

Les règles de prédiction : un guide de lecture

A. Perrier, T. Perneger

Clinique de Médecine 1, Département de Médecine,
Hôpitaux Universitaires de Genève (AP)
Institut de Médecine Sociale et Préventive,
Faculté de Médecine de Genève
Unité Qualité des Soins, Hôpitaux Universitaires de Genève (TP),
Genève, Suisse.

Voir Éditorial p. 177.

Tirés à part : A. Perrier, Clinique de Médecine 1, Hôpital cantonal,
24, rue Micheli-du-Crest, CH-1211 Genève 14, Suisse.
Arnaud.perrier@medecine.unige.ch

Réception version *princeps* à la Revue : 20.12.2002.

Retour aux auteurs pour révision : 13.01.2003.

Réception 1^{ère} version révisée : 17.01.2003.

Acceptation définitive : 20.01.2003.

Publié par anticipation sur www.splf.org/rmr/ le 18.04.2003.

Résumé

Introduction Les scores cliniques ou règles de prédiction sont des instruments qui permettent de quantifier la contribution respective des éléments de l'anamnèse, de l'examen clinique et des examens complémentaires à l'établissement d'un diagnostic ou d'un pronostic chez un patient donné.

Etat des connaissances Le développement de règles de prédiction nécessite plusieurs étapes : 1) la dérivation de la règle à partir d'une population donnée dont on connaît les caractéristiques cliniques (variables de prédiction) et le devenir en termes de diagnostic ou de pronostic (la variable à prédire) ; 2) la validation de la règle dans une population distincte de l'échantillon de dérivation, afin d'évaluer si elle conserve son pouvoir de prédiction ; et 3) l'évaluation de l'impact de la règle de prédiction sur le comportement du clinicien et les conséquences pour les patients.

Perspectives Cet article fournit une grille d'appréciation critique d'une règle de prédiction en prenant l'exemple de deux scores cliniques destinés à prédire la probabilité clinique d'embolie pulmonaire.

Conclusions Les règles de prédiction ou scores cliniques sont devenus un élément indispensable de la trousse à outils du clinicien.

Mots-clés : Médecine factuelle • Règles de prédiction • Embolie pulmonaire.

Rev Mal Respir 2003 ; 20 : 249-56

Clinical prediction rules: a reader's guide

A. Perrier, T. Perneger

Summary

Introduction Clinical prediction rules or scores are instruments that allow to quantify the respective contributions of medical history, clinical examination and laboratory tests to a diagnosis or a prognosis in a given patient.

State of the art The development of clinical prediction rules requires three steps: 1) derivation of the rule from a population with known clinical characteristics and diagnosis or prognosis; 2) validation of the rule in a population distinct from the derivation sample; and 3) evaluation of the impact of the rule on clinician behavior and patient outcomes.

Perspectives This paper provides a guide for the critical appraisal of a prediction rule illustrated by two clinical scores for the assessment of the clinical probability of pulmonary embolism.

Conclusions Clinical prediction rules or scores have become indispensable components of the clinician's toolbox.

Key-words: Evidence-based medicine • Prediction rules • Pulmonary embolism.

Rev Mal Respir 2003 ; 20 : 249-56
Arnaud.perrier@medecine.unige.ch

Cas clinique

Vous accueillez aux urgences de votre hôpital un patient de 65 ans en bonne santé habituelle, qui se plaint de l'apparition rapide d'une douleur basi-thoracique droite augmentée en inspiration sans dyspnée, fièvre, toux, expectorations ou hémoptysie. Il y a 3 semaines, il a été opéré d'une hernie inguinale bilatérale par laparoscopie et les suites post-opératoires ont été simples. Il n'a pas d'antécédent de thrombose veineuse profonde ou d'embolie pulmonaire, et il ne connaît pas de problème thromboembolique veineux dans sa famille. A l'examen clinique, il s'agit d'un patient en bon état général et apyrétique. La tension artérielle est à 150/80 mmHg, la fréquence cardiaque à 88/minute, et la fréquence respiratoire à 20/minute. Il n'y a pas de turgescence jugulaire, l'auscultation cardio-pulmonaire est normale et vous ne notez pas de signes de thrombose veineuse profonde. Les valeurs des gaz sanguins artériels sont les suivantes ; pH 7,46 ; PaCO₂ 35 mmHg, PaO₂ 64 mmHg. La radiographie du thorax est sans particularité. Vous savez que la stratégie diagnostique pour cette suspicion d'embolie pulmonaire va dépendre de la probabilité clinique, que vous estimez intuitivement forte. Toutefois, vous aimeriez conforter cette intuition à l'aide d'une estimation plus objective et explicite, compte-tenu de la complexité du diagnostic de cette affection.

