DISCLAIMER

- **This includes use for the following relevant business purposes:**
  - **Communication and storage within Astellas for business and/or scientific purposes**
  - **Preparation and use of training materials within Astellas**
  - **External communications to regulatory authorities**
  - **External meetings, such as advisory board, steering committee, or investigator meetings (provided that a confidential disclosure agreement is in place with attendees prior to use)**
  - **Responses by Medical Science Liaison and/or Medical Information Department to unsolicited requests by health care providers (as allowed by local regulations)**

Poster number: PO2340

A machine learning algorithm to identify patients with possible non-dialysis-dependent chronic kidney disease

Thierry Lobbedez[1], Karim Dardim[2], Arnaud Panes[3], Gwendoline Poinsot[3], Jérôme Fernandes[4], Laurence Dubel[5], Josephine Wolfram[6]

[1]University Hospital, Caen, France; [2]Association Limousine pour l'Utilisation du Rein Artificiel à Domicile (ALURAD), Limoges, France; [3]HEVA, Lyon, France; [4]Centre Hospitalier de la Côte Basque, Bayonne, France; [5]Astellas Pharma Europe, Levallois-Perret, France; [6]Astellas Pharma Europe B.V., Leiden, The Netherlands

KIDNEY WEEK 2021

FINANCIAL DISCLOSURES

Presenter: Thierry Lobbedez

Disclosure: no conflicts of interest to declare

Co-author: Karim Dardim

Disclosure: no conflicts of interest to declare

Co-author: Arnaud Panes

Disclosure: no conflicts of interest to declare

Co-author: Gwendoline Poinsot

Disclosure: no conflicts of interest to declare

Co-author: Jérôme Fernandes

Disclosure: no conflicts of interest to declare

Co-author: Laurence Dubel

Disclosure: employee of Astellas Pharma Europe

Co-author: Josephine Wolfram

Disclosure: employee of Astellas Pharma Europe B.V.

Funding statement: This study was sponsored by Astellas Pharma Inc. Medical writing support was provided by Iona Easthope, DPhil, of Cello Health MedErgy, funded by Astellas Pharma Inc.

BACKGROUND/OBJECTIVE

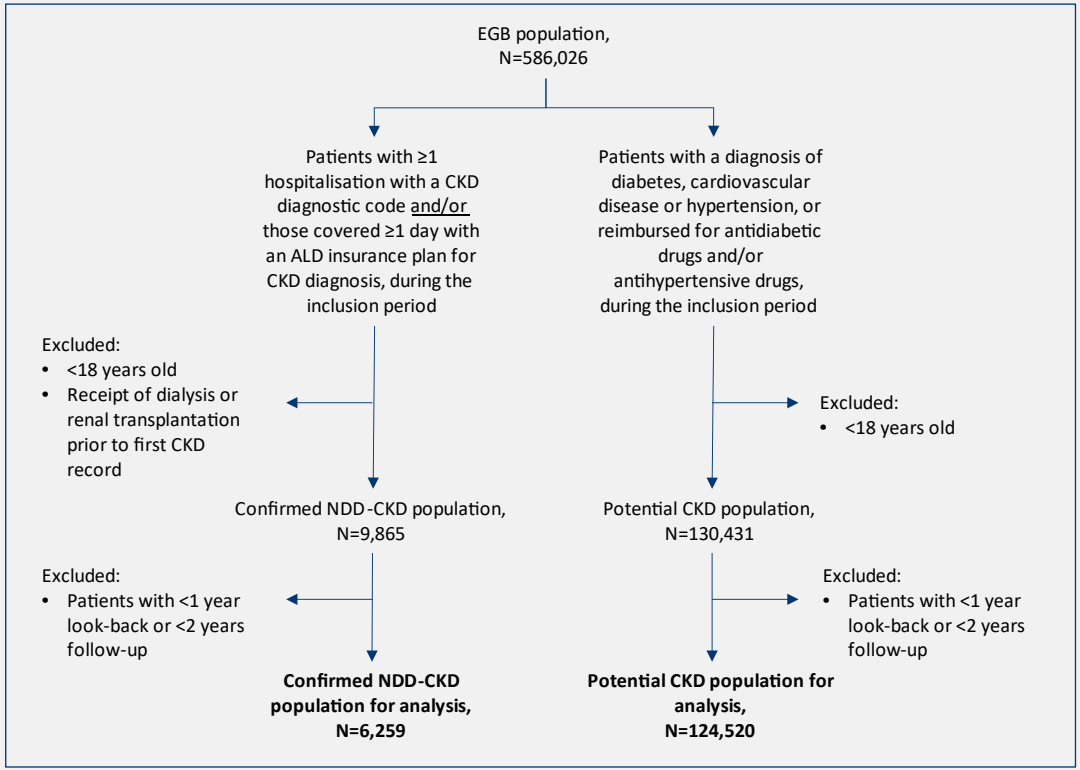
- There is a lack of data describing the proportion of patients with non-dialysis-dependent chronic kidney disease (NDD-CKD) in France
- The DAKOTAH study was designed to estimate the prevalence and incidence of NDD-CKD, among other objectives, in the French population
- As part of the study, we aimed to **identify patients from the general population with possible NDD-CKD**, who did not have a recorded ICD-10 diagnosis of CKD, **by developing a machine learning algorithm**

