

Utilité d'un modèle prédictif pour un diagnostic plus précoce du cancer du poumon ?

Question clinique

Quelles sont, en pratique de médecine générale, les données présentes dans les dossiers de patients ultérieurement identifiés avec un cancer du poumon, pouvant servir à l'élaboration d'un modèle prédictif permettant un diagnostic plus précoce ?

Référence Iyen-Omofoman B, Tata LJ, Baldwin DR, et al. Using socio-demographic and early clinical features in general practice to identify people with lung cancer earlier. *Thorax* 2013;68:451-9.

Analyse Alain Van Meerhaeghe, pneumologue, Hôpital Vésale Charleroi

Texte sous la responsabilité de la rédaction francophone

Contexte

La survie à 5 ans du cancer du poumon est très dépendante du stade estimé par la classification TNM (Tumor, Node, Metastase). Elle évolue en effet de 73 % pour les stades pT1a à 13 % pour les pT4 dans le cadre des cancers non à petites cellules¹. Comme l'aspect curatif est très dépendant du stade au moment du diagnostic et qu'au Royaume-Uni la survie est moins bonne que dans les pays comparables, des épidémiologistes de Nottingham ont développé et validé un modèle prédictif permettant un diagnostic plus précoce. L'apport potentiel de ce travail par rapport à un modèle antérieur² est qu'il rejette les symptômes se développant dans la période précédant immédiatement le diagnostic, recherchant ainsi une validation de variables prédictives plus précoces.

Résumé de l'étude

Population étudiée

- population âgée de plus de 40 ans reprise dans la base de données « The Health Improvement Network » (THIN), représentative de la pratique de la médecine générale au Royaume-Uni
- pour construire le modèle prédictif : 12 074 cas de cancer du poumon (diagnostiqués entre janvier 2000 et fin juillet 2009, survenus au moins après 1 an d'inclusion dans la base de données) et 120 731 cas contrôle, étudiés rétrospectivement
- pour la validation prospective du modèle prédictif : 1 826 293 patients âgés > 39 ans à la date du 29 juillet 2009 ; 1 728 cas de cancer du poumon observés sur un an de suivi.

Protocole d'étude

- étude cas-témoins pour construire un modèle prédictif puis étude prospective de validation du modèle
- variables prédictives analysées: âge (par tranches d'âge de 5 ans), sexe, statut socio-économique (score de déprivation de Townsend, par quintile) et antécédents de tabagisme
- symptômes analysés : ceux repris dans les guidelines de NICE³ à savoir hémoptysies et tout autre symptôme persistant et inexpliqué (toux, douleur thoracique/scapulaire, dyspnée, perte de poids, rauçité de voix) ainsi que les 6 symptômes et diagnostics les plus souvent retrouvés dans les dossiers des patients (infections des voies aériennes supérieures et inférieures, infections pulmonaires non spécifiques, constipation, dépression, BPCO)
- relevé des protocoles des radiographies de thorax, des biologies, du nombre de consultations, de tous les symptômes, manoeuvres diagnostiques et consultations dans la période comprise entre 2 ans et 4 mois avant le diagnostic du cancer du poumon.

Mesure des résultats

- modèle prédictif établi en analyse multivariée par régression logistique en retenant les variables associées au diagnostic de cancer du poumon en analyse univariée ($p < 0,05$)
- validation du modèle prédictif : en faisant varier les cut-off pour les variables, calcul de différentes valeurs du couple sensibilité/spécificité du modèle et du pouvoir discriminant global du modèle (**courbe ROC, AUC**) ; comparaison avec le modèle proposé par NICE, basé uniquement sur les symptômes³.