Etape 1 : formulation de la question

Vous connaissez l'existence de règles de prédiction ou scores cliniques à visée diagnostique [1-3] ou pronostique [4-6], mais vous ignorez s'il en existe une qui permette une estimation standardisée de la probabilité clinique d'embolie pulmonaire. La recherche d'une telle règle de prédiction est donc votre question.

Etape 2 : recherche de littérature

Vous faites votre recherche dans la base de données Medline au moyen de l'interface PubMed (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>). Vous entrez d'abord les mots-clé suivants : « pulmonary embolism » et « prediction rule » et obtenez 6 articles [7-12], dont un ne correspond pas à l'objet de votre recherche [10], et deux sont des revues générales [8, 9]. Des trois articles restants, les deux plus pertinents vous semblent être celui de Chagnon et coll. [12] et celui de Wells et coll. [11]. L'article de Stollberger et coll. concerne en effet un échantillon moins important de patients [7]. Vous décidez de compléter votre recherche en employant les mots-clé « pulmonary embolism » et « score ». Toutefois, vu que ce dernier terme est vague, vous mettez une limite à votre recherche en exigeant que ce terme figure dans le titre. Vous trouvez 3 articles qui avaient échappé à votre

recherche précédente [13-15], dont un concerne un score pronostique [14], et un autre est ancien et rapporte un score établi sur 51 patients seulement. Le dernier vous paraît pertinent [13], d'autant qu'il s'agit de l'un des scores comparés dans l'article de Chagnon et coll. [12]. Vous retenez donc l'article de Wicki et coll. [13] et celui de Wells et coll. [11] pour une étude plus approfondie.

Etape 3 : évaluation critique

Vous voilà prêts à évaluer la littérature pertinente à votre question de façon critique. Plusieurs guides de lecture ont été publiés afin de vous assister dans cette tâche [16-19]. Nous allons maintenant analyser les deux articles que vous avez trouvés en appliquant une grille de lecture se rapportant aux règles de prédiction (tableau I).

Le modèle de prédiction est-il cliniquement crédible ?

Toutes les données cliniques pertinentes ont-elles été évaluées pour être incluses dans le modèle ?

La réponse à cette question nécessite une compréhension de la méthode employée pour obtenir une règle de prédiction. Une règle de prédiction peut être définie comme un instrument clinique qui permet de quantifier la contribution respective des éléments de l'anamnèse, de l'examen clinique et est examens complémentaires à l'établissement d'un diagnostic

Tableau I.

Critères méthodologiques de validité et d'utilité d'une règle de prédiction [16-18].

1) Le modèle de prédiction est-il cliniquement crédible ?

— Toutes les données cliniques pertinentes ont-elles été évaluées pour être incluses dans le modèle ?

— Les données composant le score sont-elles faciles à récolter de façon reproductible ?

— La méthodologie pour l'établissement de la règle de prédiction dans l'échantillon de dérivation est-elle valide ?

— Le score est-il facile à calculer ?

2) La règle permet-elle une prédiction suffisamment exacte ?

— La règle permet-elle des prédictions au moins aussi exactes que celles obtenues par le clinicien sans l'aide du score ?

— Dans le cas d'une règle de prédiction diagnostique, le diagnostic a-t-il été posé sur des critères objectifs ne faisant pas intervenir la règle elle-même, dans l'échantillon de dérivation et de validation ?

3) La règle de prédiction est-elle généralisable à d'autres patients ?

— La règle a-t-elle été validée dans un échantillon de patients et un endroit distincts de ceux dans lesquels elle a été établie ?

— Le type de patients et le cadre dans lequel le score a été validé sont-ils décrits de façon détaillée ?

4) L'application de la règle est-elle efficace sur le plan clinique ?

— Les effets de l'utilisation clinique de la règle de prédiction sur le devenir des patients ont-ils été mesurés ?

ou d'un pronostic chez un patient donné [20]. De tels scores sont généralement établis à partir de grandes bases de données qui ont enregistré de façon fiable la variable à prédire, pronostique ou diagnostique, et les caractéristiques cliniques des patients, les variables de prédiction ou facteurs prédictifs (« prédictors »). Il va de soi que la qualité du score ainsi obtenu dépendra de la pertinence clinique des variables évaluées dans le modèle de prédiction. Considérons notre question et examinons les règles de prédiction de Wicki et coll. [13] et Wells et coll. [11] (tableau II). Si la plupart des cliniciens chevronnés s'accorderont sur la pertinence des variables retenues dans ces deux scores, certains seront peut-être surpris par le petit nombre d'éléments qui les composent. L'explication en est la suivante. La technique la plus employée pour construire un modèle de prédiction est l'analyse par régression multivariée ou régression logistique. La première étape consiste à évaluer toutes les variables de prédiction disponibles une à une afin de voir si elles sont ou non statistiquement associées à l'événement que l'on cherche à prédire (analyse univariée). Puis, toutes les variables ayant un pouvoir de prédiction prises isolément sont intégrées dans une analyse multivariée dont le but est d'identifier les facteurs prédictifs indépendants. A ce stade, un certain nombre de variables au pouvoir de prédiction faible ou redondantes avec d'autres éléments cliniques vont

Tableau II.

