dans une population constituée de patients atteints d'un CBP et de témoins sains 11 COVs d'intérêt associés au CBP, parmi lesquels 5 sont concordants avec les COVs précédemment décrits dans la littérature. Nous avons également mis en évidence une association avec le CBP de 6 nouveaux COVs appartenant à la famille des siloxanes jusqu'alors non décrits. Les techniques de recueil et d'analyse utilisées ont permis d'obtenir des résultats fiables et reproductibles. Aussi, nos résultats suggèrent, sous réserve des limites de l'étude, que la modélisation d'un nez électronique après identification des COVs d'intérêt est faisable avec de bonnes performances diagnostiques.

Ces résultats doivent néanmoins être confirmés sur des effectifs plus larges : c'est l'objectif de l'étude CATOCOV en cours de recrutement.

6. BIBLIOGRAPHIE

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
2. Defossez G, Le Guyader-Peyrou S, Uhry Z, Grosclaude P, Dantony E. Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Étude à partir des registres des cancers du réseau Francim. juill 2019 [cité 29 mars 2020]; Disponible sur: <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers/cancer-du-sein/documents/rapport-synthese/estimations-nationales-de-l-incidence-et-de-la-mortalite-par-cancer-en-france-metropolitaine-entre-1990-et-2018-volume-1-tumeurs-solides-etud>
3. Colonna M. Epidémiologie du cancer du poumon en France : incidence, mortalité et survie (tendance et situation actuelle). *Rev Mal Respir Actual.* sept 2016;8(5):308-18.
4. IARC. Les cancers attribuables au mode de vie et à l'environnement en France métropolitaine. Lyon Int Agency Res Cancer [Internet]. 2018 [cité 15 avr 2020]; Disponible sur: http://gco.iarc.fr/resources/paf-france_fr.php
5. Haute Autorité de Santé - Pertinence du dépistage du cancer broncho-pulmonaire en France - Point de situation sur les données disponibles - Analyse critique des études contrôlées randomisées [Internet]. [cité 4 avr 2020]. Disponible sur: https://www.has-sante.fr/jcms/c_2001613/fr/pertinence-du-depistage-du-cancer-broncho-pulmonaire-en-france-point-de-situation-sur-les-donnees-disponibles-analyse-critique-des-etudes-controlees-randomisees
6. Soubani AO. The evaluation and management of the solitary pulmonary nodule. *Postgrad Med J.* 1 sept 2008;84(995):459-66.
7. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol.* 1 janv 2016;11(1):39-51.
8. International Early Lung Cancer Action Program Investigators, Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, et al. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med.* 26 oct 2006;355(17):1763-71.
9. Doroshow DB, Herbst RS. Treatment of Advanced Non–Small Cell Lung Cancer in 2018. *JAMA Oncol.* 1 avr 2018;4(4):569-70.
10. Wilson JMG, Jungner G, Organization WH. Principes et pratique du dépistage des maladies [Internet]. Genève : Organisation mondiale de la Santé; 1970 [cité 16 avr 2020]. Disponible sur: <https://apps.who.int/iris/handle/10665/41503>

11. Haute Autorité de Santé. Comment évaluer a priori un programme de dépistage ? [Internet]. 2004. Disponible sur: https://www.has-sante.fr/jcms/c_433375/fr/comment-evaluer-a-priori-un-programme-de-depistage
12. Gill RR, Jaklitsch MT, Jacobson FL. Controversies in Lung Cancer Screening. *J Am Coll Radiol JACR*. févr 2016;13(2 Suppl):R2-7.
13. Party BTSS of CS of GBIW. Guidelines on the selection of patients with lung cancer for surgery. *Thorax*. 1 févr 2001;56(2):89-108.
14. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 4 août 2011;365(5):395-409.
15. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N Engl J Med*. 6 févr 2020;382(6):503-13.
16. Haute Autorité de Santé. Dépistage du tabagisme et prévention des maladies liées au tabac [Internet]. Haute Autorité de Santé. [cité 20 avr 2020]. Disponible sur: https://www.has-sante.fr/jcms/c_2635050/fr/depistage-du-tabagisme-et-prevention-des-maladies-liees-au-tabac
17. Oudkerk M, Devaraj A, Vliegenthart R, Henzler T, Prosch H, Heussel CP, et al. European position statement on lung cancer screening. *Lancet Oncol*. 1 déc 2017;18(12):e754-66.
18. Pastorino U, Boffi R, Marchianò A, Sestini S, Munarini E, Calareso G, et al. Stopping Smoking Reduces Mortality in Low-Dose Computed Tomography Screening Participants. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer*. 2016;11(5):693-9.
19. Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA*. 2 nov 2011;306(17):1865-73.
20. Moyer VA. Screening for Lung Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 4 mars 2014;160(5):330-8.
21. Leleu O, Auquier M, Carre O, Chauffert B, Petigny V, Trancart B, et al. Dépistage du cancer du poumon par scanner thoracique basse irradiation dans la Somme : résultats à 1 an. *Rev Mal Respir*. 1 janv 2018;35:A17.
22. Marshall D, Simpson KN, Earle CC, Chu C-W. Economic decision analysis model of screening for lung cancer. *Eur J Cancer*. 1 sept 2001;37(14):1759-67.
23. Gendarme S, Perrot É, Reskot F, Bhoowabul V, Fourre G, Souquet P-J, et al. [Economic impact of lung cancer screening in France: A modeling study]. *Rev Mal Respir*. sept 2017;34(7):717-28.