ICD-10, International Classification of Diseases 10th Revision

SELECTION OF PATIENT POPULATIONS

- DAKOTAH was a retrospective, non-interventional study of NDD-CKD patients, using data from the Echantillon Généraliste des Bénéficiaires (EGB) database, which is a sample of the French population protected by health insurance
- The study inclusion period was January 01, 2012 to December 31, 2017
- Two initial populations of interest were identified directly from the EGB data:
  - 'Confirmed NDD-CKD'
  - 'Potential CKD' (patients who were considered at high risk of CKD)

Selection of confirmed NDD-CKD and potential CKD populations



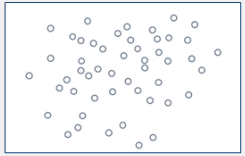
ALD, affection de longue durée; CKD, chronic kidney disease; ICD-10, International Classification of Diseases 10th Revision; NDD, non-dialysis-dependent

STEP 1: SELECTION OF POTENTIAL CKD PATIENTS

- The potential CKD population was split into two parts for:
  - algorithm training
  - algorithm testing

Development of an algorithm to detect patients with possible CKD

STEP 1 **Selection of all potential CKD patients**



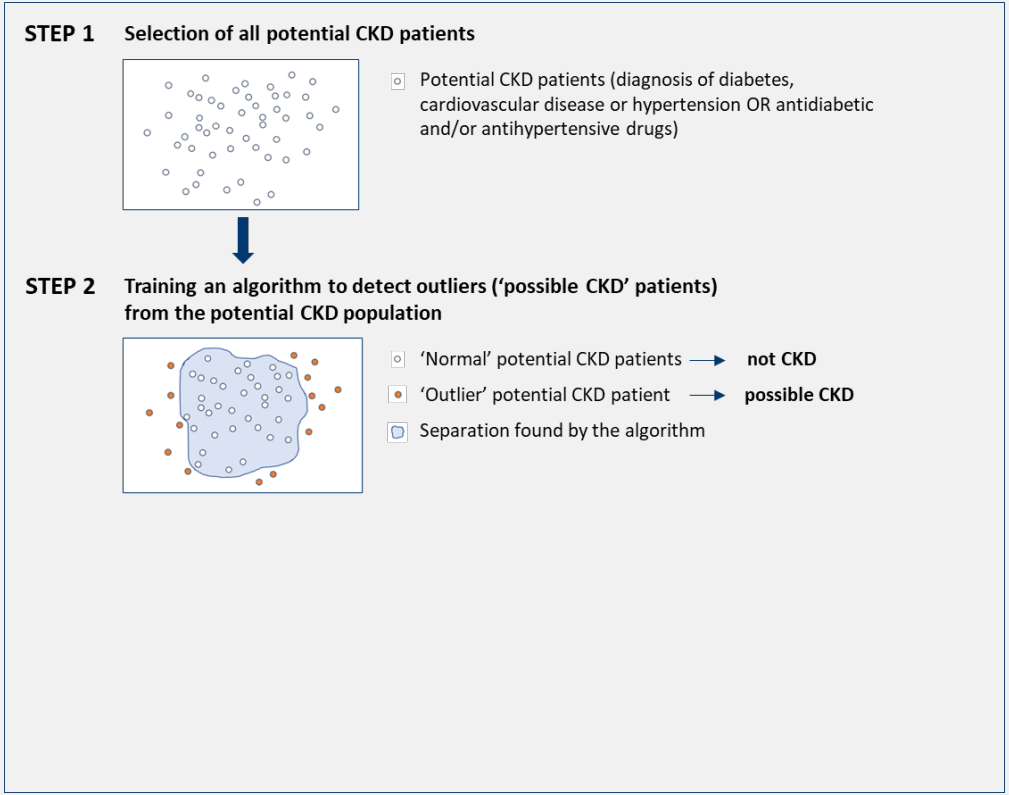
Potential CKD patients (diagnosis of diabetes, cardiovascular disease or hypertension OR antidiabetic and/or antihypertensive drugs)

