Université de Poitiers
Faculté de Médecine et Pharmacie

ANNEE 2015 Thèse n°

THESE

POUR LE DIPLOME D'ETAT

DE DOCTEUR EN MEDECINE

(décret du 16 janvier 2004)

présentée et soutenue publiquement

le 6 juillet 2015 à Poitiers

par Madame Astrid HARVARD de HAUTECLOCQUE

Validation externe indépendante de scores de prédiction du risque rénal chez des patients diabétiques de type 2

Composition du Jury

Président : Madame le Professeur Virginie Migeot

Membres :

- Monsieur le Professeur Pierre Ingrand
- Madame le Docteur Stéphanie Ragot
- Madame le Docteur Marion Albouy-Llaty

Directeur de thèse : Monsieur le Professeur Samy Hadjadj

Université de Poitiers
Faculté de Médecine et Pharmacie

ANNEE 2015 Thèse n°

THESE

POUR LE DIPLOME D'ETAT

DE DOCTEUR EN MEDECINE

(décret du 16 janvier 2004)

présentée et soutenue publiquement

le 6 juillet 2015 à Poitiers

par Madame Astrid HARVARD de HAUTECLOCQUE

Validation externe indépendante de scores de prédiction du risque rénal chez des patients diabétiques de type 2

Composition du Jury

Président : Madame le Professeur Virginie Migeot

Membres :

- Monsieur le Professeur Pierre Ingrand
- Madame le Docteur Stéphanie Ragot
- Madame le Docteur Marion Albouy-Llaty

Directeur de thèse : Monsieur le Professeur Samy Hadjadj

UNIVERSITE DE POITIERS

Faculté de Médecine et de Pharmacie

Le Doyen,

Année universitaire 2014 - 2015

LISTE DES ENSEIGNANTS DE MEDECINE

Professeurs des Universités-Praticiens Hospitaliers

1. AGIUS Gérard, bactériologie-virologie
2. ALLAL Joseph, thérapeutique
3. BATAILLE Benoît, neurochirurgie
4. BENSADOUN René-Jean, cancérologie – radiothérapie **(en disponibilité 1 an à compter de janvier 2014)**
5. BRIDOUX Frank, néphrologie
6. BURUCOA Christophe, bactériologie - virologie
7. CARRETIER Michel, chirurgie générale
8. CHEZE-LE REST Catherine, biophysique et médecine nucléaire
9. CHRISTIAENS Luc, cardiologie
10. CORBI Pierre, chirurgie thoracique et cardio-vasculaire
11. DEBAENE Bertrand, anesthésiologie réanimation
12. DEBIAIS Françoise, rhumatologie
13. DROUOT Xavier, physiologie
14. DUFOUR Xavier, Oto-Rhino-Laryngologie
15. EUGENE Michel, physiologie **(surnombre jusqu'en 08/2016)**
16. FAURE Jean-Pierre, anatomie
17. FRITEL Xavier, gynécologie-obstétrique
18. GAYET Louis-Etienne, chirurgie orthopédique et traumatologique
19. GICQUEL Ludovic, pédopsychiatrie
20. GILBERT Brigitte, génétique
21. GOMBERT Jean-Marc, immunologie
22. GOUJON Jean-Michel, anatomie et cytologie pathologiques
23. GUILHOT-GAUDEFFROY François, hématologie et transfusion
24. GUILLET Gérard, dermatologie
25. GUILLEVIN Rémy, radiologie et imagerie médicale
26. HADJADJ Samy, endocrinologie et maladies métaboliques
27. HAUET Thierry, biochimie et biologie moléculaire
28. HERPIN Daniel, cardiologie
29. HOUETO Jean-Luc, neurologie
30. INGRAND Pierre, biostatistiques, informatique médicale
31. IRANI Jacques, urologie
32. JABER Mohamed, cytologie et histologie
33. JAAFARI Nematollah, psychiatrie d'adultes
34. JAYLE Christophe, chirurgie thoracique t cardio-vasculaire
35. KARAYAN-TAPON Lucie, cancérologie
36. KEMOUN Gilles, médecine physique et réadaptation **(en détachement 2 ans à compter de janvier 2014)**
37. KITZIS Alain, biologie cellulaire
38. KRAIMPS Jean-Louis, chirurgie générale
39. LECRON Jean-Claude, biochimie et biologie moléculaire
40. LEVARD Guillaume, chirurgie infantile
41. LEVEZIEL Nicolas, ophtalmologie
42. LEVILLAIN Pierre, anatomie et cytologie pathologiques
43. MACCHI Laurent, hématologie
44. MARECHAUD Richard, médecine interne
45. MAUCO Gérard, biochimie et biologie moléculaire
46. MEURICE Jean-Claude, pneumologie
47. MIGEOT Virginie, santé publique
48. MILLOT Frédéric, pédiatrie, oncologie pédiatrique
49. MIMOZ Olivier, anesthésiologie - réanimation
50. MORICHAU-BEAUCHANT Michel, hépato-gastro-entérologie
51. NEAU Jean-Philippe, neurologie
52. ORIOT Denis, pédiatrie
53. PACCALIN Marc, gériatrie
54. PAQUEREAU Joël, physiologie
55. PERAULT Marie-Christine, pharmacologie clinique
56. PERDRISOT Rémy, biophysique et médecine nucléaire
57. PIERRE Fabrice, gynécologie et obstétrique
58. POURRAT Olivier, médecine interne
59. PRIES Pierre, chirurgie orthopédique et traumatologique
60. RICCO Jean-Baptiste, chirurgie vasculaire
61. RICHER Jean-Pierre, anatomie
62. RIGOARD Philippe, neurochirurgie
63. ROBERT René, réanimation
64. ROBLOT France, maladies infectieuses, maladies tropicales
65. ROBLOT Pascal, médecine interne
66. RODIER Marie-Hélène, parasitologie et mycologie
67. SENON Jean-Louis, psychiatrie d'adultes **(surnombre jusqu'en 08/2017)**
68. SILVAIN Christine, hépato-gastro- entérologie
69. SOLAU-GERVAIS Elisabeth, rhumatologie
70. TASU Jean-Pierre, radiologie et imagerie médicale
71. TOUCHARD Guy, néphrologie
72. TOURANI Jean-Marc, cancérologie
73. WAGER Michel, neurochirurgie

Maîtres de Conférences des Universités-Praticiens Hospitaliers

1. ALBOUY-LLATY Marion, santé publique
2. ARIES Jacques, anesthésiologie - réanimation
3. BEBY-DEFAUX Agnès, bactériologie - virologie
4. BEN-BRIK Eric, médecine du travail
5. BILAN Frédéric, génétique
6. BOURMEYSTER Nicolas, biologie cellulaire
7. CASTEL Olivier, bactériologie - virologie - hygiène
8. CATEAU Estelle, parasitologie et mycologie
9. CREMNITER Julie, bactériologie - virologie
10. DAHYOT-FIZELIER Claire, anesthésiologie - réanimation
11. DIAZ Véronique, physiologie
12. FAVREAU Frédéric, biochimie et biologie moléculaire
13. FRASCA Denis, anesthésiologie - réanimation
14. HURET Jean-Loup, génétique
15. LAFAY Claire, pharmacologie clinique
16. SAPANET Michel, médecine légale
17. SCHNEIDER Fabrice, chirurgie vasculaire
18. THILLE Arnaud, réanimation
19. TOUGERON David, hépato-gastro-entérologie

Professeur des universités de médecine générale

GOMES DA CUNHA José

Professeurs associés de médecine générale

BINDER Philippe
VALETTE Thierry

Maîtres de Conférences associés de médecine générale

BIRAULT François
BOUSSAGEON Rémy
FRECHE Bernard
GIRARDEAU Stéphane
GRANDCOLIN Stéphanie
PARTHENAY Pascal
VICTOR-CHAPLET Valérie

Enseignants d'Anglais

DEBAIL Didier, professeur certifié
PERKINS Marguerite, maître de langue étrangère

Professeurs émérites

1. DORE Bertrand, urologie (08/2016)
2. FAUCHERE Jean-Louis, bactériologie – virologie (08/2015)
3. GIL Roger, neurologie (08/2017)
4. MAGNIN Guillaume, gynécologie-obstétrique (08/2016)
5. MARCELLI Daniel, pédopsychiatrie (08/2017)
6. MENU Paul, chirurgie thoracique et cardio-vasculaire (08/2017)

Professeurs et Maîtres de Conférences honoraires

1. ALCALAY Michel, rhumatologie
2. BABIN Michèle, anatomie et cytologie pathologiques
3. BABIN Philippe, anatomie et cytologie pathologiques
4. BARBIER Jacques, chirurgie générale (ex-émérite)
5. BARRIERE Michel, biochimie et biologie moléculaire
6. BECQ-GIRAUDON Bertrand, maladies infectieuses, maladies tropicales (ex-émérite)
7. BEGON François, biophysique, Médecine nucléaire
8. BOINOTCatherine, hématologie - transfusion
9. BONTOUX Daniel, rhumatologie (ex-émérite)
10. BURIN Pierre, histologie
11. CASTETS Monique, bactériologie -virologie – hygiène
12. CAVELLIER Jean-François, biophysique et médecine nucléaire
13. CHANSIGAUD Jean-Pierre, biologie du développement et de la reproduction
14. CLARAC Jean-Pierre, chirurgie orthopédique
15. DABAN Alain, cancérologie radiothérapie (ex-émérite)
16. DAGREGORIO Guy, chirurgie plastique et reconstructrice
17. DESMAREST Marie-Cécile, hématologie
18. DEMANGE Jean, cardiologie et maladies vasculaires
19. FONTANEL Jean-Pierre, Oto-Rhino Laryngologie (ex-émérite)
20. GOMBERT Jacques, biochimie
21. GRIGNON Bernadette, bactériologie
22. GUILLARD Olivier, biochimie et biologie moléculaire
23. JACQUEMIN Jean-Louis, parasitologie et mycologie médicale
24. KAMINA Pierre, anatomie (ex-émérite)
25. KLOSSEK Jean-Michel, Oto-Rhino-Laryngologie
26. LAPIERRE Françoise, neurochirurgie (ex-émérite)
27. LARSEN Christian-Jacques, biochimie et biologie moléculaire
28. MAIN de BOISSIERE Alain, pédiatrie
29. MARILLAUD Albert, physiologie
30. MORIN Michel, radiologie, imagerie médicale
31. POINTREAU Philippe, biochimie
32. REISS Daniel, biochimie
33. RIDEAU Yves, anatomie
34. SULTAN Yvette, hématologie et transfusion
35. TALLINEAU Claude, biochimie et biologie moléculaire
36. TANZER Joseph, hématologie et transfusion (ex-émérite)
37. VANDERMARCQ Guy, radiologie et imagerie médicale

Abréviations

ADVANCE : Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation study

AINS : anti-inflammatoire non stéroïdien

AIT : accident ischémique transitoire

AOMU : artériopathie oblitérante des membres inférieurs

ASC : aire sous la courbe

ATCD : antécédent

AVC : accident vasculaire cérébral

CHU : centre hospitalier universitaire

CKD-EPI : Chronic Kidney Disease Epidemiology Collaboration

CRP : protéine C réactive

CV : cardiovasculaire

DCS : Diabetes Cohort Study

DFGe : débit de filtration glomérulaire estimé

DIABHYCAR : non-insulin-dependent DIABetes, HYpertension, microalbuminuria or proteinuria, CARdiovascular events, and ramipril

DT2 : diabète de type 2

EUA : excrétion urinaire d'albumine

HbA1c : hémoglobine glyquée

HDL : high density lipoprotein

HTA : hypertension artérielle

IDM : infarctus du myocarde

IEC : inhibiteur de l'enzyme de conversion

IMC : indice de masse corporelle

IRC : insuffisance rénale chronique

IRCT : insuffisance rénale chronique terminale

LDL : low density lipoprotein

LED : lupus érythémateux disséminé

MRC : maladie rénale chronique

NIDDM : non insulin dependant diabetes mellitus

NZDCS : New Zealand Diabetes Cohort Study

OCDE : Organisation de Coopération et de Développement Économiques

OMS : organisation mondiale de la santé

PAD : pression artérielle diastolique

PAS : pression artérielle systolique

PI : prognostic index

PR : polyarthrite rhumatoïde

Rapport A/C : rapport albumine urinaire sur créatinine urinaire

RENAAL : Reduction of Endpoints in NIDDM with Angiotensin II Antogonist Losartan Study

ROC : receiver operating characteristic

SURDIAGNE : SUivi Renal, DIAbète de type 2 et GENEtique

UKPDS : United Kingdom Prospective Diabetes Study

Table des matières

1. Introduction 8
1.1 Le diabète de type 2 8
1.1.1 Définition 8
1.1.2 Epidémiologie 8
1.1.3 Complications 8
1.2 Aspects méthodologiques des modèles et des scores de risque 9
1.2.1 Elaboration et développement 9
1.2.2 Principes et méthodes de validation interne 10
1.2.3 Principes et typologie des études de validation externe 11
1.3 Contexte 12
1.4 Problématique 16
1.5 Objectifs 16
2. Matériels et méthodes 17
2.1 Critères de sélection des scores de risque 17
2.2 Présentation des modèles prédictifs et des scores de risque de l'étude 20
2.2.1 Etude NZDCS 20
2.2.2 Etude ADVANCE 21
2.3 Populations d'analyse 22
2.3.1 SURDIAGENE 22
2.3.2 DIABHYCAR 23
2.3.3 Principe de sélection 23
2.4 Définitions des évènements étudiés et des variables des scores de risque 23
2.4.1 Données cliniques 23
2.4.2 Données biologiques 25
2.5 Analyses statistiques 26
2.5.1 Evaluation de la similarité entre les populations de développement et de validation 26
2.5.2 Estimation du risque dans la population de validation 26
2.5.3 Evaluation des performances prédictives du score 27
3. Résultats 28
3.1 Sélection de la population d'analyse 28

3.2 Validation externe du score NZDCS 29
3.2.1 Populations d'étude 29
3.2.2 Caractéristiques à l'inclusion 30
3.2.3 Evènements rénaux au cours du suivi 33
3.2.4 Associations entre les paramètres du score et les évènements rénaux 33
3.2.5 Estimation du risque d'évènement prédit par le score 34
3.2.6 Evaluation des performances prédictives 36
3.3 Validation du score ADVANCE 41
3.3.1 Populations d'étude 41
3.3.2 Caractéristiques à l'inclusion 42
3.3.3 Evènements rénaux au cours du suivi 45
3.3.4 Associations entre les paramètres du score et les évènements rénaux 45
3.3.5 Estimation du risque d'évènement prédit par le score 46
3.3.6 Evaluation des performances prédictives 48
3.4 Comparaison des scores de risque rénal 52
4. Discussion 53
4.1 Principaux résultats de l'étude 53
4.2 Avantages et limites de l'étude 53
4.3 Mise en perspective des résultats de l'étude 56
4.4 Au-delà de la validation externe 58
4.5 Conclusion 59
5. Références 60
6. Résumé 69
7. Serment 70
Résumé 71

1. Introduction

1.1 Le diabète de type 2

1.1.1 Définition

Le diabète de type 2 (DT2) est la forme la plus fréquente de diabète. Il est défini par une glycémie supérieure à 1,26g/l (7,0mmol/l) après un jeûne de 8 heures et vérifiée à deux reprises ; ou par la présence de symptômes de diabète (polyurie, polydipsie, amaigrissement) associée à une glycémie supérieure ou égale à 2g/l (11,1 mmol/l) ; ou par une glycémie supérieure ou égale à 2g/l (11,1 mmol/l) 2 heures après une charge orale de 75g de glucose (critères de l'Organisation Mondiale pour la Santé, OMS).

1.1.2 Epidémiologie

Véritable pandémie au niveau mondial, on estime à 371 millions le nombre de sujets diabétiques adultes en 2012[1]. En 2014, la prévalence mondiale du diabète était estimée à 9% chez les adultes âgés de 18 ans et plus[2]. En 2009 en France, la prévalence du diabète traité était estimée à 4,4% de la population, soit 2,9 millions de personnes, se situant ainsi dans la moyenne par rapport à ses voisins européens[3]. En raison du vieillissement de la population générale, de l'allongement de l'espérance de vie des diabétiques et de l'augmentation de la fréquence de l'obésité, la prévalence du diabète va certainement continuer à progresser. L'OMS prévoit qu'en 2030, le diabète sera la septième cause de décès dans le monde[4].

1.1.3 Complications

Longtemps asymptomatique, le DT2 peut provoquer au cours de son évolution de graves complications microvasculaires (rétinopathie, néphropathie et neuropathie) et macrovasculaires (infarctus du myocarde, artérite et accident vasculaire cérébral).

Le diagnostic de néphropathie diabétique est suspecté cliniquement lorsque le diabète est relativement ancien (ancienneté supérieure à 10-15 ans) et qu'il est associé à une macroalbuminurie > 300 mg/24h ou à un débit de filtration glomérulaire estimé (DFGe) < 60

ml/min/1.73m2, en l'absence de signes rénaux ou extra rénaux orientant vers un autre type d'atteinte. Lors du diagnostic du DT2, on estime que 3% des patients présentent déjà une maladie rénale[5]. Aux Etats-Unis mais également dans de nombreux autres pays, la principale cause d'insuffisance rénale chronique (IRC) est la néphropathie diabétique[6]. On estime par ailleurs que 40% des patients DT2 présenteront une néphropathie diabétique[7].

1.2 Aspects méthodologiques des modèles et des scores de risque

1.2.1 Elaboration et développement[8,9]

Les scores pronostiques ont pour but d'évaluer la probabilité de survenue d'un évènement ultérieur dont la définition doit être objective et précise. Leur élaboration nécessite de mettre en évidence les caractéristiques individuelles cliniques et biologiques associées à l'évolution et l'apparition des complications d'une maladie, grâce à la construction d'un modèle multivarié correspondant à la meilleure combinaison de paramètres pronostiques de l'issue de santé étudiée à partir des données individuelles d'une population donnée. Les modèles sont le plus souvent développés de manière rétrospective à partir des données d'études de cohorte longitudinale dont les avantages sont un moindre coût et une réalisation plus simple pour des résultats pouvant être d'une grande puissance statistique ; les problématiques de cette méthodologie sont liées aux données manquantes et à l'hétérogénéité des définitions des critères de sélection, des paramètres et des évènements étudiés. Il existe des modèles développés prospectivement, pour lesquels les critères de sélection, les paramètres et les évènements étudiés sont définis de manières identiques pour tous les sujets ; les principales limites de cette méthodologie sont la moindre généralisation des résultats du fait de critères de sélection spécifiques, les coûts plus élevés et une mise en place plus complexe.

On regroupe sous les termes de modèle prédictif ou de modèle de prédiction ou de modèle pronostique, l'ensemble des paramètres accompagnés de leur coefficient de régression et/ou de leur risque relatif issus d'un modèle multivarié dont la variable à expliquer est l'évènement d'intérêt. Ces modèles peuvent être des modèles de régression logistiques lorsque la variable d'intérêt est binaire, ou bien des modèles de régression pour données de survie de type Cox ou Weibull lorsque la variable d'intérêt est censurée[10]. Après l'analyse univariée des associations entre les paramètres d'intérêt et l'évènement prédit, seules les variables significativement liées pour un seuil prédéfini (par exemple $p<0,2$) sont intégrées au modèle multivarié maximal. La

sélection des variables constituant le modèle final repose sur une procédure pas à pas descendante d'exclusion des variables non significatives au seuil généralement défini de 5%. Il est généralement recommandé de ne pas catégoriser les variables continues afin de perdre le moins d'information possible. La gestion des données manquantes est délicate et peut être source d'un manque de puissance si ces dernières ne sont pas correctement analysées. En effet, éliminer les sujets présentant des données manquantes pour les variables du modèle n'est pas la meilleure alternative puisqu'elle revient à ne prendre en compte que les sujets dont toutes les covariables sont disponibles, laissant supposer que l'existence des valeurs manquantes est due au hasard. Les données manquantes n'étant pas forcément de répartition aléatoire dans un jeu de données, différentes techniques d'imputation des données manquantes ont été développées[11,12].

Le score de risque est une transformation et une simplification du modèle afin de le rendre plus facile d'utilisation en pratique clinique courante. Il existe de nombreuses manières de construire un score depuis un modèle selon différentes méthodologies d'assignation de point pour les paramètres, parmi lesquelles : diviser le coefficient β par son erreur standard puis l'arrondir à l'entier le plus proche[13–15], multiplier le coefficient β par quatre puis l'arrondir à l'entier le plus proche[16], ou encore arrondir l'odds ratio à l'entier le plus proche[17]. Il est par ailleurs intéressant de constater que ces méthodes peuvent constituer une « boîte noire » statistique dans certains articles proposant des scores de risque[18,19]. Les scores de risque finaux proposés sont diverses, de la simplification extrême du modèle sous formes de système de points à sommer sur un total, à l'équation de risque brute comportant plusieurs lignes de calcul, on peut parfois trouver une table de simulation des probabilités de risque créée pour des patients-types.