Résultats

Etude cas-témoins

- moyenne du temps de suivi avant le diagnostic de 9,5 années pour les cas et de 9,1 années pour les témoins
- variables retenues dans le modèle avec régression logistique : âge, sexe, quintiles du score de Townsend, tabagisme (statut et catégorie des plus grands consommateurs (> 40 cigarettes/j)), nombre de consultations demandées pour d'autres motifs ; symptômes et diagnostics retenus : toux, hémoptysie, dyspnée, perte de poids, infections des voies aériennes basses, infections pulmonaires non spécifiques, BPCO, douleur thoracique/scapulaire, rauçité de voix, et infection des voies aériennes supérieures (= variables explicatives)
- cote du cancer du poumon augmentant avec l'âge, la gravité des déprivations socio-économiques, la consommation de cigarettes (surtout fumeurs actifs) et le sexe mâle.

Validation du modèle prédictif

- versus modèle développé par NICE : pour identifier le même nombre de vrais positifs, le modèle de NICE nécessite un nombre significativement plus élevé de patients devant subir une radiographie de thorax avec AUC de courbe ROC à 0,64 (pour 0,88 pour le modèle ici testé).

Conclusion des auteurs

Les auteurs concluent que leur nouveau modèle est substantiellement plus performant que le modèle proposé par le guide de pratique de NICE et que les autres modèles comparables. Il peut prédire les cas de cancer du poumon suffisamment tôt pour qu'il soit détecté plus probablement à un stade curatif en permettant aux médecins généralistes de mieux stratifier le risque chez leurs patients. Une étude clinique est nécessaire pour quantifier le bénéfice absolu chez ces patients et le rapport coût/efficacité de ce modèle dans la pratique.

Financement de l'étude bourse de doctorat du conseil de recherche sociale et économique (R.-U.).

Conflits d'intérêt des auteurs aucun n'est signalé.

Considérations sur la méthodologie

Cette étude d'observation rétrospective (étude cas-témoin) puis prospective (cohorte de validation) cherche à déterminer le rôle prédictif précoce de certaines variables socio-démographiques et cliniques dans le risque de développer un cancer pulmonaire. Comme toute étude d'observation, elle peut souffrir de la présence de facteurs de confusion inconnus et de biais non détectés. Pour la construction du modèle, l'étude cas-témoin présente plusieurs avantages dont une réalisation rapide et peu coûteuse ; elle est adaptée à l'étude de maladies ayant une longue période de latence et à fréquence faible. Enfin, elle permet l'étude simultanée de plusieurs facteurs de risques.

Des points essentiels de méthodologie doivent être respectés pour éviter au maximum les biais et les facteurs de confusion. Tout d'abord la définition des cas doit être exacte et précise, ici dans le cadre des cancers pulmonaires la valeur de vérité des diagnostics est très élevée et fiable. Les critères d'inclusion et d'exclusion doivent être appliqués de manière identique aux cas et aux témoins ; ce qui est réalisé dans cette étude. Une préoccupation constante de toute étude épidémiologique est la généralisation des résultats. Le choix de cas issus d'une base de données réputée représentative de la pratique de la médecine générale au Royaume-Uni est fondamental et, l'appariement aléatoire des témoins issus de la même pratique est également un pré-requis tout à fait respecté.

Face à des pathologies dont l'incidence est relativement faible, il est recommandé d'augmenter le nombre de témoins appariés par cas pour augmenter la **puissance** de l'étude. Cependant, au-delà de 4 témoins par cas le gain de puissance est très limité et nous ne voyons pas très bien pourquoi les auteurs sont allés jusqu'à 10 témoins par cas.

La validation et l'estimation globale du pouvoir discriminant du modèle par une courbe ROC dans une cohorte prospective rencontre les critères de qualité pour le but à atteindre.

Les auteurs ont divisé la période précédant l'établissement du diagnostic définitif en deux parties distinctes (4-12 mois et 13-24 mois). L'analyse des symptômes, de la biologie et des consultations avant le diagnostic montraient des odds ratios plus élevés pour le risque de cancer du poumon dans la période 4-12 mois que dans la période 13-24 mois. Toutes les analyses dont il est question dans les résultats donnés ci-dessus, se sont donc correctement focalisées sur la période 4-12 mois avant le diagnostic.