CKD, chronic kidney disease

STEP 2: ALGORITHM TRAINING

- A ‘possible CKD’ population was extracted from the potential CKD population using an unsupervised machine learning algorithm*
- A distance metric between patients was defined based on variables that may be associated with CKD:
  - Sex
  - Number and duration of hospitalizations for renal failure
  - Number of general practitioner visits
  - Medications
  - Number/type of biological examinations
- Using the distance metric, patients having similar characteristics were positioned close to one another
- The algorithm learned to construct a spherical boundary around the non-CKD population, to create a decision rule for possible CKD versus non-CKD. Outliers were considered possible CKD

Development of an algorithm to detect patients with possible CKD

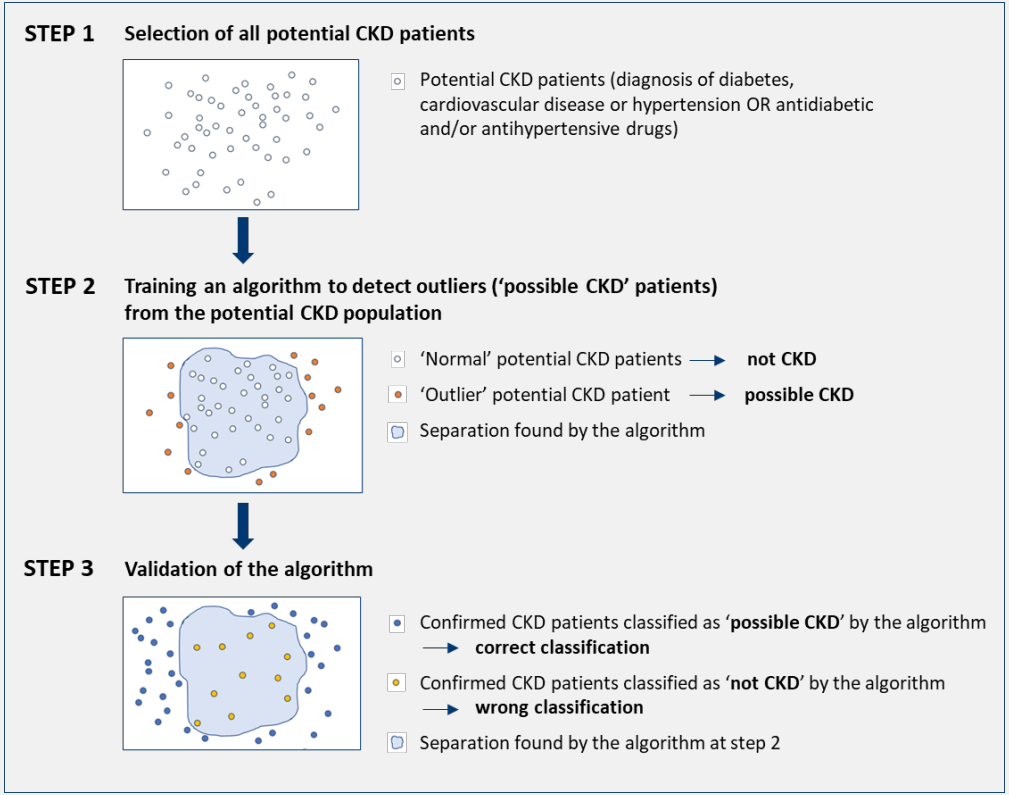


*One Class Support Vector Machine [SVM]; Python 0.23.2. CKD, chronic kidney disease