1.2.2 Principes et méthodes de validation interne

Les coefficients des paramètres du modèle final sont généralement biaisés et les degrés de signification qui leur sont associés ne correspondent pas au risque d'erreur réel car le modèle a été élaboré et ses paramètres estimés sur un même jeu de données. La validité d'un ensemble de facteurs pronostiques dépend de sa reproductibilité sur un échantillon distinct de patients. La validation d'un score ou d'un modèle nécessite d'étudier la discrimination (capacité à discriminer les sujets présentant l'évènement de ceux ne le présentant pas) et la calibration (performance du score quel que soit le risque d'apparition de l'évènement). Il existe deux principales méthodes de validation interne qui sont des techniques utilisant les mêmes données

que celles ayant permis son élaboration. La méthode le plus souvent rencontrée est la validation croisée ou *split sample 2:1* qui consiste à scinder de manière aléatoire l'échantillon en deux sous-échantillons, un jeu de données pour le développement (par exemple, les deux tiers) et l'autre pour la validation (l'autre tiers). Cette méthode réduit cependant la puissance statistique et la méthode à privilégier est celle du rééchantillonnage ou *bootstrap* [20].

1.2.3 Principes et typologie des études de validation externe

La pertinence clinique d'un modèle de risque n'est pas seulement définie par de bonnes performances prédictives lors de son élaboration et de sa validation interne, une confirmation empirique de ses performances prédictives, en l'appliquant à une population différente de celle dans laquelle le modèle a été développé, reste nécessaire[21,22]. Un modèle prédictif est d'autant plus généralisable, et donc applicable en clinique courante, qu'il est performant dans un contexte éloigné de celui dans lequel il a été initialement développé[23]. Toutefois, la population dans laquelle va être réalisée la validation externe doit correspondre à une population cible comparable, autant à risque pour les évènements prédits que la population ayant permis l'élaboration du modèle[22]. Une validation externe n'est pas l'adaptation des coefficients des paramètres prédictifs du modèle testé (*model updating*) ni la recherche de nouveaux paramètres prédictifs de l'issue étudiée plus adaptés à la nouvelle population (*model derivation*) mais consiste en l'application du modèle final, avec ses paramètres prédictifs et leurs coefficients, ou du score de risque, dans une nouvelle population puis de la quantification de ses performances prédictives, calibration et discrimination. Il existe différents types d'étude de validation. La validation géographique permet l'étude fiable et robuste de la généralisation du modèle à une population de patients plus large, contrairement à la validation temporelle pour laquelle la population de validation est issue du même contexte géographique à un moment différent et permet donc seulement l'étude de la généralisation du modèle au cours du temps[20,23,24]. Les différences attendues entre la population de développement du modèle et la population de validation, en termes de caractéristiques à l'inclusion, de définition des facteurs prédictifs et des évènements étudiés, sont plus importantes dans une étude de validation externe géographique, à l'origine de l'important niveau de preuve de la généralisation du modèle ainsi validé[24]. Il est également suggéré qu'une validation externe indépendante, c'est-à-dire non conduite par les auteurs du modèle initial, pourrait diminuer le risque de biais dans les résultats[22].

1.3 Contexte

La maladie rénale chronique (MRC, tableau 1) est un problème de santé publique majeur, en progression dans de nombreux pays[25,26], à l'origine de graves complications telles que les maladies cardiovasculaires (CV) et l'insuffisance rénale chronique terminale (IRCT)[27].

Tableau 1: Classification de la maladie rénale chronique

Stade	DFG* (ml/min/1,73m²)	Définition
1	≥ 90	Maladie rénale chronique† avec DFG normal ou augmenté
2	Entre 60 et 89	Maladie rénale chronique† avec DFG légèrement diminué
3	Entre 30 et 59	Insuffisance rénale chronique modérée
4	Entre 15 et 29	Insuffisance rénale chronique sévère
5	< 15	Insuffisance rénale chronique terminale

*Débit de filtration glomérulaire

†Avec marqueurs d'atteinte rénale : protéinurie clinique, hématurie, leucocyturie, ou anomalies morphologiques ou histologiques, ou marqueurs de dysfonction tubulaire, persistant depuis plus de trois mois.

L'élaboration de stratégies permettant aux professionnels de santé l'identification précoce des patients à risque de MRC est une priorité. Un grand nombre de modèles prédictifs du risque d'apparition ou de progression de la MRC ont ainsi été développés[28] (tableaux 2a et 2b).

La plupart des scores prédictifs du risque rénal élaborés en population générale sont plus souvent prédictifs de la présence de MRC alors que les modèles développés dans populations à risque plus ciblées sont plus généralement prédictifs de la progression de la MRC [28]. Les événements considérés dans les modèles de prédiction du risque rénal varient selon le stade de MRC étudié et pour définir l'apparition d'une IRCT, l'évènement le plus souvent pris en compte est la diminution du DFGe en dessous du seuil de 60 ml/min/1,73m² (stades 3 à 5) et la présence d'une néphropathie avec une albuminurie telle que le rapport albuminurie sur créatinurie (rapport A/C) soit ≥ 30mg/g ou ≥ 3mg/mmol[29]. La progression de la MRC est souvent étudiée par les évènements suivants : doublement de la créatinine sérique (≥ 200 µmol/l[28] ou doublement de la créatinine initiale pendant au moins un mois[30]), IRCT définie telle qu'une IRC nécessitant une méthode de suppléance rénale (la mise sous dialyse ou la transplantation rénal), et le décès de cause rénale[31].

Parmi les facteurs de risque reconnus de progression de la MRC, le diabète, dont la prévalence est en constante augmentation, constitue la principale cause d'IRCT aux Etats-Unis et dans de

nombreux pays[26,32]. Ainsi, le diabète est le facteur de risque le plus fréquemment intégré aux modèles prédictifs du risque rénal retrouvés dans la littérature[28]. Un certain nombre de ces modèles ont d'ailleurs été développés dans des populations de patients DT2: the Reduction of Endpoints in NIDDM with Angiotensin II Antogonist Losartan (RENAAL) Study[33], the Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation Study (ADVANCE)[34], the United Kingdom Prospective Diabetes Study (UKPDS)[35], the Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT) End Stage Renal Disease model[36], et the New Zealand (NZ) Diabetes Cohort Study (DCS)[37]. Ces scores de risque ont été principalement conçus pour prédire le risque de complications de MRC, telles que la survenue d'IRCT, fatale ou non. Bien que ces modèles aient été élaborés selon la méthodologie consensuelle comportant les étapes nécessaires de validation interne et d'évaluation des performances prédictives[38], leur validité clinique ne peut être que renforcée en les appliquant dans d'autres contextes populationnels[23].

Tableau 2a : Modèles de prédiction d'apparition de la maladie rénale chronique

Etude	Facteurs prédictifs inclus dans le modèle	Evènements rénaux prédits
Bang et al. 2007 – *score SCORED*[39]	Age, sexe, anémie, HTA, diabète, ATCD CV, AOMI	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Kshirsagar et al. 2008 – *score ARIC/CHS*[40]	Age, sexe, anémie, HTA, diabète, ATCD CV, AOMI	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Fox et al. 2010 – *score FRAMINGHAM 3*[41]	Age, sexe, PAS, HTA, traitement HTA, tabac, IMC, HDL, diabète, DFGe, aldostérone, homocystéine	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Hippisley Cox et al. 2010 – *score QKidney*[42]	Age, ethnie, déprivation sociale, tabac, IMC, PAS, diabète, PR, ATCD CV, traitement HTA, AOMI, LED, ATCD familiaux néphropathie, ATCD calculs rénaux, AINS (stratification du modèle sur le sexe)	MRC stades 3b-5 (DFGe < 45 ml/min/1,73m² ou transplantation, dialyse, néphropathie, protéinurie persistante) ; MRC stade 5 (DFGe < 15 ml/min/1,73m² ou transplantation, dialyse, néphropathie, protéinurie persistante)
Chien et al. 2010 – *score TAIWAN 1*[43]	Age, IMC, PAD, diabète de type 2, ATCD AVC	MRC stade 2 (DFGe diminué mais ≥ 60 ml/min/1,73m²)
Halbesma et al. 2011 – *score PREVEND*[44]	Age, DFGe, EUA, CRP, PAS, HTA	MRC stades 3-5 (pente annuelle de déclin de la fonction rénale et DFGe < 60 ml/min/1,73m²)
Ando et al. 2011 – *score JAPAN/HIV*[45]	Age, taux de lymphocytes T CD4, diabète, protéinurie, DFGe	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Thakkinstian et al. 2011 – *score THAILANDE*[46]	Age, diabète, HTA, ATCD calculs rénaux	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
O'Seaghdha et al. 2012 – *score FRAMINGHAM 4c*[47]	Age, HTA, diabète, DFGe, albuminurie	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Alssema et al. 2012 – *score ROTTERDAM-HOORN*[48]	Age, IMC, tour de taille, traitement HTA, tabac, ATCD familiaux CV, ATCD familiaux diabète	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)
Kwon et al. 2012 – *score KOREAN*[49]	Age, sexe, anémie, HTA, diabète, ATCD CV, protéinurie	MRC stades 3-5 (DFGe < 60 ml/min/1,73m²)

Tableau 2b : Modèles de prédiction de progression de la maladie rénale chronique

Etude	Facteurs prédictifs inclus dans le modèle	Evènements rénaux prédits
Keane et al. 2006 – *score RENAAL*[33]	Rapport A/C, albumine sérique, créatinine sérique, hémoglobine, HbA1c	IRCT (dialyse ou transplantation), décès de cause rénale
Kent et al. 2007[50]	Age, sexe, créatininémie, EUA, PAS	Doublement de la créatinine sérique initiale, IRCT (dialyse)
Wakai et al. 2006[51]	Age, sexe, PAS, hématurie, albuminémie, grade néphropathie à IgA	IRCT (dialyse)
Goto et al. 2009[52]	Age, sexe, PAS, protéinurie, hématurie, albumine, DFGe, grade néphropathie à IgA	IRCT (dialyse)
Johnson et al. 2008[53]	Age, sexe, DFGe, diabète, anémie, HTA	IRCT (dialyse ou transplantation rénale)
Landray et al. 2010[54]	Créatinémie, phosphates, rapport A/C, sexe, NT-pro-BNP, âge, tabac, troponines	IRCT (dialyse ou transplantation), décès de cause rénale
Hallan et al. 2009[55]	Age, sexe, activité physique, diabète, PAS, HTA, HDL, DFG, rapport A/C	IRCT (dialyse ou transplantation)
Tangri et al. 2011 – *score CANADA*[56]	Age, sexe, DFGe, albuminurie, calcium, phosphates, bicarbonate et albumine	IRCT (dialyse ou transplantation)
Desai et al. 2011[36]	Age, sexe, ethnie, IMC, insuline, DFGe, nitrites, rapport A/C, hémoglobine, ATCD CV, ferritinémie, CRP, troponines, NT-pro-BNP	IRCT (dialyse ou transplantation), décès de cause rénale
Jardine et al. 2012 – *score ADVANCE*[34]	Sexe, DFGe, rapport A/C, PAS, HbA1c, rétinopathie diabétique, âge à la fin du secondaire	Doublement de la créatinine sérique ≥ 200µol/l, IRCT (dialyse ou transplantation), décès de cause rénale
Hayes et al. 2013 – *UKPDS OMs 2*[35]	Ethnie, âge, sexe, IMC, DFGe, hémoglobine, LDL, leucocytes, albuminurie, PAS, ATCD amputation, ATCD cécité	Créatinine sérique ≥ 250µmol/l, décès de cause rénale
Elley et al. 2013 – *score NZDCS*[37]	Sexe, ethnie, âge, durée du diabète, albuminurie, créatinémie, PAS, HbA1c, tabac, ATCD CV	IRCT (dialyse ou transplantation), décès de cause rénale

Abréviations : maladie rénale chronique, MRC; insuffisance rénale chronique terminale, IRCT; rapport albuminurie sur créatinurie, rapport A/C; débit de filtration glomérulaire estimé, DFGe; excrétion urinaire d'albumine, EUA; pression artérielle systolique, PAS; pression artérielle diastolique, PAD; antécédent, ATCD; cardiovasculaire, CV; artériopathie oblitérante des membres inférieurs, AOMI; hypertension artérielle, HTA; accident vasculaire cérébral, AVC; indice de masse corporelle, IMC; hémoglobine glyquée, HbA1c; high density lipoprotéine, HDL; low density lipoprotéine, LDL; protéine C réactive, CRP; polyarthrite rhumatoïde, PR; lupus érythémateux disséminé, LED; anti-inflammatoire non stéroïdien, AINS ; NT-pro-BNP, N-terminal pro-brain natriuretic peptide.

1.4 Problématique

Devant la multitude d'études concernant le développement de nouveaux modèles de prédiction du risque, élaborés dans autant de contextes différents, l'alternative méthodologique recommandée actuellement est d'effectuer une adaptation des modèles préexistants suite à leur validation externe, plutôt que de développer des modèles supplémentaires[23]. Afin de limiter la publication exponentielle de nouveaux modèles prédictifs, certains auteurs émettent l'idée qu'éditeurs et *reviewers* devraient être plus exigeants quant à la justification de développer de nouveaux modèles et qu'ils devraient plutôt encourager la recherche visant l'évaluation des modèles existants[22].

A ce jour, les études de validation externe concernent principalement les maladies cardiovasculaires et la cancérologie, seul un nombre limité de scores de risque rénal a fait l'objet d'une telle validation[28]. En l'état actuel des connaissances, les scores ayant fait l'objet d'une validation externe dans des populations de patients diabétiques de type 2[57–60] sont des modèles prédictifs du risque cardiovasculaire et aucune étude n'a encore ciblé une validation externe indépendante de modèles prédictifs du risque rénal chez des patients DT2.

1.5 Objectifs

L'objectif principal de cette étude était de réaliser la validation externe indépendante de deux scores de risque rénal, le Diabetes Cohort Study (DCS) 5 -year Renal Risk Model[37] et le Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation Study (ADVANCE) 5 -year Renal Risk Model[34], chez les patients DT2 des études SURDIAGENE[30] et DIABHYCAR[61]. L'objectif secondaire était de comparer la calibration et la discrimination des deux scores dans chacune des populations de validation afin de déterminer le score le plus performant.

2. Matériels et méthodes

2.1 Critères de sélection des scores de risque

Afin de réaliser un état des lieux des modèles de prédiction du risque rénal publiés, une revue de la littérature a été effectuée ; le processus de sélection des scores de risque rénal inclus dans l'étude est présenté dans la figure 1. La stratégie consistait à rechercher dans la base documentaire PubMed MEDLINE, les articles princeps et les revues de la littérature publiés entre 2006 et 2014, de préférence en langue anglaise, par différentes combinaisons de termes concernant les éléments cliniques « chronic kidney disease », « end stage renal disease », « renal disease », « type 2 diabetes », et des termes d'ordre méthodologique « prediction model », « predicting », « predict », « risk score », « risk assessment », « scoring », « model developpment ». Les titres, puis les résumés et enfin articles complets (*full text*) étaient revus et retenus lorsqu'ils semblaient pertinents et publiés dans des revues de qualité. Si le contenu d'un article n'était pas disponible, il était récupéré ainsi que leurs données supplémentaires (*supplmentary files*) le cas échéant. Des points d'étapes de la recherche bibliographique étaient régulièrement organisés afin d'en évaluer la pertinence par deux évaluateurs (S.H, A.d.H) lors de réunions scientifiques. Les références bibliographiques des travaux sélectionnés ont été également revues manuellement et cela pouvaient conduire à la sélection d'un nouvel article non retrouvé par la première recherche dans la base documentaire.

Parmi les scores de risque rénal sélectionnés, les modèles développés au sein de populations DT2 suivants étaient éligibles pour l'étude : the Reduction of Endpoints in NIDDM with Angiotensin II Antogonist Losartan (RENAAL) Study[33], the Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation Study (ADVANCE)[34], the Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT) End Stage Renal Disease model[36], the United Kingdom Prospective Diabetes Study (UKPDS)[35], et the New Zealand (NZ) Diabetes Cohort Study (DCS)[37]. En effet, les modèles de prédiction élaborés en population générale n'étudiaient pas les évènements rénaux pertinents à rechercher chez des patients DT2, une population à risque plus élevé d'évènements rénaux. De plus, les évènements rénaux modélisés étaient le plus souvent liés à l'apparition de maladie rénale plutôt qu'à son aggravation ou sa chronicité (tableaux 2a et 2b). Les scores finalement inclus pour cette étude devaient proposer des équations de risque pour lesquelles les variables étaient disponibles dans les populations SURDIAGENE et DIABHYCAR mais aussi présenter des définitions des

évènements prédits et des facteurs prédictifs identiques ou très proches à celles de ces deux populations. Enfin, les articles publiés devaient également avoir défini un score de risque issu de leur modèle. Le tableau 3 résume les principales caractéristiques des deux modèles NZDCS et ADVANCE retenus pour l'analyse en suivant les critères de sélection précédents.

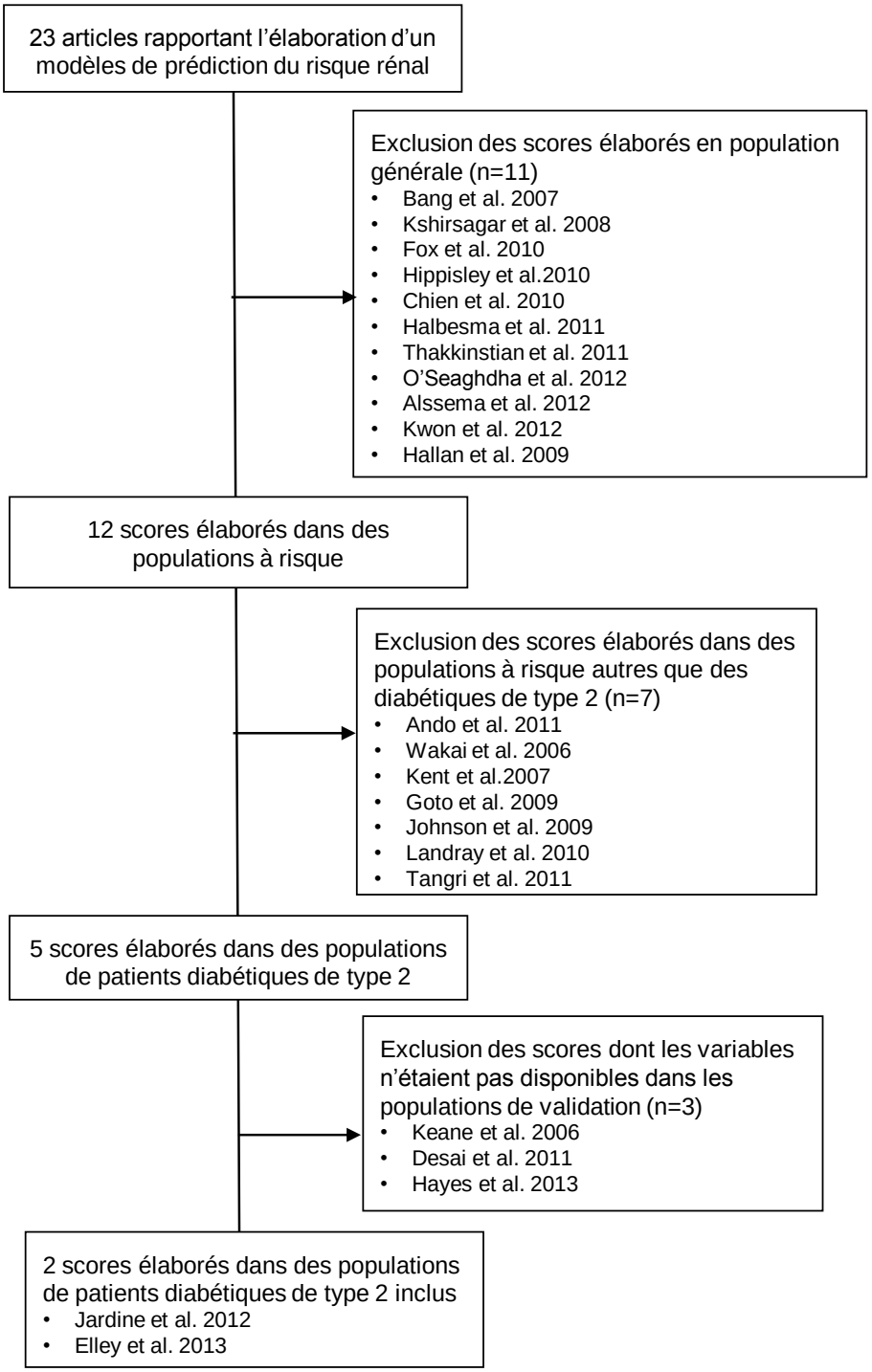

Figure 1 : Processus de sélection des scores de prédiction du risque rénal.