Estimation de la probabilité clinique d'embolie pulmonaire selon les scores de prédiction clinique de Wells [11] et de Wicki [13].

Score de Wicki	Points	Score de Wells	Points
Antécédent de TVP ou EP	+2	Antécédent de TVP ou EP	+1,5
Rythme cardiaque > 100/min	+1	Rythme cardiaque > 100/min	+1,5
Chirurgie récente	+3	Chirurgie récente ou immobilisation	+1,5
Âge		Signe de TVP	+3
60-79 ans	+1	Diagnostic alternatif moins probable que l'EP	+3
≥ 80 ans	+2		
PaCO₂			
< 36 mmHg	+2	Hémoptysie	+1
36-39 mmHg	+1		
PaO₂			
< 49 mmHg	+4	Cancer	+1
49-59,9 mmHg	+3		
60-71,2 mmHg	+2		
71,3-82,4 mmHg	+1		
Atélectasie	+1		
Élévation d'une coupole diaphragmatique	+1		
Probabilité clinique		Probabilité clinique	
Faible 0-4 points		Faible 0-1 point	
Moyenne 5-8 points		Moyenne 2-6 points	
Forte ≥ 9 points		Forte ≥ 7 points	

disparaître du modèle définitif. Dans le modèle de Wicki et coll. [13], 29 caractéristiques cliniques ont été évaluées, dont 21 étaient significativement associées à l'embolie pulmonaire en analyse univariée, mais le modèle résultant de l'analyse multivariée n'en a retenu que huit. La dyspnée, l'hémoptysie et la tachypnée, notamment, n'ont ainsi pas été incluses dans le score final. On relèvera d'ailleurs que ce processus ne produit jamais un seul modèle, mais génère en réalité de multiples modèles aux ingrédients différents, entre lesquels le choix dépendra du pouvoir de prédiction, mais aussi de la pertinence clinique et de la facilité à récolter les éléments retenus. Le lecteur intéressé trouvera une discussion plus détaillée mais compréhensible en référence [21]. En résumé, tant le score de Wicki [13] que celui de Wells [11] sont basés sur des éléments cliniques pertinents, extraits de bases de données où étaient enregistrées toutes les variables d'intérêt pour prédire l'embolie pulmonaire.

Les données composant le score sont-elles faciles à récolter de façon reproductible ?

Un score ne permet une estimation fiable que si les éléments qui le composent sont faciles à récolter. Ainsi, Wicki et coll. [13] ont choisi de retenir dans leur modèle la mesure des gaz du sang artériel, complètement objective, plutôt que la fréquence respiratoire, plus difficile à mesurer de façon reproductible. Ce choix limite toutefois l'application de ce score aux patients chez lesquels la gazométrie peut être effectuée à l'air ambiant. D'autres éléments de ce score sont plus dépendants de l'observateur, en particulier les éléments d'interprétation de la radiographie thoracique. Dans la règle de Wells [11], un élément sans doute discutable à cet égard est la nécessité de se prononcer sur la vraisemblance d'un diagnostic alternatif à celui de l'embolie pulmonaire, en raison de son caractère subjectif.

La méthodologie pour l'établissement de la règle de prédiction dans l'échantillon de dérivation est-elle valide ?

Il existe quelques pièges méthodologiques importants à connaître. Ainsi, il faut pouvoir vérifier qu'il n'y a pas un nombre excessif de données cliniques manquantes dans la base de données, qui diminueraient considérablement la fiabilité de l'analyse multivariée. De même, la taille de l'échantillon doit être suffisante pour permettre une mesure précise de la capacité de prédiction du modèle. Dans notre exemple et pour les deux scores considérés, les intervalles de confiance de la prévalence d'embolie pulmonaire dans les catégories de faible probabilité et de probabilité intermédiaire sont étroits, alors qu'ils sont larges pour la classe des fortes probabilités cliniques, qui regroupe moins de patients. L'exactitude et la fiabilité d'un modèle dépend également du nombre d'événements à prédire qui sont effectivement survenus dans la population étudiée. La littérature suggère en effet qu'il faut 5 à 10 événements pour chaque variable prise en considération dans le modèle, sans

quoi on risque d'obtenir un modèle « surajusté ». Dans le cas du score de Wicki qui évaluait 29 variables, il fallait donc entre 145 et 290 cas d'embolie pulmonaire, et il y en avait 296. En revanche, il peut être difficile de chercher à prédire un événement qui ne surviendrait que chez une très faible proportion des patients, car le nombre de sujets nécessaires pour établir un modèle de prédiction peut devenir rédhibitoire. Malheureusement, les publications sont souvent sélectives et ne donnent pas la liste complète des variables testées. Enfin, il est important que l'évaluation des éléments cliniques qui serviront à dériver le score soit faite par des cliniciens ignorant si le patient a ou non la pathologie recherchée.