Une méthodologie similaire à celle développée dans le Thoracic Surgery Scoring System (thoracscore) a permis de prédire le risque de cancer du poumon pour chaque patient⁴.

Interprétation des résultats

Les résultats suggèrent qu'une association de variables présentes dans les dossiers de médecine générale peut permettre un diagnostic plus précoce du cancer du poumon dans une telle population.

La valeur prédictive de ce modèle utilisant des variables issues de la pratique courante est potentiellement utile et le recours à ce modèle non coûteux.

Il convient cependant de comparer l'apport d'une telle approche avec le dépistage. Les résultats de l'essai randomisé du National Lung Screening Trial (NLST)⁵ montrent que le dépistage du cancer pulmonaire par CT-scan à faible dose versus radio de thorax dans une population de 53 454 fumeurs de plus de 30 paquets/années entraîne une différence absolue de mortalité spécifique de 0,28 %. Le NND par cette technique pour éviter un décès par cancer bronchique sur 6,5 ans est de 353 (IC à 95 % de 198 à

1 612). En ce qui concerne la mortalité globale, le NND pour éviter un décès est de 219 (IC à 95 % de 112 à 5 626).

A l'heure actuelle, le nombre de faux positifs aboutissant à des manoeuvres diagnostiques et thérapeutiques laisse ouverte la question des rapports coût/efficacité et coût/bénéfice d'une telle stratégie. Les radiations induites par les techniques d'imagerie avec risque de cancérisation secondaire sont également source de questions.

Enfin, le nombre de surdiagnostics (découverte de cancers qui n'auraient pas affecté la vie du sujet) est en cours d'évaluation. Dans l'étude NLST, il existe un excès de 120 cas diagnostiqués dans le groupe soumis au dépistage par le CT-scan à faible dose comparé à la radiographie de thorax, mais un suivi à long terme est nécessaire avant de pouvoir conclure.

Une revue systématique récente n'apporte pas de réponse plus précise⁶.

Des études randomisées en cours comme le PCLO (prostate, lung, colorectal, ovarian cancer screening) vont peut-être permettre d'affiner nos connaissances sur l'utilité du dépistage.

Les apports potentiels du diagnostic précoce tel qu'étudié dans l'étude ici analysée, par rapport à un dépistage systématique, sont, outre l'absence de coût pour l'estimation faite par le modèle, la réduction du taux des faux positifs et du surdiagnostic. Une évaluation prospective au sein d'essais randomisés devra être faite afin d'estimer la balance bénéfices/risques de cette approche et de déterminer si les bénéfices de réduction des manoeuvres diagnostiques et l'absence de surdiagnostic sont supérieurs aux conséquences d'un retard diagnostique qui serait observé chez certains patients.

Conclusion de Minerva

Cette étude d'observation méthodologiquement solide suggère qu'il est possible de construire un modèle prédictif du risque de cancer du poumon dans une patientèle de médecine générale à partir de données simples récoltées en pratique quotidienne.

Pour la pratique

Le guide de pratique de NICE³ recommande d'utiliser un modèle basé sur les symptômes pour dépister plus précocement un cancer du poumon en médecine générale. Ce modèle conduit à effectuer 421 radiographies thoraciques pour identifier 1 cas de cancer du poumon. Le modèle ici construit et validé dans une population britannique, basé sur des données socio-démographiques et cliniques, diminuerait le nombre de faux-positifs et donc de radiographies à réaliser pour identifier un cas de cancer du poumon.

Deux problèmes doivent être cependant encore être abordés : une confrontation au réel par un ou des essais randomisés et une comparaison avec les stratégies de dépistage qui n'ont pas encore apporté de réponse définitive.

Références voir site web www.minerva-ebm.be