Tableau 3 : Principales caractéristiques des modèles de prédiction du risque rénal retenus dans l'étude

Etude	Evénements rénaux	Taille de l'échantillon	Type d'étude	Modèle statistique	Facteurs prédictifs	Performances prédictives †		Type de validation	Nombre d'évènements
						Discrimination	Calibration		
ADVANCE 5 -year Risk Model[34]	DCrS ou IRCT ou mort rénale	11140	Essai Clinique	Cox	Sexe, DGFe, rapport A/C, PAS, HbA1c, rétinopathie diabétique, âge à la fin du premier cycle de l'enseignement secondaire‡	0,847 (0,815-0,880)	P=0.9	Validation interne (*bootstrap sampling*)	188 dont 129 DCrS et 59 IRCT ou morts rénales
DCS 5 -year Renal Risk Model[37]	IRCT ou mort rénale	25736	Cohorte	Cox	Sexe, ethnie, âge au diagnostic du DT2, durée du DT2, CrS, albuminurie, PAS, HbA1c, ATCD MCV	0,914 (0,881-0,937)	P=0,663	Validations interne et externe	637 dont 369 morts rénales et 268 IRCT

Abréviations : créatinine sérique, CrS ; doublement de la créatinine sérique, DCrS ; insuffisance rénale terminale, IRT ; débit de filtration glomérulaire estimé, DFGe ; rapport albuminurie/créatinurie, rapport A/C ; pression artérielle systolique, PAS ; hémoglobine glyquée, HbA1c ; antécédent, ATCD ; maladie cardio-vasculaire, MCV ; diabète de type 2, DT2 ; non communiquée, NC.

†Performances prédictives du modèle lors de son développement exprimée par la statistique C (IC 95%) et le degrés de significativité du test du Chi ^{2} de Hosmer et Lemeshow pour données de survie .

‡ Correspond à la classe de 3e.

2.2 Présentation des modèles prédictifs et des scores de risque de l'étude

2.2.1 Etude NZDCS

Le modèle de prédiction du risque rénal de l'étude DCS était un modèle de Cox exprimant l'apparition dans les 5 ans d'une IRCT ou d'un décès d'origine rénale en fonction des 11 paramètres suivants : sexe, ethnie, créatininémie, DFGe, âge au diagnostic du diabète, durée du diabète, albuminurie, PAS, HbA1c, tabac et ATCD CV (tableau 4).

Tableau 4 : Modèle de prédiction du risque d'insuffisance rénale chronique terminale de l'étude NZDCS

Variables	Hazard ratios ajustés
Sexe (Femme)	1,49
Age au diagnostic du DT2 (années)	1,02
Ethnie	
Caucasien	1,00
Maori	2,27
Pacifique	1,34
Asie de l'Est	0,87*
Indo-asiatique	0,96*
Autre	0,8*
Durée du DT2 (années)	1,04
Créatinine sérique (pour 10 mg/dl)	1,36
Albuminurie †	
Normoalbuminurie	1,00
Microalbuminurie	1,90
Macroalbuminurie	3,68
Albuminurie avancée	9,88
HbA1c (%)	1,19
PAS (pour 10 mmHg)	1,05
Tabac	*
Jamais	1,00
Ancien	1,00
Actif	1,29
Antécédents cardiovasculaires	1,65

Abréviations : DT2, diabète de type 2; HbA1c, hémoglobine glyquée; PAS, pression artérielle systolique.
† Une microalbuminurie comme un rapport A/C ≥ 3 mg/mmol et < 30 mg/mmol, une macroalbuminurie un rapport A/C ≥ 30 mg/mmol et < 100 mg/mmol et une albuminurie avancée un rapport A/C) ≥ 100 mg/mmol. Pour la population ADVANCE, une macroalbuminuie est définie comme un rapport A/C ≥ 30 mg/mmol et la classe "avancée" n'a pas été définie.
* p>0,05.

L'équation du score de risque à 5 ans proposée était :

Score = 1-0,99558399^exp [(Femme*0,401069) + (Ethnie « Maori »*0,821956) + (Ethnie « Pacific »*0,296149) + (Ethnie « East Asian »*-0,138466) + (Ethnie «Indo-Asian »*-0,040645) + (Ethnie «autre »*-0,223648) + (Age au diagnostic du DT2*0,024199) + (Durée du DT2*0,043467) + (Créatinémie (µmol/L)/10 *0,305281) + (Microalbuminurie*0,644200) + (Macroalbuminurie*1,303188)+ (Albuminurie avancée* 2,290089) + (PAS (mmHg)/10*0,049809) + (HbA1c(%)/10*0,158597) + (Ancien fumeur*-0,003952) + (Fumeur actuel*0,252272) + (Antécédent CV*0,498804) - 7,0671736].

Dans ce modèle DCS, une microalbuminurie était définie comme un rapport A/C ≥ 3 mg/mmol et < 30 mg/mmol, une macroalbuminurie un rapport albuminurie sur créatinurie A/C ≥ 30 mg/mmol et < 100 mg/mmol et une albuminurie avancée un rapport A/C ≥ 100 mg/mmol.

2.2.2 Etude ADVANCE

Le modèle de prédiction du risque rénal de l'étude ADVANCE était un modèle de Cox décrivant l'apparition dans les 5 ans d'un doublement de la créatinine, d'une IRCT ou d'un décès d'origine rénale en fonction des 8 paramètres suivants : sexe, ethnie, DFGe, rapport A/C, PAS, HbA1c, présence d'une rétinopathie diabétique et âge à la fin du premier cycle de l'enseignement secondaire (tableau 5).

L'équation du score de risque à 5 ans proposée était :

Score = 1 - 0,99407^exp [0,89405 (si homme) - 0,04873* DFGe (ml/min/1,73 mm^2) + 0,44790*log rapport A/C (µg/mg) + 0,01163*PAS (mmHg) + 0,11618*HbA1c (%) + 0,42146 (si rétinopathie diabétique) + 0,41098 (si âge à la fin du premier cycle de l'enseignement secondaire ≤15 ans) - 1,13663],

Tableau 5 : Modèle de prédiction du risque d'insuffisance rénale chronique terminale de l'étude ADVANCE

Variables	Hazard ratios ajustés
Sexe (Homme)	2,45
DFGe (par 10 ml/min/1,73m²)	0,61
Rapport A/C (log µg/mg)	1,57
HbA1c (%)	1,12
PAS (par 10 mmHg)	1,12
Rétinopathie diabétique	1,52
Age à la fin du premier cycle de l'enseignement secondaire (≤15 ans vs. ≥16 ans)	1,51

Abréviations : DFGe, Débit de filtration glomérulaire estimé; Rapport A/C, rapport albuminurie sur créatinurie ; HbA1c, hémoglobine glyquée; PAS, pression artérielle systolique.

2.3 Populations d'analyse

2.3.1 SURDIAGENE

SUivi Renal, DIAbete de type 2 et GENEtique (SURDIAGENE) est une étude de cohorte prospective monocentrique dont l'objectif était d'identifier les déterminants génétiques et environnementaux contribuant aux complications microvasculaires et macrovasculaires du DT2[30]. Les patients DT2 ont été inclus et suivis de manière régulière lors de consultations, hospitalisations de jour, hospitalisations traditionnelles, dans les services de diabétologie du Centre Hospitalier Universitaire (CHU) de Poitiers, France, entre 2002 et 2011. Les critères de diagnostic de diabète de type 2 étaient les suivants[63]: un âge au diagnostic supérieur ou égal à 35 ans, une insulino-requérance apparue plus de 2 ans après le diagnostic de diabète, l'absence de diabète secondaire et l'absence de cétonurie au moment du diagnostic. Les critères d'exclusion étaient la présence d'un diabète de type 1 ou d'un diabète secondaire, ou la présence d'une néphropathie évoquant une atteinte rénale non diabétique, ou une durée de suivi inférieure à 1 an, ou un lieu de résidence en dehors de la région de Poitiers. Pour cette étude, les patients présentant une épuration extra-rénale à l'inclusion ont été exclus. L'étude a été approuvée par le Comité de Protection des Personnes Ouest 3 et le consentement écrit a été recueilli pour chaque patient.

2.3.2 DIABHYCAR

L'étude DIABHYCAR (non-insulin-dependent DIABetes, HYpertension, microalbuminuria or proteinuria, CARdiovascular events, and ramipril) est un essai contrôlé randomisé versus placebo multicentrique en double aveugle testant l'effet de 1,25 mg de ramipril sur la morbi-mortalité cardiovasculaire et rénale[61]. Les patients ont été inclus entre 1995 et 1998 dans 16 pays d'Europe et d'Afrique du Nord. Les patients éligibles devaient présenter un diabète de type 2 (traité par au moins un antidiabétique oral), être âgé d'au moins 50 ans et avoir une concentration urinaire d'albumine supérieure ou égale à 20 mg/l (sur deux échantillons successifs). Les critères d'exclusion étaient une créatinine sérique supérieure à 150 µmol/l; un traitement par insuline ou inhibiteur de l'enzyme de conversion (IEC) ou antagoniste des récepteurs de l'angiotensine II; des antécédents d'insuffisance cardiaque congestive chronique ou d'infarctus de myocarde (IDM) dans les trois derniers mois; une infection urinaire en cours; allergie ou intolérance connue aux IEC. Le protocole de l'étude a été approuvé par le comité d'éthique de CHU d'Angers et le consentement écrit a été recueilli pour chaque patient. Seules les données concernant les patients inclus en France étaient disponibles pour les analyses.

2.3.3 Principe de sélection

Dans la population de patients considérés pour les analyses, les sujets ayant fait l'évènement d'intérêt dans les 5 ans suivant l'inclusion étaient considérés comme des « cas » et les patients n'ayant pas fait l'évènement étudié dans les 5 ans suivant l'inclusion étaient considérés comme des « témoins ». Les analyses ont pris en compte les patients SURDIAGENE et DIABHYCAR pour lesquels il n'y avait pas de données manquantes pour les variables intégrées aux scores de risque rénal étudiés.

2.4 Définitions des évènements étudiés et des variables des scores de risque

2.4.1 Données cliniques

Les définitions des événements rénaux d'intérêt survenant au cours du suivi étaient identiques pour les deux populations d'étude :

- IRCT définie comme la nécessité d'un traitement de suppléance rénale par greffe rénale ou dialyse,
- Doublement de la créatinine défini comme un doublement de la créatinine initiale pendant au moins un mois,
- Mort d'origine rénale définie selon les codes de l'International Classification of Diseases version 9 (ICD-9) ≥ 580 et ≤ 593.9.

Dans la cohorte SURDIAGENE, un comité d'adjudication composé d'experts a validé ces évènements de manière indépendante tous les 2 ans depuis 2007 ; dans l'étude DIABHYCAR, l'adjudication des évènements était réalisée par un comité d'expert dédié (*Critical Event Commitee*) dont la fréquence des réunions dépendait du nombre d'évènements à valider.

Les données cliniques présentes dans les équations de risque comprenaient :

- le sexe,
- l'âge,
- la pression artérielle,
- l'ancienneté du diabète,
- la consommation tabagique,
- la rétinopathie diabétique,
- les antécédents cardio-vasculaires
- la durée du diabète,
- l'âge au diagnostic du diabète,
- l'ethnie
- le tabac

Dans l'étude DIABHYCAR, la rétinopathie diabétique était définie par la présence d'antécédent de photocoagulation laser pour rétinopathie diabétique.

Dans la cohorte SURDIAGENE, les antécédents cardio-vasculaires comprenaient les atteintes suivantes :

- une atteinte coronaire (IDM, revascularisation coronarienne par angioplastie ou pontage),
- une atteinte carotidienne définie par un acte de revascularisation des vaisseaux à destinée cérébrale,

- une anomalie neurologique focale, entrainant le décès ou de durées supérieures à 24 heures (Accident Vasculaire Cérébral, AVC) ou résolutives en moins de 24 heures (Accident Ischémique Transitoire, AIT),
- une artériopathie des membres inférieurs (amputation au-delà du métatarsophalangien ou un acte de revascularisation).

Dans l'étude DIABHYCAR, les antécédents cardio-vasculaires regroupaient IDM, angor d'effort, accidents vasculaires cérébraux ou artériopathie des membres inférieurs.

La variable « âge à la fin du premier cycle de l'enseignement secondaire» du modèle ADVANCE n'était pas disponible pour les populations d'étude et a été fixée à 16 ans pour l'analyse.

2.4.2 Données biologiques

Les données biologiques mesurées ou calculées présentes dans les scores étudiés et disponibles dans les deux populations étaient:

- l'hémoglobine glyquée (HbA1c, méthode BioRad),
- la créatinine sérique (méthode colorimétrique),
- l'excrétion urinaire d'albumine (EUA, méthode néphélémétrie),
- et le débit de filtration glomérulaire estimé (DFGe) par la formule du Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI)[64].

Le rapport A/C a été directement calculé pour les patients de la cohorte SURDIAGENE car les variables nécessaires au calcul étaient disponibles (albuminurie et créatinurie). Pour les patients de l'étude DIABHYCAR, seule la concentration urinaire d'albumine était disponible (mg/l), les équivalences utilisées afin d'obtenir l'EUA (mg/jour) puis le rapport A/C (mg/g) étaient[29,65]:

- équivalence entre la concentration urinaire d'albumine et l'EUA : *concentration urinaire d'albumine de 20 mg/l = EUA de 30 mg/jour*
- équivalence entre l'EUA et le rapport A/C : *rapport A/C en mg/g = EUA en mg/jour*
- d'où l'équivalence : *concentration urinaire d'albumine de 20 mg/l = rapport A/C de 30 mg/g (ou 3 mg/mmol*).

2.5 Analyses statistiques

Cette étude de validation externe indépendante rétrospective a été conduite en tenant compte du niveau d'information disponible concernant l'élaboration des modèles et des scores dans les publications princeps, selon les principes de validation externe issus de la littérature[20–24,66–68]. L'ensemble des étapes de la validation externe effectuée sont décrites ci-dessous.

2.5.1 Evaluation de la similarité entre les populations de développement et de validation

Cette étape a permis de décrire de manière globale le degré de ressemblance, ou au contraire de dissemblance, des populations de développement et de validation d'un score. Une première approche a consisté en la description de la distribution des principales caractéristiques des patients des populations, des variables constituant le score et des fréquences de survenue des évènements rénaux d'intérêt (description du *case mix* soit l'éventail ou spectre des cas) . Une autre approche a consisté en la description de l'hétérogénéité des associations entre les évènements rénaux d'intérêt et les facteurs pronostiques des scores étudiés entre les populations de validation afin d'évaluer si les facteurs pronostiques des scores étaient également des facteurs associés aux évènements rénaux dans les populations de validation. Cette approche a permis d'identifier les facteurs de risque et les facteurs protecteurs d'évènements rénaux communs entre ces populations. Les associations univariées entre les paramètres des scores et les évènements rénaux ont été recherchées par régression logistique pour les paramètres qualitatifs et par régression linéaire pour les paramètres quantitatifs.

2.5.2 Estimation du risque dans la population de validation

L'étape préalable et nécessaire à l'analyse des performances prédictives d'un score résultait en l'application de l'équation de prédiction de risque à chaque sujet afin d'estimer le risque individuel d'apparition d'évènement. La valeur obtenue était une prédiction du risque d'apparition de l'évènement à un horizon de temps fixé à 5 ans. Une fois le risque d'apparition de l'évènement estimé pour les sujets de la population de validation, des groupes de risque ont été créés afin de faciliter la comparaison entre risque observé et risque estimé par le score. Les groupes de risque ont été créés selon les déciles de risque estimé[20].

2.5.3 Evaluation des performances prédictives du score

Calibration

Le graphique de calibration (*calibration plot)* présentait la proportion d'évènements observés en ordonnées (valeurs comprises entre 0 et 1) en fonction du risque prédit de survenue de l'évènement (valeurs comprises entre 0 et 1) selon le score en abscisses. Cette méthode permettait de visualiser l'adéquation du score dans la nouvelle population grâce à la position des points par rapport à la diagonale de 45°qui représentait la calibration parfaite.

Le test du Chi 2 de Hosmer et Lemeshow a permis d'évaluer statistiquement la calibration en vérifiant si la somme des différences entre nombre de cas observés et prédits dans les n groupes de risque considérés suit une loi du Chi 2 à n-2 degrés de liberté. Un résultat non significatif indiquait que les nombres de cas observés et prédits ne différaient pas significativement et que la calibration était donc satisfaisante. Au contraire, un résultat significatif traduisait une plus faible calibration[69].

Discrimination

La statistique de concordance ou statistique C (*C-index, C statistic*) a permis d'évaluer la discrimination en évaluant la proportion de paires d'individus (paire constituée d'un sujet avec évènement et d'un sujet sans évènement) pour lesquelles la prédiction du risque de survenue est plus élevée pour le sujet avec l'évènement que pour le sujet sans l'évènement d'intérêt. Dans le cas d'une variable binaire, cette statistique correspondait à l'aire sous la courbe (ASC) ROC (*receiver operating characterisctic*) du modèle $logit\ (Y=1) = a + b*logit\ (\tilde{Y})$, soit la probabilité de survenue de l'évènement Y en fonction de sa prédiction $\tilde{Y}$ donnée par le score de risque[70]. La courbe ROC illustrait la sensibilité du score (taux de vrais positifs) en fonction de sa spécificité (taux de faux positifs) pour des seuils consécutifs de risque prédit. L'ASC ROC devait être comprise entre 0.5 (pas de discrimination) et 1 (discrimination parfaite) et son intervalle de confiance à 95% ne devait pas comprendre 0.5 pour être significatif[71].

Les analyses ont été effectuées avec le logiciel SAS 9.3 (Statistical Analysis System Institute Inc.).

3. Résultats

3.1 Sélection de la population d'analyse

L'évaluation des scores a nécessité une prise en compte des temps de suivi et des délais d'apparition de l'évènement rénal étudié spécifique, permettant de connaître de manière certaine le statut d'un patient vis-à-vis de la survenue de l'évènement rénal étudié dans la période des 5 années suivant l'inclusion. Les patients ayant fait l'évènement rénal d'intérêt dans les 5 premières années de suivi étaient considérés comme des cas pour l'analyse, les patients n'ayant pas fait l'évènement dans les 5 premières années de suivi étaient considérés comme des témoins pour l'analyse, et les patients ayant un suivi inférieur à 5 ans ont été exclus de l'analyse. Le principe de sélection des patients est présenté dans la figure 2.

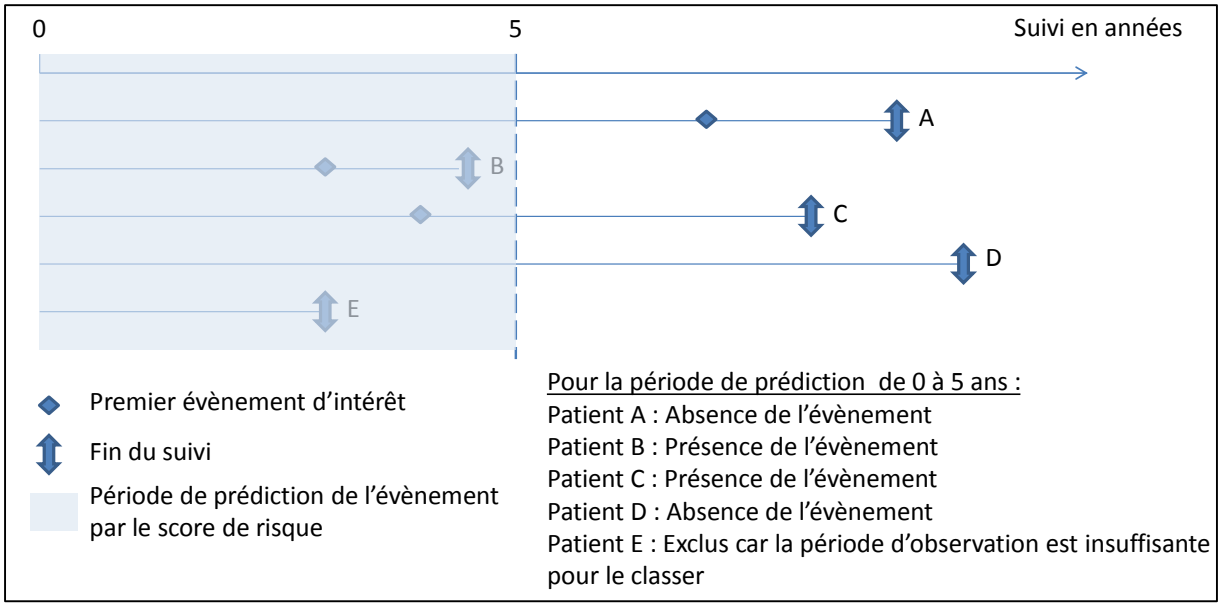

Figure 2 : Principe de sélection de la population d'analyse pour l'évaluation d'un score de risque à 5 ans.