Le score est-il facile à calculer ?

Une règle de prédiction nécessitant des calculs compliqués a peu de chances d'être d'une réelle aide au clinicien sur le terrain, même s'il est vrai que les ordinateurs de poche peuvent désormais effectuer ce genre de tâches. C'est pourquoi le calcul de la plupart des scores récents repose sur la simple addition de points, le nombre de points attribué à chaque variable reflétant son poids dans la prédiction. Par exemple, dans le score de Wells [11], l'existence de signes de thrombose veineuse profonde a 3 fois plus de poids que celle d'une hémoptysie. Ces pondérations sont obtenues à partir des coefficients de régression [21].

La règle permet-elle une prédiction suffisamment exacte ?

La règle permet-elle des prédictions au moins aussi exactes que celles obtenues par le clinicien sans l'aide du score ?

La prédiction clinique empirique ou implicite a été réhabilitée dans la suspicion d'embolie pulmonaire bien avant l'apparition de scores. C'est l'étude « Prospective Investigation On Pulmonary Embolism Diagnosis » (PIOPED) qui a démontré en 1990 que le clinicien parvenait à stratifier les patients en trois groupes correspondant à une prévalence croissante d'embolie pulmonaire (faible probabilité clinique, 9 % ; probabilité intermédiaire, 30 % ; probabilité forte, 68 %) (*fig. 1*) [22]. Deux études genevoises [23, 24] dont l'une multicentrique [24] ont confirmé cette bonne performance de l'estimation empirique et, tout récemment, l'étude multicentrique française ESSEP [25] est venue renforcer encore la validité du jugement implicite (*fig. 1*). Ainsi, pour avoir une quelconque utilité, tout score dans ce domaine sera tenu de faire aussi bien que l'évaluation empirique, et c'est bien le cas des scores de Wicki [13] et de Wells [11] (*fig. 1*).

On notera toutefois que l'évaluation implicite a aussi ses limites, sans quoi la création de scores prédictifs n'aurait que peu d'intérêt. En effet, l'intuition du clinicien est entachée de biais cognitifs de toutes sortes [26, 27] et les scores ont donc pour but de rendre explicites les éléments du raisonnement clinique. Le biais lié au souvenir de cas dramatiques est ici

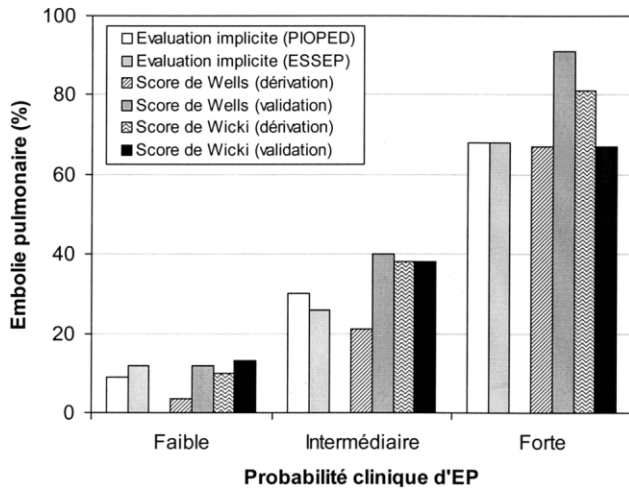


Fig. 1. Valeur prédictive de la probabilité clinique empirique [22-25], et des scores de Wells et de Wicki dans les échantillons de dérivation [11, 13] et de validation [12].

particulièrement illustratif : le clinicien qui aura vu décéder d'une embolie pulmonaire massive une jeune patiente sans aucun facteur de risque de maladie thromboembolique veineuse et à la clinique peu évocatrice va surestimer systématiquement la probabilité de ce diagnostic chez tous ses patients suivants. D'autre part, le jugement clinique dépend de l'expérience, et la qualité de l'évaluation implicite dépend du niveau de formation, ce qui a été bien documenté dans le domaine de l'embolie pulmonaire [28]. Ainsi, on peut espérer que la concordance inter-observateurs de l'évaluation par un score sera meilleure que par la simple intuition. De plus, leur caractère explicite en fait des outils didactiques utiles dans la formation des cliniciens moins expérimentés. À l'inverse, les scores établis sur des bases statistiques captent efficacement les facteurs prédictifs du « patient moyen », mais ces modèles peuvent être inefficaces dans des cas particuliers. L'existence d'une thrombophilie connue, par exemple, n'intervient ni dans le score de Wicki [13] ni dans celui de Wells [11], alors qu'une telle notion ferait légitimement revoir à la hausse la probabilité clinique. Le clinicien doit donc toujours se réserver un « droit de veto » vis-à-vis du score.