STEP 3: ALGORITHM TUNING AND VALIDATION (1)

- The algorithm was applied to both the potential and the confirmed CKD populations for analysis
- The following assumptions were made:
  - Most of the potential CKD population does not have CKD, so the proportion of patients classified by the algorithm as possible CKD from the potential population should be low
  - The proportion of patients classified as possible CKD by the algorithm from the confirmed CKD population should be high

Development of an algorithm to detect patients with possible CKD

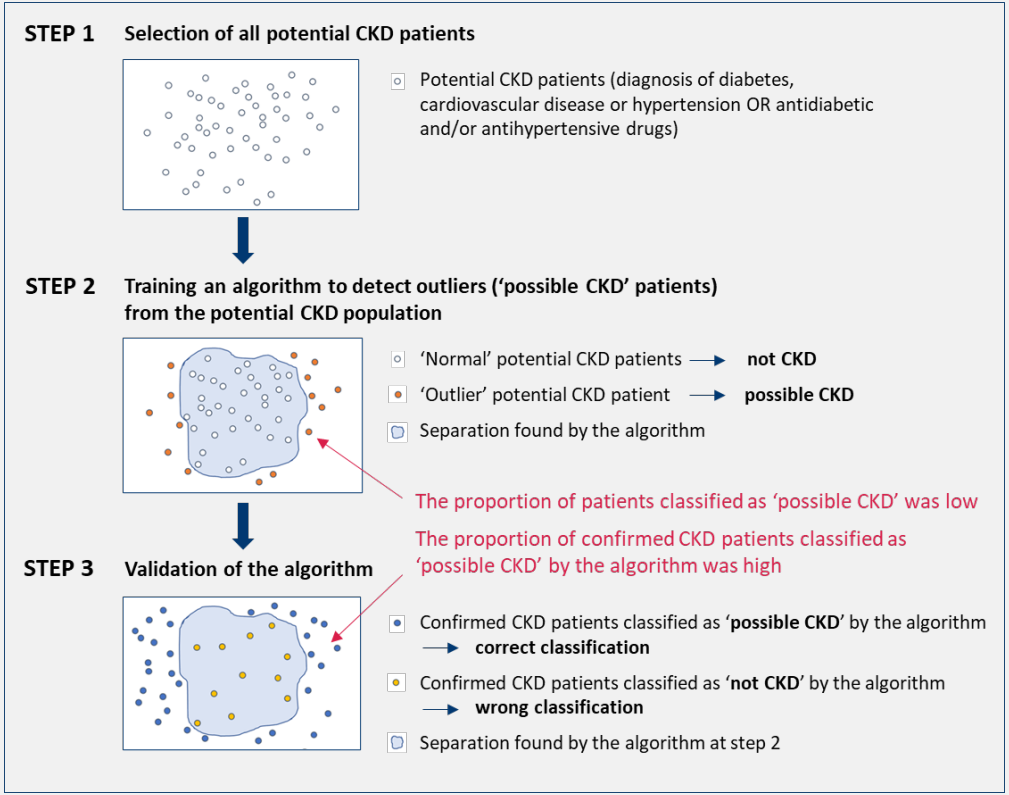


CKD, chronic kidney disease

STEP 3: ALGORITHM TUNING AND VALIDATION (2)

- From the potential CKD population for analysis, the algorithm detected 21% (26,064/124,520) as having possible CKD
  - The size of this possible CKD population (26,064) was almost three-fold greater than the size of the confirmed NDD-CKD population (9,865)
- From the confirmed CKD population for analysis, the algorithm classified 64.6% (4,044/6,259) as having possible CKD