3.2 Validation externe du score NZDCS

3.2.1 Populations d'étude

Les populations considérées pour la validation externe du score NZDCS issue des études SURDIAGENE et DIADHYCAR pour lesquelles le temps de suivi était suffisant, et qui ne présentaient pas de données manquantes pour les variables intégrées au score consistaient en 827 et 1302 patients respectivement (figures 3 et 4).

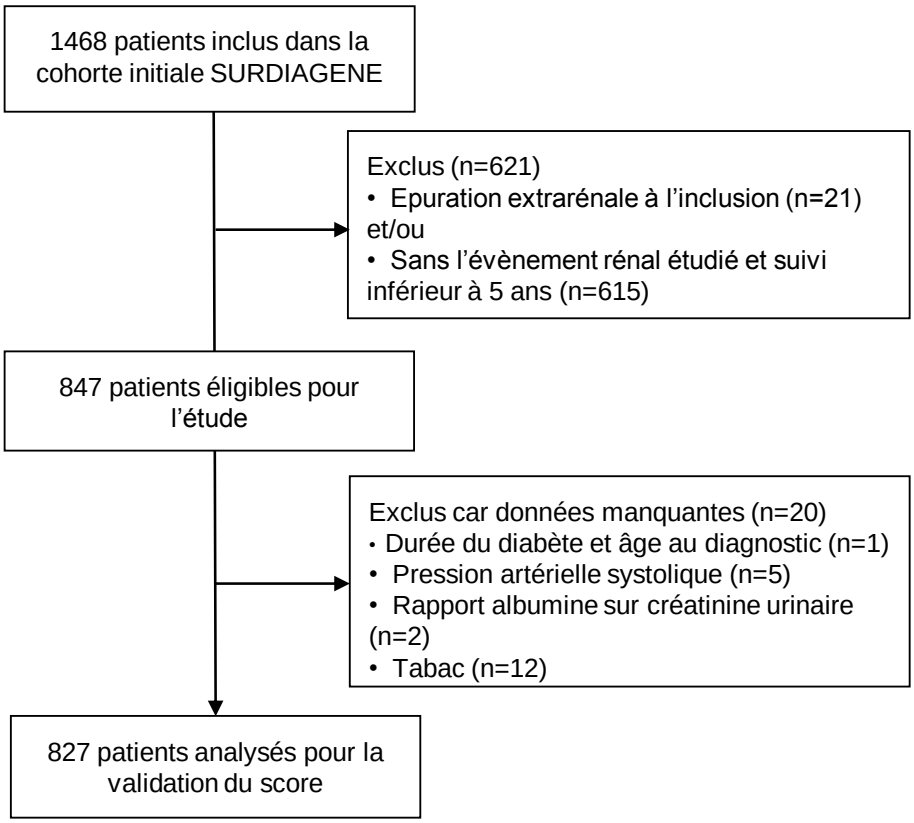

Figure 3 : Diagramme de flux des patients de la cohorte SURDIAGENE considérés pour la validation externe du score NZDCS.

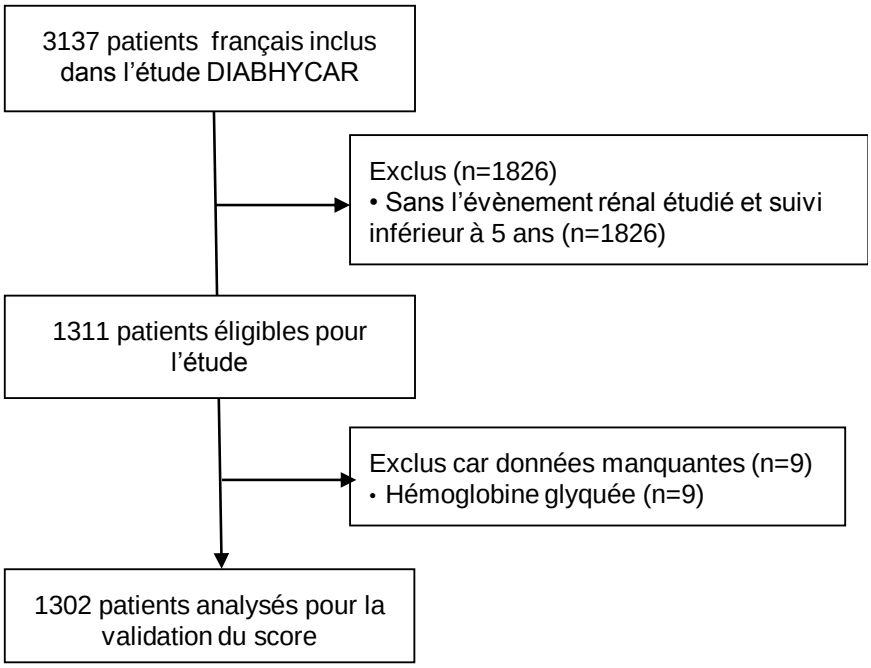

Figure 4 : Diagramme de flux des patients de l'étude DIABHYCAR considérés pour la validation externe du score NZDCS.

3.2.2 Caractéristiques à l'inclusion

L'ensemble des caractéristiques des patients et des fréquences de survenue des évènements rénaux étudiés (critère combiné IRCT fatale ou non) au cours du suivi sont présentées dans le tableau 6.

Tableau 6 : Caractéristiques à l'inclusion et évènements rénaux au cours du suivi des populations d'étude.

	DCS (n=25736)	SURGDIAGENE (n=827)	DIABHYCAR (n=1302)
Evènements rénaux			
IRCT	268 (1,0%)	51 (6%)	19 (1,5%)
Décès de cause rénale	369 (1,4%)	6 (0,7%)	4 (0,3%)
Evènement composite DCS [1]	637 (2,5%)	46 (5,5%)	19 (1,5%)
Démographie			
Age (années)	61 ± 13	63 ± 10	65 ± 8
Femmes	12984 (50%)	369 (45%)	345 (27%)
Ethnie			
Caucasien	14594 (57%)	800 (97%)	1302 (100%)
Africain	ND	14 (1,5%)	0 (0%)
Asiatique	1541 (6%)	4 (0,5%	0 (0%)
Autre [2]	9601 (38%)	9 (1%)	0 (0%)
Fonction rénale			
Créatinine sérique (µmol/l)	80 [70-98]	82 [70-99]	88 [75-100]
DFGe (ml/min/1.73m²) *	77 [64-91]	77 [60-92]	76 [63-89]
Albuminurie †			
Normoalbuminurie	16586 (65%)	445 (54%)	0 (0%)
Microalbuminurie	7042 (27%)	253 (31%)	1013 (78%)
Macroalbuminurie	1354 (5%)	62 (7%)	199 (15%)
Albuminurie avancée	754 (3%)	67 (8%)	90 (7%)
Diabète			
HbA1c (%)	7,2 [6,4-8,6]	7,7 [6,8-8,7]	7,5 [6,5-9,0]
Ancienneté du diabète (années)	4 [1-9]	12 [6-20]	9 [4-15]
Autres facteurs de risque cardiovasculaires			
IMC (kg/m²)	31 [27-35]	30 [27-34]	29 [26-32]
PAS (mmHg)	138 ± 19	133 ± 18	145 ± 14
Tabac			
Jamais	14533 (56%)	444 (54%)	ND
Actif	3958 (15%)	88 (10%)	187 (14%)
Ancien	7245 (28%)	295 (36%)	ND
Antécédents cardiovasculaires	6392 (25%)	276 (33%)	292 (22%)

Les variables qualitatives sont présentées en effectifs (pourcentages) et les variables quantitatives en moyenne ± écart-type ou médiane [25e -75e percentiles].
Les modalités de variables non définies pour les études sont notées ND.
Abréviations : IRCT, insuffisance rénale chronique terminale; DFGe, débit de filtration glomérulaire estimé; rapport A/C, rapport albuminurie sur créatinurie; HbA1c, hémoglobine glyquée; IMC, indice de masse corporelle; PAS, pression artérielle systolique.
[1] Premier évènement rénal composite au cours du suivi : IRCT ou décès de cause rénale.
[2] La catégorie "Autre" correspond aux catégories autres que "Caucasien", "Africain" et "Asiatique" pour la cohorte SURDIAGENE et aux catégories "Maori" et "Pacifique" et "Autres" pour la cohorte DCS.
* Calculé avec la formule du CKD-EPI.
† Les classes d'albuminurie sont définies telles que dans la cohorte DCS : une microalbuminurie comme un rapport A/C ≥ 3 mg/mmol et < 30 mg/mmol, une macroalbuminurie un rapport A/C ≥ 30 mg/mmol et < 100 mg/mmol et une albuminurie avancée un rapport A/C ≥ 100 mg/mmol. Pour la population ADVANCE, une macroalbuminuie est définie comme un rapport A/C ≥ 30 mg/mmol et la classe "avancée" n'était pas définie.

Comparaison des caractéristiques des populations DCS et SURDIAGENE

Les patients de l'étude NZDCS et de la cohorte SURDIAGENE étaient peu différents en termes d'âge moyen (61 ans et. 63 ans, respectivement) et de proportions de femmes (50% et 45 %, respectivement) mais les origines ethniques étaient très différentes avec notamment des ethnies plus variées dans l'étude DCS (5 ethnies principales vs. 3 dans SURDIAGENE). Les paramètres de la fonction rénale étaient relativement proches avec des valeurs médianes de créatinémie et de DFGe similaires à l'inclusion (80 µmol/l et 77 ml/min/1,73m² vs 82 µmol/l et 77 ml/min/1,73m² dans DCS et SURDIAGENE, respectivement); néanmoins, la proportion de sujets normo-albuminuriques était plus importante chez les patients DCS (65% vs 54% dans SURDIAGENE). Dans la cohorte SURDIAGENE, l'ancienneté du diabète était plus marquée (médiane de 12 ans vs 4 ans dans DCS) et le diabète était moins bien équilibré (HbA1c =7,7% vs 7,2% dans DCS). Les facteurs de risque cardiovasculaires étaient répartis de manière hétérogène selon les paramètres considérés. En comparaison avec les patients DCS, les patients SURDIAGENE présentaient une PAS moyenne plus faible (133 mmHg vs 138 mmHg), un IMC médian comparable (30 kg/m² vs 31 kg/m²), une moindre proportion de tabagisme actif (10% vs 15%) et plus d'ATCD CV (33% vs 25%).

Comparaison des caractéristiques des populations DCS et DIABHYCAR

Les patients de l'étude NZDCS et de l'essai DIABHYCAR n'étaient pas comparables en terme de démographie, les patients DIABHYCAR étaient en effet plus âgés en moyenne (65 ans vs 61 ans dans DCS) et la proportion de femmes était plus faible (27% vs 50% dans DCS). Les origines ethniques étaient également très différentes avec des ethnies plus diverses dans la population DCS (une seule ethnie caucasienne dans l'essai DIABHYCAR). L'atteinte rénale à l'inclusion était plus grave chez les patients DIABHYCAR par rapport aux patients DCS puisqu'ils étaient tous albuminuriques et que leur créatinémie médiane était plus élevée (88 µmol/l vs 80 µmol/l dans DCS). Dans la population DIABHYCAR, le diabète était d'une durée médiane plus longue (9 ans vs 4 ans dans DCS) et d'équilibre moins satisfaisant (HbA1c=7,5% vs 7,2% dans DCS). La répartition de certains facteurs de risque CV suivants étaient comparables : tabagisme actif (14% et 15 % dans DIABHYCAR et DCS, respectivement) et ATCD CV (22% et 25 % dans DIABHYCAR et DCS, respectivement). Cependant, les patients DIABHYCAR étaient plus souvent hypertendus (PAS moyenne = 145 mmHg vs 138 mmHg dans DCS) et présentaient un IMC légèrement moins élevé (29 kg/m² vs 31 kg/m² dans DCS).

3.2.3 Evènements rénaux au cours du suivi

Au cours du suivi, 637/25736 (2.5%) évènements composites IRCT non fatale ou décès de cause rénale sont survenus chez les patients DCS, 46/827 (5.5%) chez les patients SURDIAGENE, et 19/1302 (1.5%) chez les patients DIABHYCAR. La proportion de décès de cause rénale la plus élevée était retrouvée dans la population DCS en comparaison avec les populations SURDIAGENE et DIABHYCAR (N=369/25736 (1.4%), N=6/827 (0.7%) et N=4/1302 (0.3%), respectivement). L'apparition de l'évènement IRCT non fatale était plus fréquente au sein des patients de la cohorte SURDIAGENE par rapport aux populations DCS et DIABHYCAR (N=51/25736 (6%), N=268/827 (1.0%) et N=19/1302 (1.5%), respectivement)

3.2.4 Associations entre les paramètres du score et les évènements rénaux

L'ensemble des associations univariées entre les paramètres constituant le score NZDCS et la survenue de l'évènement rénal composite IRCT fatale ou décès de cause rénale sont présentées dans le tableau 7. Chez les sujets de la cohorte SURDIAGENE, la plupart des paramètres constituants le score NZDCS étaient significativement associés à la survenue d'une IRCT fatale ou non, seuls l'ethnie et l'âge au diagnostic du diabète n'étaient pas associés à cet évènement (p=0,99 et p=0,33, respectivement) mais contrairement à l'étude NZDCS, le sexe féminin et l'augmentation de l'HbA1c étaient des facteurs protecteurs (OR=0,28 [0,14- 0,60] et OR=0,80 [0,64 - 0,99], respectivement). Chez les sujets de l'essai DIABHYCAR, seules trois variables du score NZDCS étaient des facteurs de risque de survenue d'une IRCT fatale ou non, le sexe (OR=2,54 (1,02 - 6,30] p=0,04), la créatininémie (OR=1,54 [1,22 - 1,95] p=0,0003) et le niveau d'albuminurie (OR microalbuminurie=13,03 [2,51 - 67,63] p=<0,0001).

Tableau 7: Associations univariées entre les paramètres du score NZDCS et l'insuffisance rénale chronique terminale fatale ou non dans les populations de validation du score.

	SURDIAGENE (n=827)		DIABHYCAR (n=1302)	
Variables	**OR [IC 95%]**	**p**	**OR [IC 95%]**	**p**
Sexe (Femme)	0,28 [0,14- 0,60]	0,0009	2,54 (1,02 - 6,30]	0,04
Age au diagnostic du DT2 (années)	0,99 [0,96 - 1,01]	0,33	0,99 [0,95 - 1,04]	0,7
Ethnie‡		0,99	NA	
Caucasien	1			
Africain	1,29 [0,16 - 10,09]			
Asiatique	*			
Autre	*			
Durée du DT2 (années)	1,05 [1,02 - 1,08]	0,0006	1,04 [0,98 - 1,09]	0,2
Créatinine sérique (par 10 mg/dl)	1,41 [1,30-1,52]	<0,0001	1,54 [1,22 - 1,95]	0,0003
Albuminurie †		<0,0001		<0,0001
Normoalbuminurie	1		1	
Microalbuminurie	4,47 [0,86 -23,19]		13,03 [2,51 - 67,63]	
Macroalbuminurie	95,98 [22,76 - 404,67]		77,77 [17,10 - 353, 70]	
HbA1c (%)	0,80 [0,64 - 0,99]	0,04	1,21 [0,96 - 1,51]	0,1
PAS (par 10 mmHg)	1,37 [1,17 - 1,60]	<0,0001	0,91 [0,65 - 1,28]	0,6
Tabac §		0,03		0,8
Jamais	1		1	
Ancien	2,33 [1,23 - 4,39]		NA	
Actif	1,20 [0,39 - 3,64]		1,12 [0,32 - 3,88]	
Antécédents cardiovasculaires	2,76 [1,51 - 5,04]	0,0009	0,92 [0,30 - 2,80]	0,9

Les odds ratio univariés (OR) sont accompagnés de leurs intervalles de confiance à 95% (IC95%) et du degré de significativité p.
Abréviations : DT2, diabète de type 2; HbA1c, hémoglobine glyquée; PAS, pression artérielle systolique; NA, non applicable.
‡ Dans l'essai DIABHYCAR, l'ensemble des sujets français sont caucasiens.
† Les classes d'albuminurie sont définies telles que dans l'étude NZDCS : une microalbuminurie comme un rapport A/C ≥ 3 mg/mmol et < 30 mg/mmol, une macroalbuminurie un rapport A/C ≥ 30 mg/mmol et < 100 mg/mmol et une albuminurie avancée un rapport A/C ≥ 100 mg/mmol. Pour la population ADVANCE, une macroalbuminuie est définie comme un rapport (A/C) ≥ 30 mg/mmol et la classe "avancée" n'était pas définie.
* En raison d'un nombre d'évènements trop faible pour les catégories « asiatique » et « autre » de la variable ethnie, l'odds ratio ne peut pas être calculé de manière satisfaisante et était estimé à <0,001 avec un intervalle de confiance à 95 % compris entre <0,001 et >999,99.
§ Dans l'essai DIABHYCAR, les deux modalités de la variable tabac sont "oui" et "non".

3.2.5 Estimation du risque d'évènement prédit par le score

L'estimation du risque d'apparition d'une IRCT fatale ou non selon le score NZDCS a consisté en l'application de l'équation de risque suivante pour chaque sujet des populations de validation : *1-0,99558399^exp ((Sexe féminin*0,401069) + (Ethnie « Maori »*0,821956) + (Ethnie « Pacific »*0,296149) + (Ethnie « East Asian »*-0,138466) + (Ethnie «Indo-Asian »*-0,040645) + (Ethnie «autre »*-0,223648) + (Age au diagnostic du DT2*0,024199) + (Durée du DT2*0,043467) + (Créatinémie (µmol/L)/10 *0,305281) + (Microalbuminurie*0,644200)*

*+ (Macroalbuminurie*1,303188)+ (Albuminurie avancée* 2,290089) + (PAS (mmHg)/10*0,049809) + (HbA1c(%)/10*0,158597) + (Ancien fumeur*-0,003952) + (Fumeur actuel*0,252272) + (ATCD CV*0,498804) - 7,0671736)).*

Dans les populations SURDIAGENE et DIABHYCAR, les risques médians d'apparition d'une IRCT fatale ou non dans les 5 ans selon le score NZDCS étaient de 0,4 % (25e-75e percentiles : 0,2%-1,1%) et de 0,7 % (25e-75e percentiles : 0,4%-1,4%), respectivement. Les groupes de risque ont été établis en fonction des déciles du risque prédit (tableau 8). La confrontation du nombre de cas prédits par le score et du nombre de cas réellement observés dans chaque groupe de risque, montrait une cohérence plus satisfaisante dans la population SURDIAGENE que dans la population DIABHYCAR avec, respectivement, la quasi -totalité des cas observés (90%, n=41/46) vs moins de ¾ des cas observés (73%, n=14/19), classés dans le groupe le plus à risque (10e décile) (figure 5).