Dans le cas d'une règle de prédiction diagnostique, le diagnostic a-t-il été posé sur des critères objectifs ne faisant pas intervenir la règle elle-même, dans l'échantillon de dérivation et de validation ?

Pour l'établissement de leur score, Wells et coll. [11] ont recouru à une base de données dans laquelle le diagnostic d'embolie pulmonaire était posé sur la base d'un algorithme complexe faisant intervenir la probabilité clinique et diverses combinaisons d'examen diagnostiques. Néanmoins, dans l'étude de Wells, on ne peut ignorer que les critères diagnostiques n'ont pas été identiques pour tous les patients, puisque la

démarche dépendait de la probabilité clinique, même si celle-ci n'a pas été établie par le score de Wells [11], publié plus tardivement (*tableau II*). Cela est vrai également du score de Wicki [13], dont l'échantillon de dérivation a également subi une stratégie diagnostique stratifiée selon la probabilité clinique jugée implicitement par le clinicien. Dans l'étude de validation de leur score faite par Wells et coll. [29], celui-ci servait à stratifier les patients et à déterminer la séquence diagnostique, et ne constitue donc pas une validation satisfaisante. En revanche, dans l'étude de validation récemment publiée par le groupe de Genève [12] (*fig. 1*), le score de Wells n'était pas connu des internes qui établissaient la probabilité clinique. Le score de Genève était employé dans le cadre d'une étude multicentrique destinée à évaluer un algorithme diagnostique incluant l'angioscanner hélicoïdal. Il a donc indirectement contribué au diagnostic final et n'est donc pas strictement validé [12].

La règle de prédiction est-elle généralisable à d'autres patients ?

La règle a-t-elle été validée dans un échantillon de patients et un endroit distincts de ceux dans lesquels elle a été établie ?

Le talon d'Achille de toute règle de prédiction réside dans son potentiel de généralisation à d'autres patients que ceux chez lesquels elle a été établie. En effet, l'analyse multivariée tend en effet à tailler un « costume sur mesure » à chaque patient, dont les caractéristiques cliniques sont connues. Lorsque l'on applique une règle de prédiction à des patients différents de ceux de l'échantillon de dérivation, elle perd généralement de son pouvoir de prédiction. Si cette perte est minime, on considérera la règle comme validée ; si elle est importante, le score devra être abandonné faute de fiabilité. Le processus de validation d'une règle de prédiction est donc une étape essentielle, qui peut se faire à différents niveaux. On peut la réaliser à l'intérieur même de l'échantillon de dérivation, en le partageant de façon aléatoire et en appliquant le score aux différents sous-groupes obtenus, afin de tester sa stabilité. C'est la validation interne, qui avait été appliquée avec succès à l'échantillon de dérivation dans l'étude de Wicki et coll. [13]. Toutefois, le niveau de preuves associé à un score qui n'aurait subi que ce type de validation est faible (*tableau III*). La *validation externe* consiste en l'application de la règle à des patients distincts de ceux qui ont servi à établir le score. Plus les patients sont variés et représentent un large spectre des patients concernés par la règle de prédiction, des moins atteints aux plus graves, plus la validation est robuste (*tableau III*). De même, la validation d'une règle dans des centres hospitaliers différents de ceux qui ont établi la règle est souhaitable. Une telle validation a été effectuée pour les deux scores [12, 29]. Dans une étude multicentrique canadienne [29], la prévalence de l'embolie pulmonaire était semblable à celle de l'échantillon de dérivation de Wells dans les catégories de probabilité clinique faible (1,3 %

Tableau III.
Niveaux de preuves concernant les règles de prédiction [19].

Niveau 1	<p>– Règles pouvant être employées dans une grande variété de cadres de pratique avec confiance, qui modifient le comportement des cliniciens et améliore les issues pour les patients.</p> <p>Au moins une validation prospective dans une population différente et une analyse démontrant un impact sur le comportement des cliniciens et des conséquences favorables pour les patients.</p>
Niveau 2	<p>– Règles pouvant être employées dans une grande variété de cadres de pratique avec confiance en ce qui concerne leur capacité de prédiction.</p> <p>– Capacité de prédiction confirmée dans au moins un étude prospective incluant un large spectre de patients et de cliniciens ou validée dans plusieurs études de petite taille mais réalisées dans des cadres différents.</p>
Niveau 3	<p>– Règles à employer avec prudence et seulement si les patients de l'étude sont similaires à ceux de son propre cadre de pratique.</p> <p>Validation dans une seule étude prospective de petite taille.</p>
Niveau 4	<p>– Règles nécessitant une confirmation avant de pouvoir être appliquées en clinique.</p> <p>Dérivation mais pas de validation ou validation interne ou dans des base de données rétrospectives.</p>