Development of an algorithm to detect patients with possible CKD

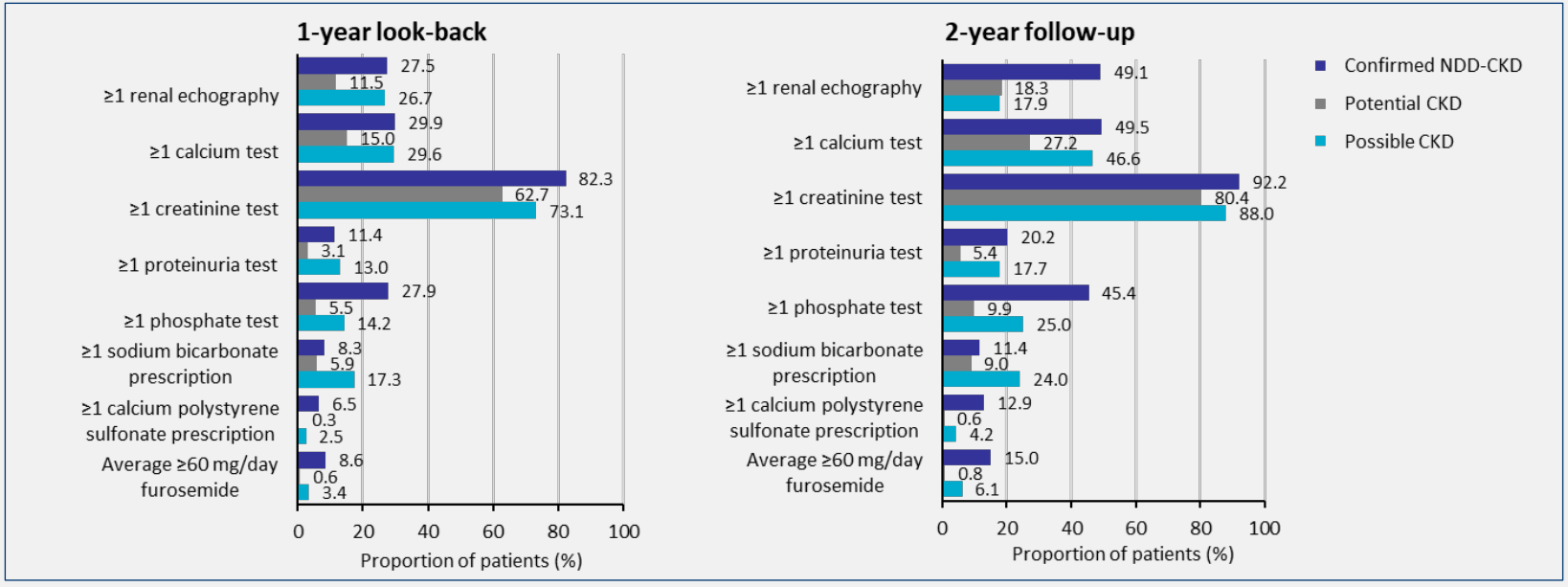


CKD, chronic kidney disease; NDD, non-dialysis-dependent

STEP 3: ALGORITHM TUNING AND VALIDATION (3)

- There were some similarities between the confirmed and possible CKD populations in relation to drug prescriptions, laboratory tests and biological examinations

Select drug prescriptions, laboratory tests and biological examinations, in the confirmed, potential and possible CKD populations



CKD, chronic kidney disease; NDD, non-dialysis-dependent

CONCLUSIONS

From a population of 124,520 patients with potential CKD, 21% (n=26,064) were identified as having possible CKD, using a machine learning algorithm

The number of patients with possible CKD was almost three-fold higher than the number recorded with confirmed CKD using hospital codes or ICD-10 diagnostic codes

This machine learning-derived decision rule could be a tool to identify undiagnosed patients with NDD-CKD

CKD, chronic kidney disease; ICD-10, International Classification of Diseases 10th Revision; NDD, non-dialysis-dependent