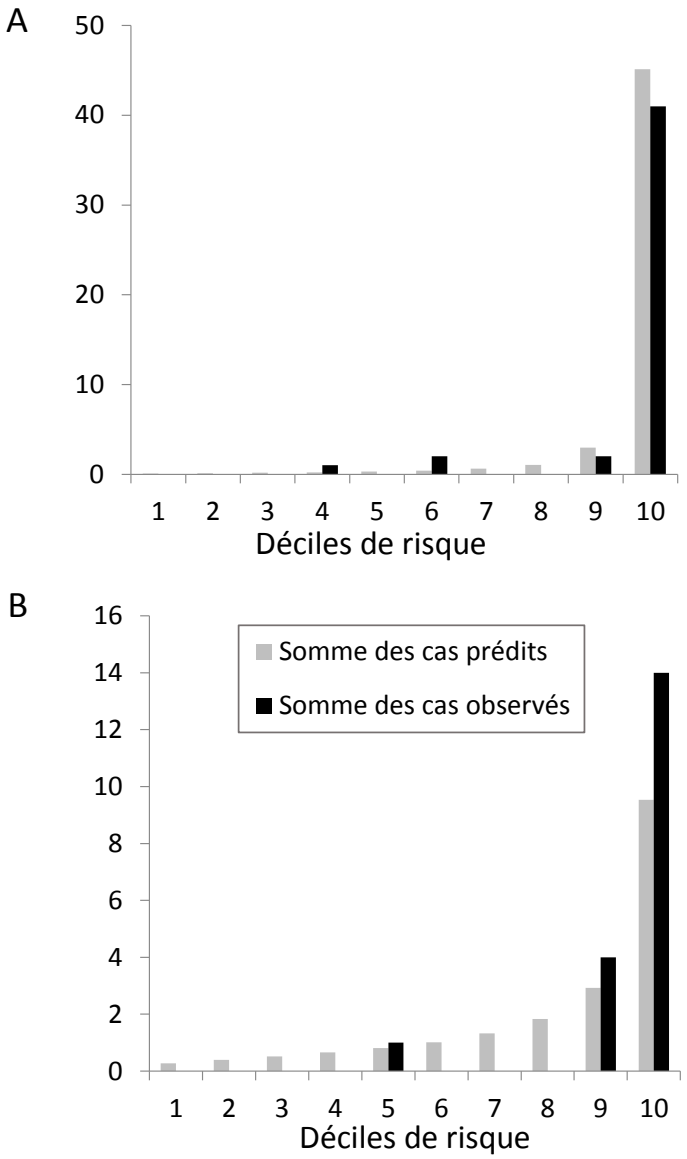

Figure 5 : Nombres de cas observés et de cas prédits d'insuffisance rénale chronique terminale fatale ou non par déciles de risque dans les populations de validation du score NZDCS (A, population SURDIAGENE ; B, population DIABHYCAR)

3.2.6 Evaluation des performances prédictives

Calibration

L'étude de la calibration du score NZDCS par l'approche du graphique de calibration (figure 6) montrait que les prédictions de cas et les proportions observées étaient plus proches dans la population SURDIAGENE que dans la population DIABHYCAR (points plus proches de la ligne de calibration idéale). Le score avait tendance à surestimer le nombre de cas prédits chez les patients SURDIAGENE avec des prédictions supérieures aux proportions de cas observés

dans la plupart des déciles de risque (tableau 8 ; figure 6, points situés en dessous de la ligne de calibration idéale). De la même manière chez les patients DIABHYCAR, le nombre total de cas prédits était surestimé par rapport au nombre total de cas observés (tableau 8). De plus, les cas observés par décile de risque étaient majoritairement regroupés dans les valeurs de prédiction les plus faibles (tableau 8), traduisant un probable défaut d'étalonnage du score dans les deux populations. La notion d'une adéquation du score NZDCS plus satisfaisante dans la population SURDIAGENE était confirmée par le test d'adéquation du Chi2 de Hosmer et Lemeshow qui était non significatif (Chi2= 6,6, p=0,58); alors qu'il était à la limite de la significativité dans la population DIABHYCAR (Chi2= 14,8, p=0,06).

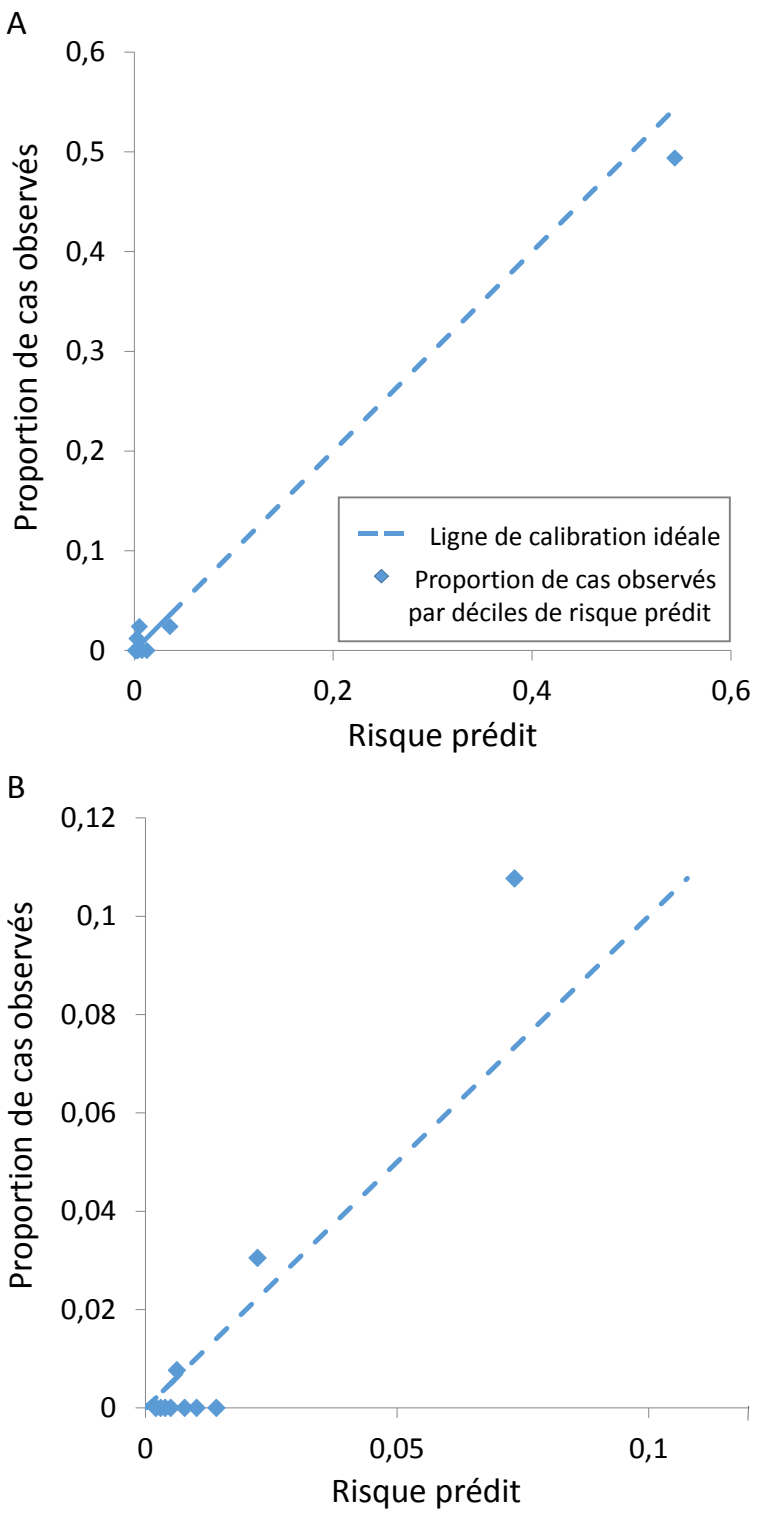

Figure 6 : Graphiques de calibration du score NZDCS dans les populations de validation (A, population SURDIAGENE ; B, population DIABHYCAR).

Tableau 8 : Calibration du score NZDCS dans les populations de validation SURDIAGENE et DIABHYCAR

	SURDIAGENE				DIABHYCAR			
Déciles de risque	**P (%)***	**N**	**Cas prédits**	**Cas observés**	**P (%)***	**N**	**Cas prédits**	**Cas observés**
1	<0,13	83	0,08	0	<0,27	130	0,27	0
2	0,13-0,18	82	0,12	0	0,27-0,34	130	0,39	0
3	0,19-0,23	83	0,17	0	0,35-0,44	130	0,51	0
4	0,24-0,31	83	0,22	1	0,45-0,55	131	0,66	0
5	0,32-0,42	82	0,30	0	0,56-0,69	130	0,81	1
6	0,43-0,57	83	0,41	2	0,70-0,87	130	1,01	0
7	0,58-0,90	82	0,62	0	0,88-1,18	130	1,32	0
8	0,91-1,76	83	1,04	0	1,19-1,66	130	1,83	0
9	1,77-7,88	83	2,97	2	1,67-2,89	131	2,92	4
10	>7,88	83	45,11	41	>2,89	130	9,53	14
		827	51,04	46		1302	19,27	19

*Bornes des déciles de risque.

Discrimination

Le score NZDCS présentait des performances de discrimination très satisfaisantes dans les deux populations de validation (figure 7). La discrimination du score était cependant meilleure chez les patients SURDIAGENE que chez les patients DIABHYCAR (ASC = 0,96 [0,92 – 0,99] et ASC = 0,92 [0,86-0,97], respectivement).

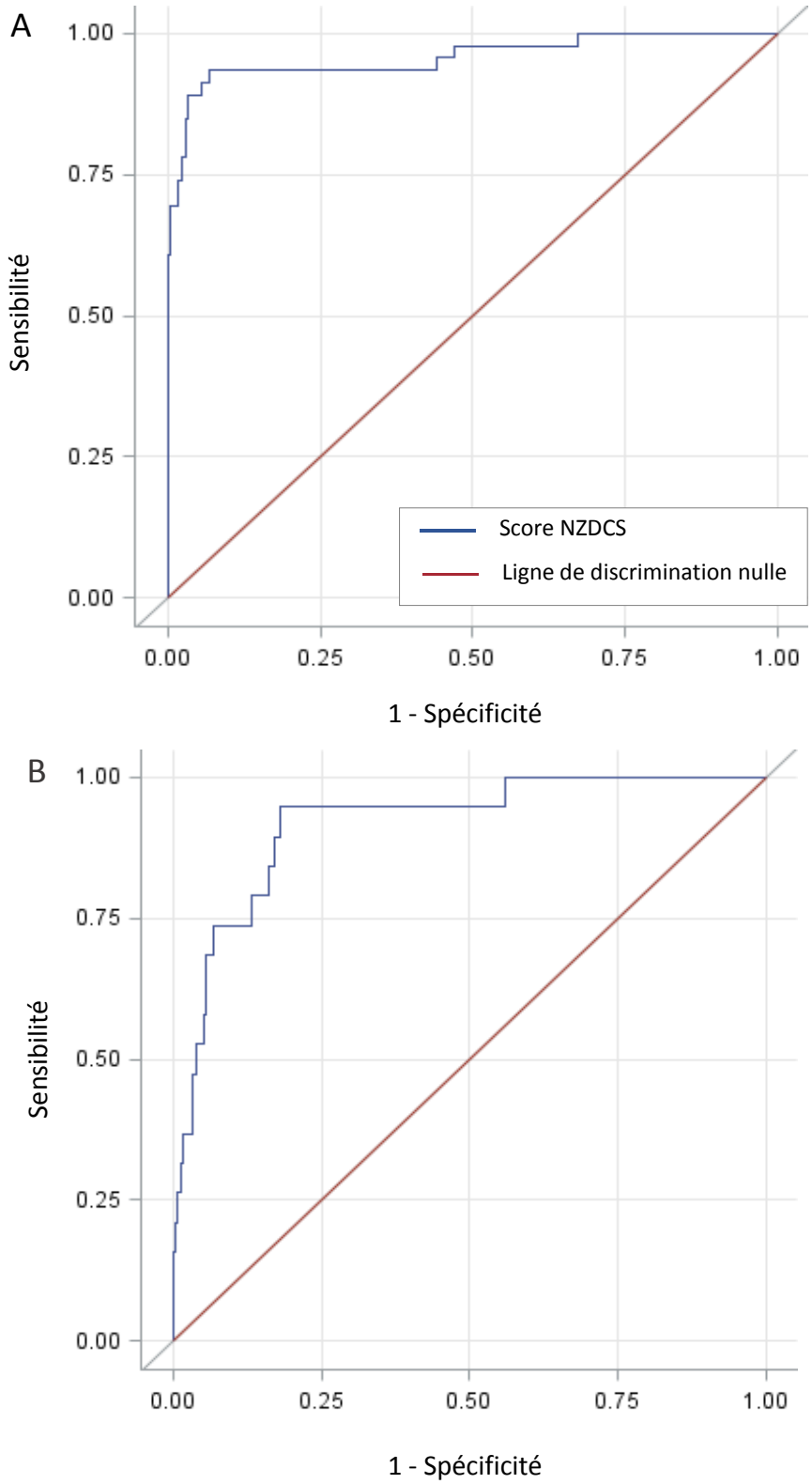

Figure 7: Courbes ROC (receiver operating characteristic) du score NZDCS pour l'insuffisance rénale chronique terminale fatale ou non (A, population SURDIAGENE ; B, population DIABHYCAR).

3.3 Validation du score ADVANCE

3.3.1 Populations d'étude

Les populations considérées pour la validation externe du score ADVANCE issue des études SURDIAGENE et DIADHYCAR pour lesquelles le temps de suivi était suffisant, et qui ne présentaient pas de données manquantes pour les variables intégrées au score consistaient en 856 et 1330 patients respectivement (figures 8 et 9).

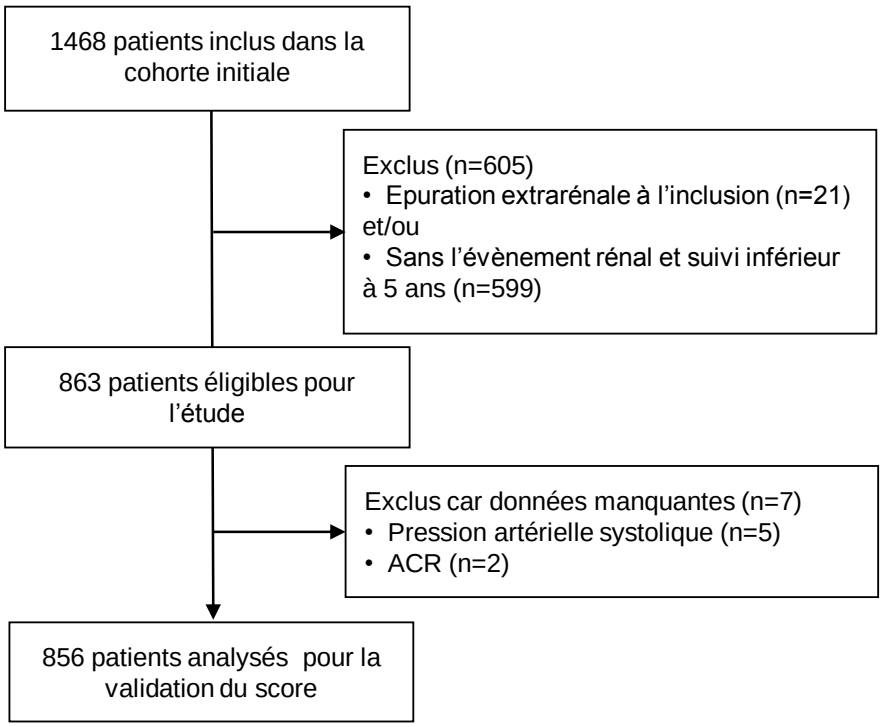

Figure 8 : Diagramme de flux des patients de la cohorte SURDIAGENE considérés pour la validation externe du score ADVANCE.

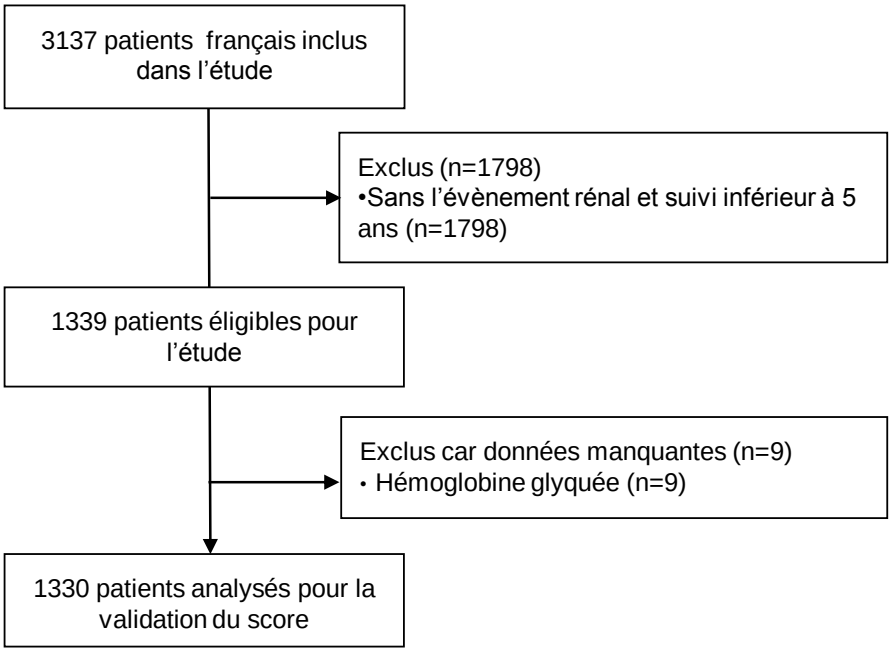

Figure 9 : Diagramme de flux des patients de l'étude DIABHYCAR considérés pour la validation externe du score ADVANCE.

3.3.2 Caractéristiques à l'inclusion

L'ensemble des caractéristiques des patients et des fréquences de survenue des évènements rénaux étudiés (doublement de la créatinine sérique ou IRCT fatale ou non) au cours du suivi sont présentées dans le tableau 9.

Tableau 9 : Caractéristiques à l'inclusion et évènements rénaux au cours du suivi des populations d'étude

	ADVANCE (n=11140)	SURGDIAGENE (n=856)	DIABHYCAR (n=1330)
Evènements rénaux			
IRCT	86 (0,8%)	52 (6,1%)	19 (1,4%)
Décès de cause rénale	37 (0,3%)	6 (0,7%)	4 (0,3%)
Doublement de la créatinine sérique	129 (1,2%)	80 (9,3%)	67 (5,0%)
Evènement combiné ADVANCE [1]	166 (1,5%)	90 (10,5%)	76 (5,7%)
Démographie			
Age (années)	66 ± 6	63 ± 10	65 ± 8
Hommes	6405 (57%)	478 (56%)	979 (74%)
Ethnie			
Caucasien	6684 (60%)	829 (96,8%)	1189 (100%)
Africain	-	14 (1,7%)	-
Asiatique	4242 (38%)	4 (0,5%)	-
Autre [2]	2228 (2%)	9 (1,0%)	-
Fonction rénale			
Créatinine sérique (mg/dl)	0,95 [0,80-1,11]	0,93 [0,79-1,12]	1,00 [0,85-1,14]
DFGe (ml/min/1,73m²) *	74,6 ± 17,6	73,6 ± 24,0	75,3 ± 17,0
Rapport (A/C) (mg/g)	15,0 [7,1-39,8]	25,5 [9,7-129,0]	115,7 [60,0-271,5]
Albuminurie †			
Normoalbuminurie	7377 (69%)	452 (53%)	0 (0%)
Microalbuminurie	2862 (27%)	264 (31%)	1023 (77%)
Macroalbuminurie	401 (4%)	140 (16%)	307 (23%)
Diabète			
HbA1c (%)	7,5 ± 1,6	7,7 ± 1,6	7,9 ± 1,8
Rétinopathie diabétique	2979 (27%)	363 (42%)	70 (5%)
Autres			
PAS (mmHg)	145 ± 22	133 ± 18	145 ± 14
PAD (mmHg)	81 ± 11	73 ± 11	80 ± 8
Tabac actif	1682 (15%)	89 (10%)	193 (14%)
Antécédents cardiovasculaires	2228 (20%)	286 (33%)	303 (23%)

Les variables qualitatives sont présentées en effectifs (pourcentage) les variables quantitatives en moyenne ± écrat-type ou médiane [25e -75e percentiles].

Abréviations : IRCT, insuffisance rénale chronique terminale; DFGe, débit de filtration glomérulaire estimé; rapport A/C, rapport albuminurie sur créatininurie; HbA1c, hémoglobine glyquée; PAS, pression artérielle systolique; PAD, pression artérielle diastolique.

[1] Premier évènement rénal composite au cours du suivi : doublement de la créatinine sérique ou IRCT ou décès de cause rénale.

[2] La catégorie "Autre" de la variable ethnie est définie telle que non caucasien, non africain ou non asiatique pour la cohorte SURDIAGENE; non caucasien et non asiatique pour la cohorte ADVANCE.

* Calculé avec la formule du CKD-EPI.

† Les classes d'albuminurie sont définies telles que dans la cohorte DCS : une microalbuminurie comme un rapport A/C ≥ 3 mg/mmol et < 30 mg/mmol, une macroalbuminurie un rapport A/C ≥ 30 mg/mmol et < 100 mg/mmol et une albuminurie avancée un rapport A/C ≥ 100 mg/mmol. Pour la population ADVANCE, une macroalbuminuie est définie comme un rapport A/C ≥ 30 mg/mmol et la classe "avancée" n'était pas définie.