24. van der Aalst CM, ten Haaf K, de Koning HJ. Lung cancer screening: latest developments and unanswered questions. *Lancet Respir Med*. 1 sept 2016;4(9):749-61.
25. I H, Cho J-Y. Lung Cancer Biomarkers. *Adv Clin Chem*. 2015;72:107-70.
26. Hofman P. Les nouveaux outils biologiques pour améliorer le dépistage du cancer bronchique. *Rev Mal Respir Actual*. oct 2019;11(3):200-6.
27. Zhou J, Huang Z-A, Kumar U, Chen DDY. Review of recent developments in determining volatile organic compounds in exhaled breath as biomarkers for lung cancer diagnosis. *Anal Chim Acta*. 15 déc 2017;996:1-9.
28. Phillips M, Bauer TL, Pass HI. A volatile biomarker in breath predicts lung cancer and pulmonary nodules. *J Breath Res*. 19 2019;13(3):036013.
29. Cicoella A. Les composés organiques volatils (COV) : définition, classification et propriétés. *Rev Mal Respir*. 1 févr 2008;25(2):155-63.
30. Rocco G, Pennazza G, Santonico M, Longo F, Rocco R, Crucitti P, et al. Breathprinting and Early Diagnosis of Lung Cancer. *J Thorac Oncol*. 1 juill 2018;13(7):883-94.
31. Fiches toxicologiques - Publications et outils - INRS [Internet]. [cité 5 mai 2020]. Disponible sur: <http://www.inrs.fr/publications/bdd/fichetox.html>
32. Sun X, Shao K, Wang T. Detection of volatile organic compounds (VOCs) from exhaled breath as noninvasive methods for cancer diagnosis. *Anal Bioanal Chem*. 1 avr 2016;408(11):2759-80.
33. Horváth I, Barnes PJ, Loukides S, Sterk PJ, Högman M, Olin A-C, et al. A European Respiratory Society technical standard: exhaled biomarkers in lung disease. *Eur Respir J* [Internet]. 1 avr 2017 [cité 25 mars 2020];49(4). Disponible sur: <http://erj.ersjournals.com/content/49/4/1600965>
34. ReCIVA® Breath Sampler [Internet]. [cité 20 avr 2020]. Disponible sur: <https://www.owlstonemedical.com/products/reciva/>
35. Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, O'Neill HJ. Volatile organic compounds in exhaled air from patients with lung cancer. *Clin Chem*. 1 août 1985;31(8):1278-82.
36. Jia Z, Patra A, Kutty VK, Venkatesan T. Critical Review of Volatile Organic Compound Analysis in Breath and In Vitro Cell Culture for Detection of Lung Cancer. *Metabolites* [Internet]. 18 mars 2019 [cité 18 avr 2020];9(3). Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6468373/>
37. Saalberg Y, Wolff M. VOC breath biomarkers in lung cancer. *Clin Chim Acta*. août 2016;459:5-9.
38. Finamore P, Scarlata S, Incalzi RA. Breath analysis in respiratory diseases: state-of-the-art and future perspectives. *Expert Rev Mol Diagn*. 2019;19(1):47-61.

39. Li W, Dai W, Liu M, Long Y, Wang C, Xie S, et al. VOC biomarkers identification and predictive model construction for lung cancer based on exhaled breath analysis: research protocol for an exploratory study. *BMJ Open*. août 2019;9(8):e028448.
40. Hulo S. Analyse des Composés Organiques Volatiles dans l'Air Exhalé comme outil diagnostique des Cancers Thoraciques (CATOCOV).
41. Centers for Disease Control and Prevention (CDC). Cigarette smoking among adults--United States, 2000. *MMWR Morb Mortal Wkly Rep*. 26 juill 2002;51(29):642-5.
42. West H (Jack), Jin JO. Performance Status in Patients With Cancer. *JAMA Oncol*. 1 oct 2015;1(7):998-998.
43. Zhang X, Gao B, Creamer AE, Cao C, Li Y. Adsorption of VOCs onto engineered carbon materials: A review. *J Hazard Mater*. 15 sept 2017;338:102-23.
44. Lourenço C, Turner C. Breath Analysis in Disease Diagnosis: Methodological Considerations and Applications. *Metabolites*. 20 juin 2014;4(2):465-98.
45. Gaj K, Pakuluk A. Volatile Methyl Siloxanes as Potential Hazardous Air Pollutants. *Pol J Environ Stud*. 20 mai 2015;24(3):937-43.
46. Lieberman MW, Lykissa ED, Barrios R, Ou CN, Kala G, Kala SV. Cyclosiloxanes produce fatal liver and lung damage in mice. *Environ Health Perspect*. févr 1999;107(2):161-5.
47. Rutjes AWS, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PMM. Case-control and two-gate designs in diagnostic accuracy studies. *Clin Chem*. août 2005;51(8):1335-41.
48. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. *Am J Epidemiol*. 15 avr 2007;165(8):874-81.
49. Steyerberg EW. Validation of Prediction Models. In: Steyerberg EW, éditeur. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* [Internet]. Cham: Springer International Publishing; 2019 [cité 8 mai 2020]. p. 329-44. (Statistics for Biology and Health). Disponible sur: https://doi.org/10.1007/978-3-030-16399-0_17
50. Patassini S, O'Neill M, Layton D, Krizek E, Palo GD, Williams M. Exhaled Breath Analysis: A Tool For Exposomics – Determining Smoking Status Using VOC Biomarkers. :1.
51. Haick H, Broza YY, Mochalski P, Ruzsanyi V, Amann A. Assessment, origin, and implementation of breath volatile cancer markers. *Chem Soc Rev*. 7 mars 2014;43(5):1423-49.
52. Casset A, de Blay F. Effets sur la santé des composés organiques volatils de l'habitat. *Rev Mal Respir*. 1 avr 2008;25(4):475-85.

53. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med.* 3 févr 2004;140(3):189-202.
54. Larose TL, Meheus F, Brennan P, Johansson M, Robbins HA. Assessment of Biomarker Testing for Lung Cancer Screening Eligibility. *JAMA Netw Open.* 2 mars 2020;3(3):e200409-e200409.
55. NHS Lothian. Non-invasive Detection of Pneumonia in Context of Covid-19 Using Gas Chromatography - Ion Mobility Spectrometry (GC-IMS) [Internet]. *clinicaltrials.gov*; 2020 mars [cité 17 mai 2020]. Report No.: NCT04329507. Disponible sur: <https://clinicaltrials.gov/ct2/show/NCT04329507>

AUTEUR :	Nom : RÉMY	Prénom : Clémence
Date de soutenance : 16 juin 2020		
Titre de la thèse : Analyse des composés organiques volatils dans l'air exhalé de patients atteints d'un cancer broncho-pulmonaire et de témoins sains.		
Thèse - Médecine - Lille 2020		
Cadre de classement : <i>Doctorat de médecine</i>		
DES + spécialité : <i>Pneumologie</i>		
Mots-clés : Cancer broncho-pulmonaire – Dépistage – Scanner – Biomarqueurs – Composés organiques volatils		
Résumé :		
<p>Contexte : Le dépistage du cancer broncho-pulmonaire (CBP) par scanner thoracique permet une réduction de mortalité par ce cancer mais présente des limites (irradiation, coût). Les composés organiques volatils (COVs), mesurés dans l'air exhalé, pourraient permettre de mieux identifier la population cible au dépistage par scanner. L'objectif est de décrire et comparer les COVs retrouvés chez des patients atteints d'un CBP et des témoins sains puis de confirmer les résultats des COVs déjà décrits dans la littérature comme associés au CBP.</p> <p>Méthode : Nous avons mené une étude cas-témoin interventionnelle multicentrique. Les patients atteints d'un CBP et naïfs de traitement ont été recrutés au CHU de Lille et les témoins sains à l'institut Pasteur de Lille. L'air exhalé a été collecté dans des tubes de sorbant via un dispositif de recueil validé puis analysé dans un laboratoire indépendant par une technique de chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC-MS).</p> <p>Résultats : Quatorze cas et 15 témoins ont été inclus. Nous avons mis en évidence 11 COVs associés au CBP parmi lesquels 5 étaient déjà décrits dans la littérature : 2-propanol ($p < 0,001$), 2,4-diméthylheptane ($p < 0,001$), benzène ($p < 0,001$), éthanol ($p < 0,001$) et toluène ($p = 0,001$). Six COVs inédits étaient également associés au CBP : cyclotétrasiloxane, octaméthyl- ($p = 0,005$), cyclotrisiloxane, hexaméthyl- ($p < 0,001$), disiloxane, 1-éthényl-1,1,3,3-tétraméthyl-3-(2-propényl)- ($p = 0,008$), disiloxane, hexaméthyl- ($p < 0,001$), triméthylsilanol ($p < 0,001$) et trisiloxane, octaméthyl- ($p < 0,001$).</p> <p>Conclusion : Nous avons identifié 11 COVs associés au CBP parmi lesquels 6 appartiennent à la famille des siloxanes et n'ont jamais été décrits dans la littérature. Ces résultats et leur implication diagnostique méritent d'être confirmés au sein d'une population plus importante.</p>		
<u>Composition du Jury :</u>		
Président : Monsieur le Professeur Régis MATRAN		
Asseseurs : Monsieur le Docteur Sébastien HULO, Monsieur le Docteur Clément GAUVAIN		
Directeur de thèse : Monsieur le Professeur Arnaud SCHERPEREEL		