contre 3,6 % dans l'échantillon de dérivation) et moyenne (16 % contre 21 %). Elle était en revanche plus faible dans la catégorie des fortes probabilités cliniques (38 % contre 67 %). Ceci s'explique sans doute par le fait que la prévalence globale de l'embolie pulmonaire était plus basse dans l'échantillon de validation (9 %) que dans celui de dérivation (16 %). Dans une étude de validation européenne [12], la prévalence de l'embolie pulmonaire dans les différents sous-groupes était significativement plus élevée (probabilité clinique faible, 12 % ; moyenne, 40 % ; forte, 91 %) (fig. 1), une différence à nouveau attribuable à la prévalence globale de l'embolie pulmonaire, qui était ici de 26 %. On voit ainsi l'influence de la prévalence de la maladie sur la performance d'un score diagnostique. Or, il n'est pas toujours aisé d'obtenir que la population à laquelle on applique un score soit comparable à celle de l'échantillon de dérivation, même lorsque les caractéristiques de la population dans laquelle la règle a été dérivée sont décrites de manière détaillée. En effet, pour chercher à prédire l'existence d'une embolie pulmonaire ou de tout autre diagnostic, il faut d'abord y avoir pensé, et les critères de suspicion peuvent varier d'un centre ou d'une personne à l'autre. Enfin, le score de Genève a démontré une performance équivalente à celle observée dans l'échantillon de dérivation dans une étude réalisée dans trois centres européens [12] (fig. 1).

Le type de patients et le cadre dans lequel le score a été validé sont-ils décrits de façon détaillée ?

C'est une condition préalable indispensable à l'application clinique d'une règle de prédiction. En effet, on ne peut

demander d'un modèle qu'il conserve son pouvoir de prédiction si on l'applique à des patients qui n'étaient pas représentés dans l'échantillon de dérivation. Par exemple, le score de Wicki [13] a été dérivé d'une population de patients admis aux urgences et ne peut donc être appliqué aux patients hospitalisés qui ont des caractéristiques trop différentes. Au contraire, l'échantillon de dérivation du score de Wells [11] incluait également des patients hospitalisés.

L'application de la règle est-elle efficace sur le plan clinique ?

Les effets de l'utilisation clinique de la règle de prédiction sur le devenir des patients ont-ils été mesurés ?

C'est le paramètre ultime pour justifier l'emploi d'une règle de prédiction, et pourtant, peu de scores sont évalués jusque-là. La règle d'Ottawa [30, 31], destinée à diminuer le recours systématique à la radiographie de la cheville en cas de traumatisme, a démontré son efficacité dans une étude regroupant 5 centres d'urgence français [32]. Le score de risque de la pneumonie acquise à domicile développé par Fine et coll. [6] a également fait l'objet d'une étude d'impact [33, 34]. Dans notre exemple, Wells et coll. ont employé leur score dans une étude prospective [29] et démontrent que l'embolie pulmonaire a pu être écartée chez 437 patients sur 930 (47 %) par l'association d'une faible probabilité clinique (estimée par le score) et du dosage plasmatique des D-dimères, un test de laboratoire bon marché, évitant ainsi à ces patients des investigations plus lourdes et coûteuses. Le score de Wicki a été employé dans une étude diagnostique récente [35], et a permis de simplifier la démarche diagnostique chez les patients dont la probabilité clinique était faible ou intermédiaire. Toutefois, aucune de ces études n'avait de groupe ou d'hôpital contrôle qui permette d'évaluer l'importance des économies ainsi réalisées ou de l'amélioration du pronostic des patients.

Résolution du cas clinique

Vous concluez que les deux règles de prédiction de l'embolie pulmonaire que vous venez d'évaluer sont valides et correspondent à un niveau de preuve de niveau 2, l'impact réel n'en ayant pas été complètement évalué. Vous décidez donc de les appliquer à votre patient. Selon les deux scores [11, 13], votre patient a une probabilité clinique d'embolie pulmonaire intermédiaire. Or, en employant votre jugement clinique implicite, vous lui aviez assigné une probabilité clinique forte. Que faire de ce désaccord ? Soulignons d'abord que de tels désaccords sont fréquents lorsque l'on recourt à une règle complètement standardisée, même si la littérature sur les scores fait très rarement état de cette difficulté. Dans un article récent, Chagnon et coll. [12] ont constaté de tels désaccords entre l'estimation du clinicien et celle obtenue en appliquant le score