Comparaison des caractéristiques des populations ADVANCE et SURDIAGENE

Les patients de l'essai ADVANCE étaient plus âgés en moyenne que les patients de la cohorte SURDIAGENE (66 ans vs 63 ans, respectivement) et présentaient des ethnies différentes (40% d'ethnies non caucasiennes dans ADVANCE vs. 3,2% dans SURDIAGENE). D'un point de vue démographique, seules les proportions d'hommes étaient similaires (56% vs 57% dans SURDIAGENE et ADVANCE, respectivement). Les paramètres de la fonction rénale étaient comparables entre les patients SURDIAGENE et ADVANCE ; valeurs médianes de créatinémie (0,93 mg/dl vs 0,95 mg/dl, respectivement) et de DFGe (73,6 ml/min/1,73 m² vs 74,6 ml/min/1,73 m², respectivement). Cependant, la proportion de sujets albuminuriques était plus importante chez les patients SURDIAGENE (47% vs 31% dans ADVANCE). Le diabète était légèrement moins bien équilibré et la rétinopathie diabétique était plus fréquente dans la population SURDIAGENE que dans la population ADVANCE (HbA1c=7,7 % vs 7,5%, et rétinopathie 42% vs.27%, respectivement). Les facteurs de risque cardiovasculaires étaient plus souvent retrouvés chez les sujets ADVANCE que chez les patients SURDIAGENE avec des niveaux tensionnels moyens plus élevés (145 mmHg vs 133 mmHg), une proportion de tabagisme actif plus importante (15% vs 10%) ; les ATCD CV étaient cependant plus fréquents dans la population SURDIAGENE (33% vs 20%).

Comparaison des caractéristiques des populations ADVANCE et DIABHYCAR

Les patients des essais ADVANCE et DIABHYCAR présentaient des âges moyens similaires (65 ans vs 66 ans, respectivement) mais la proportion d'hommes était plus importante dans l'étude DIABHYCAR (74% vs 57% dans ADVANCE). Les origines ethniques étaient également très différentes, l'ensemble de la population DIABHYCAR étant caucasienne. L'atteinte rénale à l'inclusion était plus grave chez les patients DIABHYCAR qui présentaient des paramètres systématiquement plus élevés en comparaison avec les sujets ADVANCE : la créatinémie médiane (1,00 mg/dl vs 0,95 mg/dl, respectivement), le DFGe moyen (75,3 ml/min/1,73m² vs 74,6 ml/min/1,73m², respectivement) et le rapport A/C médian (115,7 mg/g vs 15 mg/g, respectivement) et une proportion de patients macro-albuminuriques plus importante (23% vs 4%, respectivement) Dans la population DIABHYCAR, le diabète paraissait d'équilibre moins satisfaisant (HbA1C=7,9% vs 7,5% dans ADVANCE), on y retrouvait cependant moins de patients présentant une rétinopathie (5% vs 27% dans ADVANCE). Les facteurs de risque CV étaient répartis de manière similaire entre les deux populations avec des sujets plutôt hypertendus (PAS moyenne =145 mmHg dans les deux

populations), des proportions comparables de sujets présentant des ATCD CV (23% et 20% dans DIABHYCAR et ADVANCE, respectivement) et de patients tabagiques actifs (14% et 15 % dans DIABHYCAR et ADVANCE, respectivement).

3.3.3 Evènements rénaux au cours du suivi

Au cours du suivi, N=166/11140 (1,5%) évènements composites doublement de la créatinine ou IRCT fatale ou non sont survenus chez les patients ADVANCE, N=90/856 (10,5%) chez les patients SURDIAGENE, et N=76/1330 (5,7%) chez les patients DIABHYCAR. Les proportions de décès de cause rénale étaient identiques entre les populations des essais ADVANCE et DIABHYCAR, bien qu'inférieures à celle observée dans la population SURDIAGENE (N=37/11140 (0,3%), N=6/856 (0,7%), et N=4/1330 (0,3%), respectivement). Les autres évènements rénaux considérés isolément étaient systématiquement plus fréquents chez les patients de la cohorte SURDIAGENE par rapport aux populations ADVANCE et DIABHYCAR (IRCT : N=86/11140 (0,8%), N=52/856 (6,1%), et N=19/1330 (1,4%), respectivement, et doublement de la créatininémie : N=129/11140 (1,2%), N=80/856 (9,3%), et N=67/1330 (5,0%), respectivement).

3.3.4 Associations entre les paramètres du score et les évènements rénaux

L'ensemble des associations univariées entre les paramètres constituant le score ADVANCE et la survenue de l'évènement rénal composite doublement de la créatinine ou IRCT fatale ou décès de cause rénale sont présentées dans le tableau 10.

Chez les sujets de la cohorte SURDIAGENE, la plupart des paramètres constituants le score NZDCS étaient significativement associés à la survenue d'une IRCT fatale ou non et en étaient également des facteurs de risque ; seul le niveau d'HbA1c n'étaient pas associé à cet évènement (p=0,9). Chez les sujets de l'essai DIABHYCAR, seules trois variables du score ADVANCE étaient des facteurs de risque de survenue de l'évènement rénal étudié, le DFGe (OR=0,96 [0,95 - 0,98]), l'albuminurie (OR=8,94 [5,73 - 13,93]), et le niveau d'HbA1c (OR=1,20 [1,07 - 1,35]).

Tableau 10: Associations univariées entre les paramètres du score ADVANCE et l'évènement rénal composite doublement de la créatinine sérique ou insuffisance rénale chronique terminale ou décès de cause rénale dans les populations de validation du score.

	SURGDIAGENE (n=856)		DIABHYCAR (n=1330)	
Variables	**OR [IC 95%]**	**p**	**OR [IC 95%]**	**p**
Sexe (Homme)	2,86 [1,72- 4,77]	<0,0001	1,08 [0,63 - 1,84]	0,8
DFGe (10ml/min/1,73m²)*	0,54 [0,49-0,61]	<0,0001	0,96 [0,95 - 0,98]	<0,0001
Rapport A/C (log µg/mg)	7,40 [5,25 - 10,44]	<0,0001	8,94 [5,73 - 13,93]	<0,0001
HbA1c (%)	0,99 [0,86 - 1,14]	0,9	1,20 [1,07 - 1,35]	0,002
PAS (10 mmHg)	1,42 [1,26 - 1,60]	<0,0001	1,00 [0,84 - 1,18]	>0,99
Rétinopathie diabétique	3,49 [2,17 - 5,60]	<0,0001	1,59 [0,67 - 3,80]	0,3

Les odds ratio univariés (OR) sont accompagnés de leurs intervalles de confiance à 95% (IC95%) et du degré de significativité p.
Abréviations : DFGe, débit de filtration glomérulaire estimé; HbA1c, hémoglobine glyquée; PAS, pression artérielle systolique.
* Calculé selon la formule de CKD-EPI.

3.3.5 Estimation du risque d'évènement prédit par le score

L'estimation du risque d'apparition du premier évènement rénal composite doublement de la créatinémie, ou d'une IRCT non fatale ou d'un décès de cause rénale selon le score ADVANCE a consisté en l'application de l'équation de risque suivante pour chaque sujet des populations de validation : *1 - 0.99407^exp (0.89405 (si homme) - 0.04873* DFGe (mL/min/1.73 mm^2) + 0.44790*log rapport (A/C) (µg/mg) + 0.01163*PAS (mmHg) + 0.11618*HbA1c (%) + 0.42146 (si rétinopathie diabétique) + 0.41098 (si âge à la fin du premier cycle de l'enseignement secondaire ≤15 ans) - 1.13663.*

Dans les populations SURDIAGENE et DIABHYCAR, les risques médians d'apparition de l'évènement rénal d'intérêt dans les 5 ans prédits par le score ADVANCE étaient de 13 % (25e-75e percentiles : 6,6%-28%) et de 1,1 % (25e-75e percentiles : 0,6%-2,3%), respectivement. Les groupes de risque ont été établis en fonction des déciles du risque prédit (tableau 11). La confrontation des nombres de cas prédits par le score et de cas réellement observés dans chaque groupe de risque, montrait un classement des cas plus satisfaisant dans la population SURDIAGENE pour laquelle la proportion de cas observés classés dans le groupe le plus à risque (10e décile) était plus importante que dans la population DIABHYCAR (61%, n=55/90 vs 51%, n=39/76, respectivement) (figure 10).

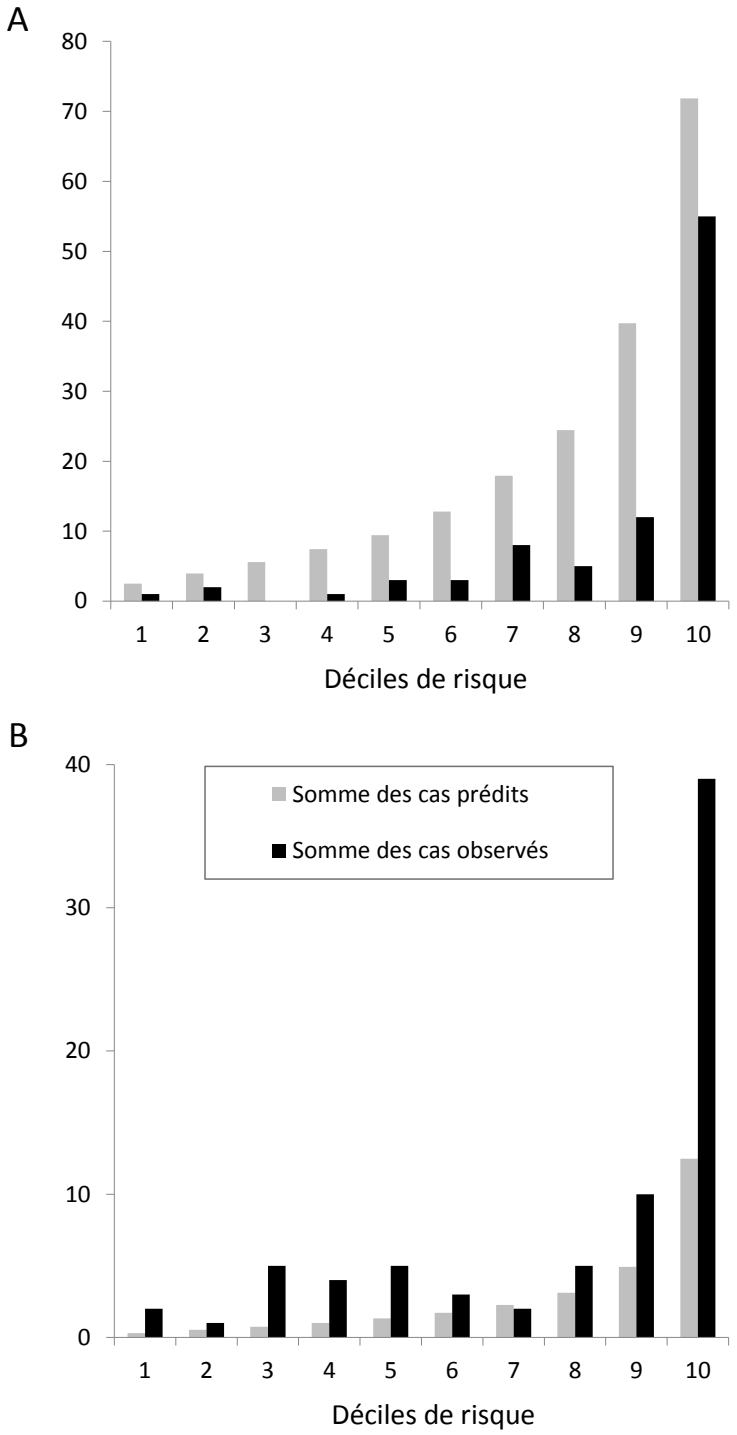

Figure 10 : Nombres de cas observés et de cas prédits de doublement de la créatinine, ou d'insuffisance rénale chronique terminale non fatale, ou de décès de cause rénale par déciles de risque dans les populations de validation du score ADVANCE (A, population SURDIAGENE ; B, population DIABHYCAR)

3.3.6 Evaluation des performances prédictives

Calibration

L'étude de la calibration du score ADVANCE par l'approche du graphique de calibration (figure 11) montrait que les observations et les prédictions n'étaient pas similaires, pour les patients SURDIAGENE comme pour les patients DIABHYCAR (points éloignés de la ligne de calibration idéale). Chez les patients SURDIAGENE, le score ADVANCE avait tendance à surestimer l'apparition de l'évènement puisque les prédictions étaient systématiquement supérieures aux proportions de cas observés dans chaque décile (tableau 11 ; figure 11, points situés en-dessous de la ligne de la calibration idéale). Au contraire, le score ADVANCE sous-estimait la survenue de l'évènement avec des valeurs de prédiction systématiquement inférieures aux proportions de cas observés dans chaque décile (tableau 11 ; figure 11, points situés au-dessus de la ligne de calibration idéale). De façon similaire dans les deux populations, les proportions de cas observés par décile de risque semblaient augmenter progressivement avec les valeurs de prédiction, et malgré un regroupement d'observations dans les valeurs de prédiction les plus faibles (tableau 11), cela démontrait une relative capacité d'étalonnage du score. Par ailleurs, les tests d'adéquation du score de Hosmer et Lemeshow étaient non significatifs (Chi2= 11,3, p=0,18 et Chi2= 7,6, p=0,47, dans les populations SURDIAGENE et DIABHYCAR, respectivement).

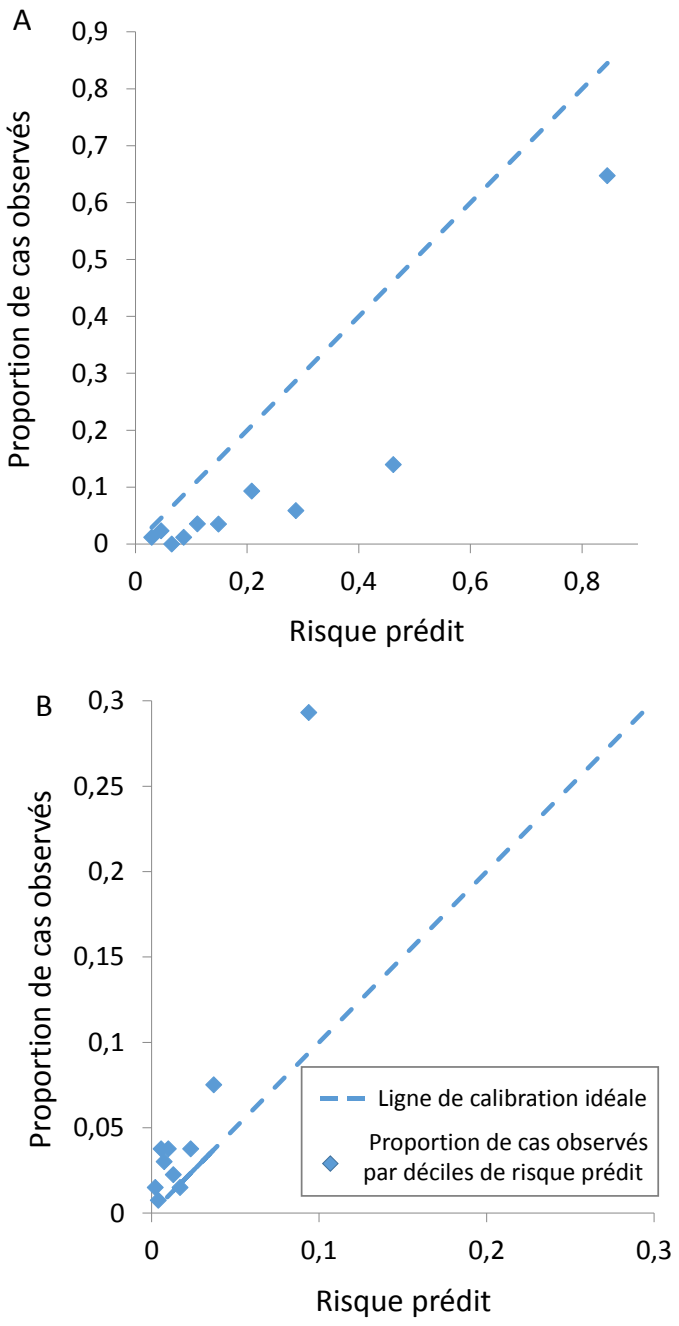

Figure 11 : Graphiques de calibration du score ADVANCE dans les populations de validation (A, population SURDIAGENE ; B, population DIABHYCAR).

Tableau 11 : Calibration du score ADVANCE dans les populations de validation SURDIAGENE et DIABHYCAR.

	SURDIAGENE				DIABHYCAR			
Déciles de risque	**P (%)***	**N**	**Cas prédits**	**Cas observés**	**P (%)***	**N**	**Cas prédits**	**Cas observés**
1	<3,9	85	2,49	1	<0,31	133	0,292	2
2	3,9-5,6	86	3,96	2	0,31-0,47	133	0,53	1
3	5,7-7,4	86	5,59	0	0,48-0,65	133	0,75	5
4	7,5-9,7	86	7,43	1	0,66-0,86	133	1,00	4
5	9,8-12,7	85	9,43	3	0,87-1,14	133	1,32	5
6	12,8-17,2	86	12,80	3	1,15-1,49	133	1,72	3
7	17,3-24,5	86	17,93	8	1,50-1,94	133	2,27	2
8	24,6-34,3	85	24,44	5	1,95-2,87	133	3,11	5
9	34,4-63,3	86	39,733	12	2,88-4,96	133	4,93	10
10	>63,3	85	71,84	55	>4,96	133	12,48	39
		856	195,65	90		1330	28,41	76

** Bornes des déciles de risque*

Discrimination

Les performances de discrimination du score ADVANCE n'étaient pas similaires dans les deux populations de validation (Figure 12). La discrimination du score était très bonne pour les patients SURDIAGENE, meilleure que pour les patients DIABHYCAR (ASC = 0,88 [0,84 – 0,92] et ASC = 0,74 [0,6-0,80], respectivement).

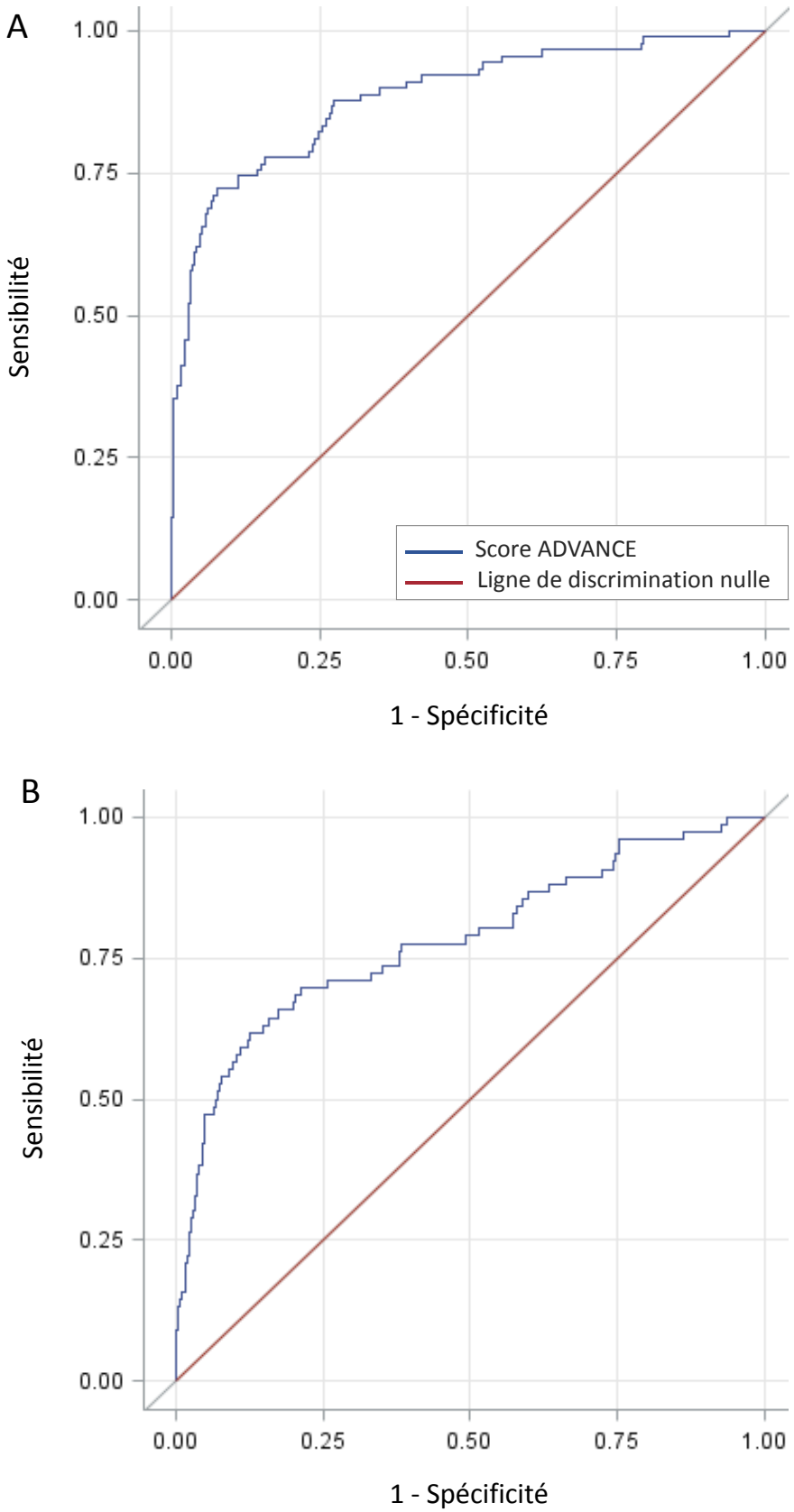

Figure 12: Courbes ROC (receiver operating characteristic) du score ADVANCE pour l'évènement rénal composite doublement de la créatinine sérique, insuffisance rénale chronique terminale non fatale ou décès de cause rénale (A, population SURDIAGENE ; B, population DIABHYCAR).

3.4 Comparaison des scores de risque rénal

Les performances prédictives des scores NZDCS et ADVANCE dans la population SURDIAGENE sont résumées tableau 12. Les performances des scores en termes de calibration n'étaient pas aussi satisfaisantes que leurs capacités de discrimination. Le score NZDCS paraissait globalement plus performant que le score ADVANCE pour les patients de la cohorte SURDIAGENE.

Tableau 12 : Synthèse des critères de calibration et de discrimination des scores NZDCS et ADVANCE dans la population de validation SURDIAGENE

	Score NZDCS	Score ADVANCE
Calibration		
Graphique de calibration	Observations proches de la ligne de calibration idéale Surestimation du risque Etalonnage non satisfaisant	Observations éloignées de la ligne de calibration idéale Surestimation du risque Etalonnage satisfaisant
Test d'adéquation : Chi 2 de Hosmer et Lemeshow	6,6 (p=0,58)	11,3 (p=0,18)
Discrimination		
Aire sous la courbe [IC 95%]	0,96 [0,92-0,99]	0,88 [0,84-0,92]

Les performances prédictives des scores NZDCS et ADVANCE dans la population DIABHYCAR sont résumées tableau 13. De la même manière que dans la population SURDIAGENE, les performances des scores en termes de calibration n'étaient pas aussi satisfaisantes que leurs capacités de discrimination et le score NZDCS paraissait globalement plus performant que le score ADVANCE pour les patients de l'essai DIABHYCAR.

Tableau 13 : Synthèse des critères de calibration et de discrimination des scores NZDCS et ADVANCE dans la population de validation DIABHYCAR

	Score NZDCS	Score ADVANCE
Calibration		
Graphique de calibration	Observations éloignées de la ligne de calibration idéale Surestimation du risque Etalonnage non satisfaisant	Observations éloignées de la ligne de calibration idéale Sous-estimation du risque Etalonnage satisfaisant
Test d'adéquation : Chi 2 de Hosmer et Lemeshow	14,8 (p=0,06)	7,6 (p=0,47)
Discrimination		
Aire sous la courbe [IC 95%]	0,92 [0,86-0,97]	0.74 [0.68-0.80]

4. Discussion

4.1 Principaux résultats de l'étude

La validation externe indépendante des scores de prédiction du risqué rénal NZDCS et ADVANCE chez les patients DT2 SURDIAGENE et DIABHYCAR a montré que leurs performances prédictives étaient différentes selon la population considérée mais également variables au sein d'une même population en termes de calibration et de discrimination. Le score NZDCS présentait une excellente discrimination associée à une moindre calibration dans les deux populations. Le score ADVANCE montrait également des performances de discrimination supérieures à ses capacités de calibration dans les deux populations de validation, mais ses capacités de discrimination étaient systématiquement plus faibles en comparaison avec le score NZDCS. Dans cette étude, le score NZDCS était globalement plus performant que le score ADVANCE pour prédire le risque l'apparition d'évènements rénaux majeurs à 5 ans pour les patients DT2 des études SURDIAGENE et DIABHYCAR.

4.2 Avantages et limites de l'étude

Dans cette étude, un grand nombre de modèles et de scores de risque rénal ont été identifiés par une recherche systématique dans la littérature. Les deux scores finalement inclus répondaient à des critères de sélection stricts et objectifs afin de répondre au mieux à la problématique tout en prenant en compte l'apparition croissante de bonnes pratiques méthodologiques encadrant la conduite d'une étude de validation externe[20–24,66–69,71].

Un autre avantage de ce travail a été d'évaluer plusieurs scores de risque simultanément et comparativement dans deux contextes populationnels différents et complémentaires. A ce jour, aucune étude n'a rapporté une telle évaluation. Par ailleurs, les populations de validation externe sélectionnées consistaient en de grands échantillons de patients DT2 dont les données étaient recueillies de manière prospective au cours de durée de suivi de plusieurs années. La sélection de sujets DT2 issus de deux types d'étude différents a contribué à renforcer la généralisation des résultats à plusieurs types de patients DT2. Les sujets DT2 de la cohorte SURDIAGENE représentaient des patients aux profils clinico-biologiques hétérogènes, dans un contexte de « vie réelle ». Au contraire des participants de l'étude DIABHYCAR, qui

reflétaient une population DT2 triée et sélectionnée, sur des critères d'atteinte rénale avancée notamment, bénéficiant d'un suivi programmé dans le cadre d'un essai contrôlé randomisé. De plus, le choix de populations de validation provenant de contextes temporels et géographiques éloignés des cadres dans lesquels les scores étaient construits, participait à la qualité de l'étude en élargissant le spectre de patients auxquels les résultats pouvaient être extrapolés.

Un certain nombre de limites sont à prendre en compte dans cette étude. La majorité des participants des études SURDIAGENE et DIABHYCAR étaient d'origine caucasienne ; cela a pu minorer la contribution de l'ethnie dans la prédiction des évènements rénaux par le score NZDCS. De plus, l'impact du sexe a également pu être mal évalué car seules les données des participants français de l'essai DIABHYCAR étaient disponibles, à l'origine d'un probable déséquilibre de la randomisation pour ce facteur expliquant la proportion plus élevée d'hommes dans cette population.

La non définition des variables des scores dans les populations de validation étaient un des principaux critères d'exclusion dans cette étude. Bien que non disponible dans les deux populations de validation, la variable de l'âge à la fin du premier cycle de l'enseignement secondaire du modèle ADVANCE a été fixée à 16 ans pour tous les sujets étudiés, conformément à l'âge moyen de fin de premier cycle de l'enseignement secondaire dans les pays de l'Organisation de Coopération et de Développement Économiques (OCDE)[72]. Un biais éventuel a été consenti pour plusieurs raisons. D'une part, la perte d'information n'a pas été considérée comme importante car cette variable était déjà dichotomisée dans le modèle (≤15 ans versus ≥ 16 ans), l'information qu'elle amenait était donc moindre que si cet âge avait été en continu. D'autre part, la variable avait été prise en compte de manière similaire dans une récente étude de validation du score cardiovasculaire ADVANCE[37] .

Du point de vue méthodologique, les analyses statistiques conduites ont été conditionnées par le niveau d'information disponible dans les publications des scores étudiés. Les données individuelles n'étant pas disponibles, l'évaluation des performances prédictives a tenu compte des informations concernant la description des caractéristiques des populations de développement des scores, des fréquences de survenue des évènements, des coefficients des paramètres des modèles, et des équations de risque. Il aurait été intéressant et complémentaire de réaliser la validation des modèles de prédiction en complément de la validation des scores. Dans le cas des scores NZDCS et ADVANCE, dont les modèles prédictifs dérivent de modèles de Cox, l'information utile pour le pronostic est le terme linéaire du modèle qui dépend des caractéristiques du sujet, appelé par certains auteurs le score pronostique (*prognostic index*,

PI)[10,73]. Lorsque la fonction de risque de base est disponible dans les publications, l'estimation du risque prédit résulte en l'application de la fonction de risque à chaque individu à un temps donné. Dans le cas contraire, estimer le PI d'un sujet permet d'approcher son pronostic et de le situer au sein du « spectre » pronostique défini par le modèle. Le risque d'évènement augmente avec la valeur du PI, une valeur élevée de PI traduit un risque plus important d'évènement soit une survie plus courte sans évènement alors qu'une une valeur faible de PI traduit un moindre risque d'évènement soit une survie plus longue sans évènement[74]. Il convient ensuite de créer des groupes de risque, puis d'évaluer la calibration et la discrimination du modèle. La calibration peut être mesurée par la valeur de la pente de la régression des PI et la visualisation des courbes de survie dans les différents groupes de risque. Dans le cas de données censurées, la discrimination peut être déterminée par des statistiques spécifiques telles que le c-index de Harell, la statistique K de Gönen et Heller, la statistique D de Royston-Sauerbrei ou encore le score de Brier[73].

Les principaux évènements considérés pour prédire la progression de la MRC sont l'apparition le doublement ou l'augmentation de la créatininémie, le passage au stade d'IRCT par la mise en dialyse ou du fait d'une transplantation rénale, ou la survenue d'un décès de cause rénale D'autres marqueurs pourraient cependant être pris en compte pour définir la progression de la MRC. En effet, le déclin de la fonction rénale dans le DT2 a été défini dans de précédents travaux par la diminution absolue du DFGe déterminée par une valeur seuil de pente annuelle du DFGe[7,75,76]. De récentes études ont proposé la diminution relative du niveau du DFGe comme alternative à l'évènement doublement de la créatininémie pour étudier la progression de la MRC chez des patients DT2[77,78]. Par ailleurs, certains auteurs interrogent la pertinence du doublement de la créatinémie en raison de sa forte dépendance à la fonction rénale initiale, au délai de survenue du doublement, et au niveau de créatininémie considéré comme initial mais également parce que les variations de créatininémie peuvent être dues à diverses états pathologiques aigüs[79]. La fiabilité de l'évènement décès d'origine rénale est également critiquée dans la même étude, principalement à cause de la difficulté de la définir et d'imputer réellement un décès à une cause rénale spécifique. Un décès dont la cause vraie est attribuable au rein ne peut, par définition, qu'être observé chez des patients en IRCT n'ayant pas pu bénéficier d'une méthode de suppléance rénale (dialyse ou transplantation), de telles situations restent rares en pratique clinique courante et le sont encore plus dans le cadre d'essais cliniques. Les états de MRC moins avancés ne sont généralement pas des causes directes de décès et la mortalité observée chez les patients en stade préterminal sont difficilement imputable de manière certaine

à une cause rénale. Par ailleurs, certains auteurs suggèrent de considérer les décès d'origine non rénale comme de potentiels évènements compétiteurs lors de l'estimation du risque d'IRCT[80,81]. Le risque de mort toute cause et de mort cardiovasculaire augmente avec la sévérité de la MRC et la maladie CV est un facteur de risque connu de décès avant l'apparition du stade terminal de la MRC[82,83], il paraît donc plus pertinent et plus juste de considérer les décès toute cause comme évènement d'intérêt lors de la prédiction d'évènements rénaux majeurs, au même titre que les critères habituels doublement de la créatininémie et apparition d'une IRCT. La notion de risque compétitif revient à prendre en compte dans les analyses les évènements qui empêchent l'évènement étudié d'avoir lieu. Les évènements compétiteurs doivent être considérés dès la conception des modèles prédictifs, le modèle de risque cause spécifique (*cause-specific hazards models*, combinaison de deux modèles de Cox)[84], et le modèle de Fine et Gray[85] sont les deux principales méthodes disponibles.

4.3 Mise en perspective des résultats de l'étude

L'évaluation de la comparabilité relative des échantillons de développement et de validation d'un score est essentiel pour interpréter les résultats d'une étude de validation externe et contribue à juger de sa « transportabilité » (*transportability)*. Cette approche reste néanmoins d'ordre qualitatif et la comparaison informelle des échantillons demeure relative au sein d'un spectre allant de l'identité parfaite à la complète différence[68]. Dans cette étude, les populations DIABHYCAR et SURDIAGENE pouvaient être considérées comme plutôt différentes de la population DCS. Les patients DT2 de la cohorte SURDIAGENE paraissaient plus à risque avec une atteinte rénale plus marquée, plus d'ATCD CV et une survenue plus fréquente de l'évènement rénal composite IRCT fatale ou non. De même, les participants de l'essai DIABHYCAR semblaient plus graves puisque globalement plus âgés, hypertendus, et dont l'atteinte rénale était plus avancée. En comparaison avec les participants de l'étude ADVANCE, les populations DIABHYCAR et SURDIAGENE présentaient une comparabilité plutôt « intermédiaire ». En effet, les patients SURDIAGENE présentaient une atteinte rénale plus avancée mais moins de facteurs de risque CV ; ils également ont montré plus d'évènements rénaux majeurs au cours du suivi (doublement de la créatinine, IRCT fatale ou non). Pour un risque CV similaire, la population DIABHYCAR présentait une maladie rénale plus grave et la survenue d'évènements majeurs était plus fréquente que dans la population ADVANCE. Ces différents degrés de gravité dans la maladie diabétique et ses complications étaient

principalement liés aux contextes d'étude et aux critères de sélection des patients. Les participants de la cohorte DCS comportait des patients DT2 inclus et suivis en structures de soins primaires, moins graves, ne nécessitant pas la même prise en charge que les patients DT2 de la cohorte SURDIAGENE, inclus et suivis dans un centre hospitalier universitaire. De même, les sujets de la cohorte DCS ne peuvent pas être comparables aux participants sélectionnés de l'essai DIABHYCAR pour leur atteinte rénale avancée, notamment. Les différences entre les populations de l'essai ADVANCE, de la cohorte SURDIAGENE et de l'essai DIABHYCAR sont également expliquées par les recrutements diverses rencontrés dans les types d'étude correspondants. Par ailleurs, la présence d'une rétinopathie diabétique a probablement été sous-estimée et minimisée aux stades les moins graves dans l'étude DIABHYCAR. En effet, contrairement à la cohorte SURDIAGENE, il s'agissait de la présence d'un ATCD de photocoagulation laser pour rétinopathie diabétique et non de la présence d'une rétinopathie diabétique (formes simples, pré-prolifératives et prolifératives).

Bien qu'il soit difficile de prédire dans quelle mesure les hétérogénéités constatées entre les échantillons impactaient les performances des scores[68], il était intéressant de constater que les paramètres des scores n'étaient pas toujours prédictifs des évènements rénaux étudiés dans les populations de validation externe. Dans la population SURDIAGENE, la quasi-totalité des facteurs inclus dans les scores NZDCS et ADVANCE étaient associés à la survenue des évènements rénaux étudiés (huit facteurs sur dix et cinq facteurs sur six, respectivement) alors qu'ils l'étaient plus rarement dans la population DIABHYCAR (trois facteurs sur dix et trois facteurs sur six, respectivement). Globalement, les échantillons de développement et de validation des scores ne présentaient pas être le même niveau de risque initial. Cette notion est appelée par certains auteurs le *spectrum transportability*, soit la généralisation du score à un large spectre de sujets[70].

Dans cette étude, les prédictions de risque des évènements rénaux par les deux scores étudiés étaient très faibles, mais du même ordre de grandeur que les fréquences observées des évènements rénaux dans les échantillons de validation externe. La calibration est considérée comme meilleure lorsque les valeurs de prédictions encadrent de manière plus étroite les fréquences observées d'évènement[70]. Par exemple dans la population SURDIAGENE, la fréquence de l'évènement rénal d'intérêt était de 5,5% et les prédictions selon le score NZDCS étaient comprises entre 0,13% et 7,88% ; alors qu'avec le score ADVANCE, pour une fréquence d'évènement rénal de 10,5%, les prédictions étaient comprises entre 3,9% et 63,3% ; indiquant que le score NZDCS était mieux calibré que le score ADVANCE dans cet échantillon.

Un score peut être parfaitement calibré dans une population donnée et présenter par ailleurs une mauvaise discrimination, lorsque par exemple, l'étendue des prédictions est très petite (de l'ordre de 2%) ; ce qui est le cas lorsque l'échantillon de développement et de validation sont très semblables. A l'inverse, dans notre étude, les scores étaient globalement très discriminants et avaient systématiquement une plus faible calibration. Le défaut relatif de calibration des scores dans nos échantillons de validation était en partie lié aux différences de fréquence de survenue des évènements d'intérêt[24], à l'origine d'écarts non négligeables entre les prédictions et les cas effectivement observés. Cela pouvait être également dû un à manque de puissance, les échantillons étant relativement plus petits que les échantillons de développement des scores[70] ; certains auteurs recommandent un effectif au moins égal à 200 comprenant au moins 100 sujets avec évènements et 100 sujets sans évènements pour effectuer une validation externe[55].

Les performances discriminatives du score NZDCS dans les populations SURDIAGENE et DIABHYCAR étaient comparables voire un peu meilleures que celles retrouvées dans la cohorte de validation externe Diabetes Care Support Services (Statistiques C 0,96 et 0,92 vs 0,91, respectivement)[37]. La performance discriminative du score ADVANCE dans la cohorte SURDIAGENE était similaire à sa validation externe dans deux autres populations de soins primaires de DT2 (Statistiques C 0,88, 0,86 et 0,88)[36], alors que le score était retrouvé nettement moins performant chez les participants de l'essai DIABHYCAR (Statistique C 0,74). Globalement, les scores étudiés étaient plus performants dans des contextes de soins primaires et chez des sujets « vie réelle » inclus dans des cohortes que chez les participants d'essais cliniques.

4.4 Au-delà de la validation externe

La validation externe d'un modèle prédictif ou d'un score peut conduire à son adaptation, ou customisation, à des populations particulières afin d'en améliorer les performances (*updating*). Il existe plusieurs méthodes d'amélioration des performances prédictives d'un modèle[24]. L'amélioration de la calibration consiste, soit à modifier globalement la relation entre le score et le risque d'évènement en adaptant l'intercept du modèle à la population de validation, soit à modifier la part contributive des paramètres du modèle par ajustement des coefficients de régression du modèle à l'aide d'un facteur correctif unique. L'amélioration de la discrimination revient à ré-estimer les coefficients de régression dans la population de validation lorsqu'ils

sont trop différents de ceux de la population de développement. L'autre manière d'améliorer la discrimination est d'évaluer l'ajout de nouvelles variables dans le modèle de prédiction.

Après l'évaluation des performances prédictives d'un score ou d'un modèle, se pose la question de son utilité en pratique clinique (*clinical usefulness*). Des études d'impact de l'utilisation des scores de risque sur l'amélioration clinique et biologique des patient mais également sur des indicateurs tels que l'incidence des maladies, les durées de séjour à l'hôpital, ou les délais de diagnostic, et donc le coût des soins et la qualité de vie, sont nécessaires et indispensables.

4.5 Conclusion

Le score de risque rénal NZDCS présentait une discrimination très satisfaisante dans les populations de validation SURDIAGENE et DIABHYCAR mais surestimait le risque rénal. Le score ADVANCE montrait une moindre discrimination dans les deux populations, surestimait le risque rénal pour les patients SURDIAGENE et le sous-estimait pour les participants de l'essai DIABHYCAR. Le score NZDCS était globalement plus performant que le score ADVANCE pour prédire le risque l'apparition d'évènements rénaux majeurs à 5 ans pour les patients DT2 des études SURDIAGENE et DIABHYCAR. L'utilisation de ce score pour prédire les évènements rénaux majeurs serait pertinente mais ne pourra être recommandée qu'à condition d'en améliorer la calibration. De futures études d'optimisation des performances de calibration du score NZDCS sont attendues pour permettre son utilisation en pratique clinique courante pour prédire la survenue d'évènements rénaux majeurs à 5 ans chez les patients DT2.

5. Références

1. Guariguata, L. By the numbers: New estimates from the IDF Diabetes Atlas Update for 2012. *Diabetes Res. Clin. Pract.* **98,** 524–525 (2012).
2. Mendis, S. & World Health Organization. *Global status report on noncommunicable diseases 2014*. (2014).
3. Institut de veille sanitaire (France). *Prévalence et incidence du diabète, et mortalité liée au diabète en France: synthèse épidémiologique*. (Institut de veille sanitaire, 2010).
4. Mathers, C. D. & Loncar, D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Med* **3,** e442 (2006).
5. Romand, D. *Prise en charge du patient diabétique présentant une atteinte de la fonction rénale*. (Société de Néphrologie).
6. *Center for Disease Control and Prevention. National Chronic Kidney Disease Fact Sheet: general information and national estimates on chronic kidney disease in the United States, 2010.* (Center for Disease Control and Prevention, 2010).
7. Rossing, K. *et al.* Progression of nephropathy in type 2 diabetic patients. *Kidney Int.* **66,** 1596–1605 (2004).
8. Royston, P., Moons, K. G. M., Altman, D. G. & Vergouwe, Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ* **338,** b604–b604 (2009).
9. Moons, K. G. M., Altman, D. G., Vergouwe, Y. & Royston, P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* **338,** b606–b606 (2009).
10. Hill, C. *et al. Analyse statistique des données de survie*. (Médecines-Sciences Flammarion).