de Wicki et coll. [13] chez 20 % des patients, la probabilité clinique étant revue à la hausse dans deux tiers des cas et à la baisse dans le tiers restant. Dans une approche rigoureuse, on devrait plutôt décourager de telles révisions de l'estimation par les scores, étant donné que le clinicien base souvent son évaluation sur les mêmes variables que les scores explicites. Toutefois, comme on l'a vu plus haut, les scores ne tiennent pas compte d'éléments potentiellement importants mais trop rarement présents pour être retenus par une analyse statistique. Il est ainsi intéressant de relever que, dans notre exemple, la capacité de prédiction du score de Wicki corrigé par le jugement du clinicien était un peu meilleure que celle du score brut [12], ce qui démontre que le clinicien valorise des éléments non pris en compte par le score. Vous pourriez donc légitimement choisir de conserver votre évaluation implicite chez ce patient. Toutefois, le fait que la règle de Wells et coll. [11] évalue également la probabilité clinique comme moyenne dans ce cas alors qu'elle incorpore déjà un élément subjectif global (la probabilité d'un diagnostic alternatif) devrait vous conduire à réfléchir à la tendance des cliniciens à surestimer la probabilité d'embolie pulmonaire, et il n'y a pas de réponse définitive à la question des désaccords entre évaluation implicite et par une règle de prédiction. A notre avis, cela n'invalide pas pour autant l'utilité des scores cliniques. En effet, ils ont le mérite d'être didactiques, et donc très utiles pour les cliniciens moins expérimentés, et explicites, ce qui permet aux cliniciens de discuter les désaccords éventuels sur une base claire.

En conclusion, les règles de prédiction sont des instruments très utiles en pratique clinique pour peu que l'on sélectionne des scores établis de façon valide dans l'échantillon de dérivation et ayant fait l'objet d'une validation externe. Enfin, l'utilité d'une règle de prédiction est définitivement établie lorsqu'on a pu démontrer qu'elle modifie réellement la prise en charge médicale et que ce changement a des conséquences positives en termes de consommation de ressources et/ou d'amélioration du devenir des patients.

Références

- 1 Flemons WW, Whitelaw WA, Brant R, Remmers JE : Likelihood ratios for a sleep apnea clinical prediction rule. *Am J Respir Crit Care Med* 1994 ; 150 : 1279-85.
- 2 Krijnen P, van Jaarsveld BC, Steyerberg EW, Man in 't Veld AJ, Schalekamp MA, Habbema JD : A clinical prediction rule for renal artery stenosis. *Ann Intern Med* 1998 ; 129 : 705-11.
- 3 McKinney WP, Heudebert GR, Harper SA, Young MJ, McIntire DD : Validation of a clinical prediction rule for the differential diagnosis of acute meningitis. *J Gen Intern Med* 1994 ; 9 : 8-12.
- 4 Fortescue EB, Kahn K, Bates DW : Development and validation of a clinical prediction rule for major adverse outcomes in coronary bypass grafting. *Am J Cardiol* 2001 ; 88 : 1251-8.
- 5 Rainer TH, Lam PK, Wong EM, Cocks RA : Derivation of a prediction rule for post-traumatic acute lung injury. *Resuscitation* 1999 ; 42 : 187-96.
- 6 Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN : A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997 ; 336 : 243-50.
- 7 Stollberger C, Finsterer J, Lutz W, Stoberl C, Kroiss A, Valentin A, Slany J : Multivariate analysis-based prediction rule for pulmonary embolism. *Thromb Res* 2000 ; 97 : 267-73.
- 8 Miniati M, Pistolesi M : Assessing the clinical probability of pulmonary embolism. *QJ Nucl Med* 2001 ; 45 : 287-93.
- 9 Brown J : Assessment of pretest risk for venous thromboembolic disease. *Emerg Med Clin North Am* 2001 ; 19 : 861-8.
- 10 Reber G, Blanchard J, Bounameaux H, Hoffmeyer P, Miron MJ, Leyvraz PF, de Moerloose P : Inability of serial fibrin monomer measurements to predict or exclude deep venous thrombosis in asymptomatic patients undergoing total knee arthroplasty. *Blood Coagul Fibrinolysis* 2000 ; 11 : 305-8.
- 11 Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, Turpie AG, Bormanis J, Weitz J, Chamberlain M, Bowie D, Barnes D, Hirsh J : Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: Increasing the models utility with the simplified d-dimer. *Thromb Haemost* 2000 ; 83 : 416-20.
- 12 Chagnon I, Bounameaux H, Aujesky D, Roy PM, Gourdier AL, Cornuz J, Perneger T, Perrier A : Comparison of two clinical prediction rules and implicit assessment for suspected pulmonary embolism. *Am J Med* 2002 ; 113 : 269-75.
- 13 Wicki J, Perneger TV, Junod A, Bounameaux H, Perrier A : Assessing clinical probability of pulmonary embolism in the emergency ward : A simple score. *Arch Intern Med* 2001 ; 161 : 92-7.
- 14 Wicki J, Perrier A, Perneger TV, Bounameaux H, Junod AF : Predicting adverse outcome in patients with acute pulmonary embolism: A risk score. *Thromb Haemost* 2000 ; 84 : 548-52.
- 15 Rosselli A, Landini G, Mugnaini C, Piani F : [clinical diagnosis of pulmonary embolism: Validity of a score system and its correlation with therapeutic choices in a case series of 51 patients]. *G Clin Med* 1990 ; 71 : 514-05.
- 16 Wasson JH, Sox HC, Neff RK, Goldman L : Clinical prediction rules. Applications and methodological standards. *N Engl J Med* 1985 ; 313 : p793-799.
- 17 Wasson JH, Sox HC : Clinical prediction rules. Have they come of age? *JAMA* 1996 ; 275 : 641-2.
- 18 Wyatt JC, Atman DG : Prognostic models: Clinically useful or quickly forgotten? *BMJ* 1995 ; 311 : 1539-41.
- 19 McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS : Users' guides to the medical literature: XXII: How to use articles about clinical decision rules. Evidence-based medicine working group. *JAMA* 2000 ; 284 : 79-84.
- 20 Laupacis A, Sekar N, Stiell IG : Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 1997 ; 277 : 488-94.
- 21 Guyatt G, Walter S, Shannon H, Cook D, Jaeschke R, Heddle N : Basic statistics for clinicians: 4. Correlation and regression. *CMAJ* 1995 ; 152 : 497-504.
- 22 The PIOPED Investigators : Value of the ventilation-perfusion scan in acute pulmonary embolism. *JAMA* 1990 ; 263 : 2753-9.
- 23 Perrier A, Bounameaux H, Morabia A, de Moerloose P, Slosman D, Didier D, Unger PF, Junod A : Diagnosis of pulmonary embolism by a