11. Greenland, S. & Finkle, W. D. A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses. *Am. J. Epidemiol.* **142,** 1255–1264 (1995).
12. Pigott, T. D. A Review of Methods for Missing Data. *Educ. Res. Eval.* **7,** 353–383 (2001).
13. Trigui, B., Barrier, A., Flahault, A. & Huguier, M. Facteurs pronostiques des cancers avancés du pancréas. Analyse multifactorielle et score prédictif de survie. in *Annales de chirurgie* **125,** 625–630 (Elsevier, 2000).
14. Assmann, G. Simple Scoring Scheme for Calculating the Risk of Acute Coronary Events Based on the 10-Year Follow-Up of the Prospective Cardiovascular Munster (PROCAM) Study. *Circulation* **105,** 310–315 (2002).
15. Clayton, T. C. Risk score for predicting death, myocardial infarction, and stroke in patients with stable angina, based on a large randomised trial cohort of patients. *BMJ* **331,** 869–0 (2005).
16. Pfister, R., Cairns, R., Erdmann, E. & Schneider, C. A. A clinical risk score for heart failure in patients with type 2 diabetes and macrovascular disease: An analysis of the PROactive study. *Int. J. Cardiol.* **162,** 112–116 (2013).
17. Amin, S. T. *et al.* Dynamic TIMI risk score for STEMI. *J. Am. Heart Assoc.* **2,** e003269 (2013).
18. Pocock, S. J. *et al.* A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. *Bmj* **323,** 75–81 (2001).
19. Lo, M. Y. *et al.* A Risk Score for Predicting Coronary Artery Disease in Women With Angina Pectoris and Abnormal Stress Test Finding. *Am. J. Cardiol.* **111,** 781–785 (2013).
20. Altman, D. G., Vergouwe, Y., Royston, P. & Moons, K. G. M. Prognosis and prognostic research: validating a prognostic model. *BMJ* **338,** b605–b605 (2009).

21. Altman, D. G. & Royston, P. What do we mean by validating a prognostic model? *Stat. Med.* **19,** 453–473 (2000).

22. Collins, G. S. *et al.* External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med. Res. Methodol.* **14,** 40 (2014).

23. Moons, K. G. M. *et al.* Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* **98,** 691–698 (2012).

24. Toll, D. B., Janssen, K. J. M., Vergouwe, Y. & Moons, K. G. M. Validation, updating and impact of clinical prediction rules: A review. *J. Clin. Epidemiol.* **61,** 1085–1094 (2008).

25. Eckardt, K.-U. *et al.* Evolving importance of kidney disease: from subspecialty to global health burden. *The Lancet* **382,** 158–169 (2013).

26. Collins, A. J. *et al.* US Renal Data System 2013 Annual Data Report. *Am. J. Kidney Dis.* **63,** A7

27. Fox, C. S. *et al.* Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. *The Lancet* **380,** 1662–1673 (2012).

28. Echouffo-Tcheugui, J. B. & Kengne, A. P. Risk Models to Predict Chronic Kidney Disease and Its Progression: A Systematic Review. *PLoS Med.* **9,** e1001344 (2012).

29. Stevens, P. E., Levin, A. & Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann. Intern. Med.* **158,** 825–830 (2013).

30. Hadjadj, S. *et al.* Prognostic Value of the Insertion/Deletion Polymorphism of the ACE Gene in Type 2 Diabetic Subjects: Results from the Non-Insulin-Dependent Diabetes, Hypertension, Microalbuminuria or Proteinuria, Cardiovascular Events, and Ramipril

(DIABHYCAR), Diabete de type 2, Nephropathie et Genetique (DIAB2NEPHROGENE), and Survie, Diabete de type 2 et Genetique (SURDIAGENE) studies. *Diabetes Care* **31,** 1847–1852 (2008).

31. Évaluation de la fonction rénale et de la protéinurie pour le diagnostic de la maladie rénale chronique chez l'adulte. Recommandations pour la pratique clinique. *Néphrologie Thérapeutique* **5,** 302–305 (2009).
32. Harjutsalo, V. & Groop, P.-H. Epidemiology and Risk Factors for Diabetic Kidney Disease. *Adv. Chronic Kidney Dis.* **21,** 260–266 (2014).
33. Keane, W. F. Risk Scores for Predicting Outcomes in Patients with Type 2 Diabetes and Nephropathy: The RENAAL Study. *Clin. J. Am. Soc. Nephrol.* **1,** 761–767 (2006).
34. Jardine, M. J. *et al.* Prediction of Kidney-Related Outcomes in Patients With Type 2 Diabetes. *Am. J. Kidney Dis.* **60,** 770–778 (2012).
35. Hayes, A. J., Leal, J., Gray, A. M., Holman, R. R. & Clarke, P. M. UKPDS Outcomes Model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. *Diabetologia* **56,** 1925–1933 (2013).
36. Desai, A. S. *et al.* Association between cardiac biomarkers and the development of ESRD in patients with type 2 diabetes mellitus, anemia, and CKD. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* **58,** 717–728 (2011).
37. Elley, C. R. *et al.* Derivation and Validation of a Renal Risk Score for People With Type 2 Diabetes. *Diabetes Care* **36,** 3113–3120 (2013).
38. Moons, K. G. M. *et al.* Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* **98,** 683–690 (2012).
39. Bang, H. *et al.* Screening for kidney disease in vascular patients: SCreening for Occult REnal Disease (SCORED) experience. *Nephrol. Dial. Transplant.* **24,** 2452–2457 (2009).

40. Kshirsagar, A. V. *et al.* A Simple Algorithm to Predict Incident Kidney Disease. *Arch. Intern. Med.* **168,** 2466 (2008).

41. Fox, C. S. *et al.* A multi-marker approach to predict incident CKD and microalbuminuria. *J. Am. Soc. Nephrol. JASN* **21,** 2143–2149 (2010).

42. Hippisley-Cox, J. & Coupland, C. Predicting the risk of Chronic Kidney Disease in Men and Women in England and Wales: prospective derivation and external validation of the QKidney® Scores. *BMC Fam. Pract.* **11,** 49 (2010).

43. Chien, K.-L. *et al.* A Prediction Model for the Risk of Incident Chronic Kidney Disease. *Am. J. Med.* **123,** 836–846.e2 (2010).

44. Halbesma, N. *et al.* Development and Validation of a General Population Renal Risk Score. *Clin. J. Am. Soc. Nephrol.* **6,** 1731–1738 (2011).

45. Ando, M., Yanagisawa, N., Ajisawa, A., Tsuchiya, K. & Nitta, K. A simple model for predicting incidence of chronic kidney disease in HIV-infected patients. *Clin. Exp. Nephrol.* **15,** 242–247 (2011).

46. Thakkinstian, A. *et al.* A simplified clinical prediction score of chronic kidney disease: a cross-sectional-survey study. *BMC Nephrol.* **12,** 45 (2011).

47. O'Seaghdha, C. M., Yang, Q., Wu, H., Hwang, S.-J. & Fox, C. S. Performance of a genetic risk score for CKD stage 3 in the general population. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* **59,** 19–24 (2012).

48. Alssema, M. *et al.* One risk assessment tool for cardiovascular disease, type 2 diabetes, and chronic kidney disease. *Diabetes Care* **35,** 741–748 (2012).

49. Kwon, K.-S. *et al.* A simple prediction score for kidney disease in the Korean population: Prediction of kidney disease. *Nephrology* **17,** 278–284 (2012).

50. Kent, D. M. *et al.* Progression risk, urinary protein excretion, and treatment effects of angiotensin-converting enzyme inhibitors in nondiabetic kidney disease. *J. Am. Soc. Nephrol. JASN* **18,** 1959–1965 (2007).

51. Wakai, K. *et al.* A scoring system to predict renal outcome in IgA nephropathy: from a nationwide prospective study. *Nephrol. Dial. Transplant. Off. Publ. Eur. Dial. Transpl. Assoc. - Eur. Ren. Assoc.* **21,** 2800–2808 (2006).

52. Goto, M. *et al.* A scoring system to predict renal outcome in IgA nephropathy: a nationwide 10-year prospective cohort study. *Nephrol. Dial. Transplant. Off. Publ. Eur. Dial. Transpl. Assoc. - Eur. Ren. Assoc.* **24,** 3068–3074 (2009).

53. Johnson, E. S., Thorp, M. L., Platt, R. W. & Smith, D. H. Predicting the risk of dialysis and transplant among patients with CKD: a retrospective cohort study. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* **52,** 653–660 (2008).

54. Landray, M. J. *et al.* Prediction of ESRD and death among people with CKD: the Chronic Renal Impairment in Birmingham (CRIB) prospective cohort study. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* **56,** 1082–1094 (2010).

55. Hallan, S. I. *et al.* Combining GFR and albuminuria to classify CKD improves prediction of ESRD. *J. Am. Soc. Nephrol. JASN* **20,** 1069–1077 (2009).

56. Tangri, N. *et al.* A predictive model for progression of chronic kidney disease to kidney failure. *JAMA J. Am. Med. Assoc.* **305,** 1553–1559 (2011).

57. Kengne, A. P. *et al.* Contemporary model for cardiovascular risk prediction in people with type 2 diabetes. *Eur. J. Cardiovasc. Prev. Rehabil.* **18,** 393–398 (2011).

58. for the ADVANCE Collaborative Group *et al.* The Framingham and UK Prospective Diabetes Study (UKPDS) risk equations do not reliably estimate the probability of cardiovascular events in a large ethnically diverse sample of patients with diabetes: the

Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation (ADVANCE) Study. *Diabetologia* **53,** 821–831 (2010).

59. Simmons, R. K. *et al.* Performance of the UK Prospective Diabetes Study Risk Engine and the Framingham Risk Equations in Estimating Cardiovascular Disease in the EPIC-Norfolk Cohort. *Diabetes Care* **32,** 708–713 (2009).
60. Tao, L., Wilson, E. C. F., Griffin, S. J. & Simmons, R. K. Performance of the UKPDS Outcomes Model for Prediction of Myocardial Infarction and Stroke in the ADDITION-Europe Trial Cohort. *Value Health* **16,** 1074–1080 (2013).
61. Marre, M. Effects of low dose ramipril on cardiovascular and renal outcomes in patients with type 2 diabetes and raised excretion of urinary albumin: randomised, double blind, placebo controlled trial (the DIABHYCAR study). *BMJ* **328,** 495–0 (2004).
62. Tangri, N. *et al.* Risk prediction models for patients with chronic kidney disease: a systematic review. *Ann. Intern. Med.* **158,** 596–603 (2013).
63. Harris, M. I., Klein, R., Welborn, T. A. & Knuiman, M. W. Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. *Diabetes Care* **15,** 815–819 (1992).
64. Levey, A. S., Stevens, L. A. & Schmid, C. H. A new equation to estimate glomerular filtration rate. *Ann. Intern. Med.* 604–612 (2009).
65. Rodby, R. A. *et al.* The urine protein to creatinine ratio as a predictor of 24-hour urine protein excretion in type 1 diabetic patients with nephropathy. The Collaborative Study Group. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* **26,** 904–909 (1995).
66. Steyerberg, E. W. *et al.* Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. *Epidemiology* **21,** 128–138 (2010).
67. Steyerberg, E. W. & Vergouwe, Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur. Heart J.* **35,** 1925–1931 (2014).

68. Debray, T. P. A. *et al.* A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J. Clin. Epidemiol.* (2014). doi:10.1016/j.jclinepi.2014.06.018

69. Tripepi, G., Jager, K. J., Dekker, F. W. & Zoccali, C. Statistical methods for the assessment of prognostic biomarkers(part II): calibration and re-classification. *Nephrol. Dial. Transplant.* **25,** 1402–1405 (2010).

70. Steyerberg, E. . in *Clinical Prediction Models: A Practical Approach to Developpement, Validation, and Updating* (2009).

71. Tripepi, G., Jager, K. J., Dekker, F. W. & Zoccali, C. Statistical methods for the assessment of prognostic biomarkers (Part I): Discrimination. *Nephrol. Dial. Transplant.* **25,** 1399–1401 (2010).

72. OECD. *Education at a Glance 2014*. (OECD Publishing, 2014). at <http://www.oecd-ilibrary.org/education/education-at-a-glance-2014_eag-2014-en>

73. Royston, P. & Altman, D. G. External validation of a Cox prognostic model: principles and methods. *BMC Med. Res. Methodol.* **13,** 33 (2013).

74. Christensen, E. Multivariate survival analysis using Cox's regression model. *Hepatol. Baltim. Md* **7,** 1346–1358 (1987).

75. Skupien, J. *et al.* The early decline in renal function in patients with type 1 diabetes and proteinuria predicts the risk of end stage renal disease. *Kidney Int.* **82,** 589–597 (2012).

76. De Hauteclocque, A. *et al.* The influence of sex on renal function decline in people with Type 2 diabetes. *Diabet. Med. J. Br. Diabet. Assoc.* **31,** 1121–1128 (2014).

77. Lambers Heerspink, H. J. *et al.* Estimated GFR Decline as a Surrogate End Point for Kidney Failure: A Post Hoc Analysis From the Reduction of End Points in Non–Insulin-Dependent Diabetes With the Angiotensin II Antagonist Losartan (RENAAL) Study and Irbesartan Diabetic Nephropathy Trial (IDNT). *Am. J. Kidney Dis.* **63,** 244–250 (2014).

78. Cooper, M. E. *et al.* Kidney Disease End Points in a Pooled Analysis of Individual Patient–Level Data From a Large Clinical Trials Program of the Dipeptidyl Peptidase 4 Inhibitor Linagliptin in Type 2 Diabetes. *Am. J. Kidney Dis.* (2015). doi:10.1053/j.ajkd.2015.03.024
79. Al-Aly, Z. Prediction of renal end points in chronic kidney disease. *Kidney Int.* **83,** 189–191 (2013).
80. Matsushita, K. *et al.* Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *The Lancet* **375,** 2073–2081 (2010).
81. Van der Velde, M. *et al.* Lower estimated glomerular filtration rate and higher albuminuria are associated with all-cause and cardiovascular mortality. A collaborative meta-analysis of high-risk population cohorts. *Kidney Int* **79,** 1341–1352 (2011).
82. Grams, M. E. & Coresh, J. Assessing risk in chronic kidney disease: a methodological review. *Nat. Rev. Nephrol.* **9,** 18–25 (2012).
83. Grams, M. E. *et al.* Vascular Disease, ESRD, and Death: Interpreting Competing Risk Analyses. *Clin. J. Am. Soc. Nephrol.* **7,** 1606–1614 (2012).
84. Prentice, R. L. *et al.* The analysis of failure times in the presence of competing risks. *Biometrics* **34,** 541–554 (1978).
85. Fine, J. P. & Gray, R. J. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J. Am. Stat. Assoc.* **94,** 496–509 (1999).

6. Résumé

Objectifs

Des scores prédictifs permettant l'identification précoce des patients diabétiques de type 2 (DT2) à risque d'insuffisance rénale chronique terminale (IRCT) existent. La confirmation de leurs performances prédictives dans d'autres contextes populationnels reste nécessaire. L'objectif principal était de réaliser la validation externe indépendante des deux scores de risque rénal, New Zealand Diabetes Cohort Study (NZDCS) et Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation Study (ADVANCE), chez les patients DT2 des études SURDIAGENE et DIABHYCAR. L'objectif secondaire était de comparer leurs performances prédictives dans ces populations.

Matériels et Méthodes

Les équations de risque ont permis d'estimer le risque rénal à 5 ans des patients. Les évènements rénaux étaient doublement de la créatininémie, IRCT fatale ou non pour le score ADVANCE ; IRCT fatale ou non pour le score NZDCS. La calibration a été évaluée par comparaison des prédictions et des observations dans les déciles de risque, l'adéquation estimée par le test de Hosmer et Lemeshow, et la discrimination mesurée par l'aire sous la courbe (ASC) ROC des modèles exprimant la prédiction des évènements en fonction de leur survenue.

Résultats

Score NZDCS : 46 (5.5%) et 19 (1.5%) évènements rénaux sont survenus dans les populations SURDIAGENE et DIABHYCAR. Le risque d'évènement était surestimé et la discrimination excellente dans SURDIAGENE (ASC=0,96 [0,92-0,99]) et dans DIABHYCAR (ASC=0,92 [0,86-0,97]).

Score ADVANCE : 90 (10,5%) et 76 (5,7%) évènements rénaux sont survenus dans les populations SURDIAGENE et DIABHYCAR. Le risque d'évènement était surestimé dans SURDIAGENE et sous-estimé dans DIABHYCAR. La discrimination était meilleure dans SURDIAGENE que dans DIABHYCAR (ASC=0,88 [0,84-0,92] et ASC=0,74 [0,68-0,80], respectivement).

Conclusion

Le score NZDCS présentait de meilleures performances discriminatives que le score ADVANCE chez les participants des études SURDIAGENE et DIABHYCAR. L'utilisation du score NZDCS pour prédire les évènements rénaux majeurs chez les patients DT2 serait pertinente mais ne pourra être recommandée qu'à condition d'en améliorer la calibration.

Mots-clefs : insuffisance rénale chronique terminale, diabète de type 2, modèle prédictif, score de risque, validation externe, calibration, discrimination

7. Serment

En présence des Maîtres de cette école, de mes chers condisciples et devant l'effigie d'Hippocrate, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la médecine. Je donnerai mes soins gratuits à l'indigent et n'exigerai jamais un salaire au-dessus de mon travail. Admis dans l'intérieur des maisons mes yeux ne verront pas ce qui s'y passe ; ma langue taira les secrets qui me seront confiés, et mon état ne servira pas à corrompre les mœurs ni à favoriser le crime. Respectueux et reconnaissant envers mes Maîtres, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses ! Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque

Résumé

Objectifs

Des scores prédictifs permettant l'identification précoce des patients diabétiques de type 2 (DT2) à risque d'insuffisance rénale chronique terminale (IRCT) existent. La confirmation de leurs performances prédictives dans d'autres contextes populationnels reste nécessaire. L'objectif principal était de réaliser la validation externe indépendante des deux scores de risque rénal, New-Zealand Diabetes Cohort Study (NZDCS) et Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation Study (ADVANCE), chez les patients DT2 des études SURDIAGENE et DIABHYCAR. L'objectif secondaire était de comparer leurs performances prédictives dans ces populations.

Matériels et Méthodes

Les équations de risque ont permis d'estimer le risque rénal à 5 ans des patients. Les évènements rénaux étaient doublement de la créatininémie, IRCT fatale ou non pour le score ADVANCE ; IRCT fatale ou non pour le score NZDCS. La calibration a été évaluée par comparaison des prédictions et des observations dans les déciles de risque, l'adéquation estimée par le test de Hosmer et Lemeshow, et la discrimination mesurée par l'aire sous la courbe (ASC) ROC des modèles exprimant la prédiction des évènements en fonction de leur survenue.

Résultats

Score NZDCS : 46 (5.5%) et 19 (1.5%) évènements rénaux sont survenus dans les populations SURDIAGENE et DIABHYCAR. Le risque d'évènement était surestimé et la discrimination excellente dans SURDIAGENE (ASC=0,96 [0,92-0,99]) et dans DIABHYCAR (ASC=0,92 [0,86-0,97]).

Score ADVANCE : 90 (10,5%) et 76 (5,7%) évènements rénaux sont survenus dans les populations SURDIAGENE et DIABHYCAR. Le risque d'évènement était surestimé dans SURDIAGENE et sous-estimé dans DIABHYCAR. La discrimination était meilleure dans SURDIAGENE que dans DIABHYCAR (ASC=0,88 [0,84-0,92] et ASC=0,74 [0,68-0,80], respectivement).

Conclusion

Le score NZDCS présentait de meilleures performances discriminatives que le score ADVANCE chez les participants des études SURDIAGENE et DIABHYCAR. L'utilisation du score NZDCS pour prédire les évènements rénaux majeurs chez les patients DT2 serait pertinente mais ne pourra être recommandée qu'à condition d'en améliorer la calibration.

Mots-clefs : insuffisance rénale chronique terminale, diabète de type 2, modèle prédictif, score de risque, validation externe, calibration, discrimination