- decision analysis-based strategy including clinical probability, d-dimer levels, and ultrasonography: A management study. *Arch Intern Med* 1996 ; 156 : 531-6.
- 24 Perrier A, Desmarais S, Miron MJ, de Moerloose P, Lepage R, Slosman D, Didier D, Unger PF, Patenaude JV, Bounameaux H : Non-invasive diagnosis of venous thromboembolism in outpatients. *Lancet* 1999 ; 353 : 190-5.
- 25 Musset D, Parent F, Meyer G, Maitre S, Girard P, Leroyer C, Revel MP, Carette MF, Laurent M, Charbonnier B, Laurent F, Mal H, Nonent M, Lancar R, Grenier P, Simonneau G : Diagnostic strategy for patients with suspected pulmonary embolism: A prospective multicentre outcome study. *Lancet* 2002 ; 360 : 1914-20.
- 26 Elstein AS : Heuristics and biases: Selected errors in clinical reasoning. *Acad Med* 1999 ; 74 : 791-4.
- 27 Kassirer JP, Kopelman RI : Cognitive errors in diagnosis: Instantiation, classification, and consequences. *Am J Med* 1989 ; 86 : 433-41.
- 28 Rosen MP, Sands DZ, Morris J, Drake W, Davis RB : Does a physician's ability to accurately assess the likelihood of pulmonary embolism increase with training? *Acad Med* 2000 ; 75 : 1199-205.
- 29 Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, Forgie M, Kovacs G, Ward J, Kovacs MJ : Excluding pulmonary embolism at the bedside without diagnostic imaging: Management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. *Ann Intern Med* 2001 ; 135 : 98-107.
- 30 Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR : A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med* 1992 ; 21 : 384-90.
- 31 Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Reardon M, Stewart JP, Maloney J : Decision rules for the use of radiography in acute ankle injuries. Refinement and prospective validation. *JAMA* 1993 ; 269 : 1127-32.
- 32 Auleley GR, Ravaud P, Giraudeau B, Kerboull L, Nizard R, Massin P, Garreau de Loubresse C, Vallee C, Durieux P : Implementation of the ottawa ankle rules in france. A multicenter randomized controlled trial. *JAMA* 1997 ; 277 : 1935-9.
- 33 Marrie TJ, Lau CY, Wheeler SL, Wong CJ, Vandervoort MK, Feagan BG : A controlled trial of a critical pathway for treatment of community-acquired pneumonia. Capital study investigators. Community-acquired pneumonia intervention trial assessing levofloxacin. *JAMA* 2000 ; 283 : 749-55.
- 34 Atlas SJ, Benzer TL, Borowsky LH, Chang Y, Burnham DC, Metlay JP, Halm EA, Singer DE : Safely increasing the proportion of patients with community-acquired pneumonia treated as outpatients: An interventional trial. *Arch Intern Med* 1998 ; 158 : 1350-6.
- 35 Perrier A, Roy PM, Aujesky D, Chagnon I, Howarth N, Gourdier AL, Leftheriotis G, Bargouth G, Cornuz J, Hayoz D, Bounameaux H : Diagnosing pulmonary embolism with clinical assessment, d-dimer, venous ultrasound and helical computed tomography: A multicenter management study. 2003 ; sous presse.
-