Validation of a complex algorithm for the diagnosis of metastatic castration-resistant prostate cancer within a claims database

Nicolas Thurienn, 1,2,3, Magali Rouyer, 1, Marine Gross-Goupil, 3, Michel Souleté, 4, Mathieu Roumiguie, 4, Sylvère Le Mouël, 5, Ludovic De Beaucoudrey, 6, Stéphanie Lamargue, 1, Emmanuelle Bignon, 1, Jérémy Jové, 1, Cécile Droz-Perreault, 1, Nicholas Moore 1,2,3, Patrick Blin 1

1 Bordeaux PharmacoEpi, INSERM U1241, Université de Bordeaux, Bordeaux, France – 2 INSERM U1219, Bordeaux, France – 3 CHU de Bordeaux, Bordeaux, France – 4 CHU de Toulouse, Toulouse, France – 5 Clinique Marcel, Pau, France – 6 Janssen,issy les Moulineaux, France

Abstract

Background: An algorithm was developed in the French nation-wide claims database (SNDS) to identify cases of metastatic castration-resistant prostate cancer (mCRPC). The usual way to validate such an algorithm is to review patients’ medical charts. An alternative in an impersonal pseudonymised database is to use all healthcare use information to identify diagnosis and/or treatment of prostate cancer, then resistant and metastatic stages.

Objectives: To assess and validate mCRPC algorithm using the wealth of data available in SNDS.

Methods: 100 of 14 050 mCRC patients identified by the algorithm and 100 of 372 273 non-mCRC patients were randomly selected within SNDS. The 6-year medical history of each of these 200 patients was reconstituted (long-term disease registration [LTD], drug dispensions, procedure codes, hospitalisations, lab tests). These 200 cases were randomly divided into 2 groups of 100 cases. Two groups of independent experts including an urologist and an oncologist each adjudicated blindly the mCRC status of 100 cases. In case of disagreement within a pair of experts, the 4 experts collectively assessed the case. Positive (PPV) and negative (NPV) predictive values of the algorithm were calculated.

Results: 110 of 100 mCRPC cases and 51 out of 100 non-mCRPC cases were concordant between the experts and the algorithm, resulting in an algorithm PPV of 0.92 and a NPV of 0.93.

Conclusions: The wealth of data available in the SNDS makes it possible to implement algorithms to detect complex diseases, and to validate them via the reconstitution of medical history. The present results show good performance of the algorithm for the identification of mCRC in the SNDS. In addition, the validation study detected some parameters that could be used to optimize the algorithm’s performance.

Declaration of Interest Statement

The CAMERRA study is carried out by the Bordeaux PharmacoEpi platform in collaboration with Janssen company and supervised by a scientific committee.

Methods

Validation Committee organisation (Figure 2)

- Random selection of 100 mCRC cases and 100 non-mCRC prostate cancer cases following the algorithm execution
- Constitution of 2 experts panels (1 urologist + 1 oncologist)
- Blind review of the 200 cases: mCRC and non-mCRC patients
- 50 random mCRC cases + 50 random non-mCRC cases, per expert pair (Figures 3 & 4)
- In case of disagreement within a pair of experts: the case was discussed by the 4 experts to reach consensus.
- Calculation of PPV and NPV of the algorithm based on committee conclusions and algorithm results.

Figure 2. Validation Committee for CAMERRA algorithm

Results

- PPV and NPV calculation (Figure 5)
  - Confusion of 92 out of 100 mCRPC cases and 93 out of 100 non-mCRPC cases
  - PPV = 0.92 and NPV = 0.93

Conclusion

- The wealth of data available in the SNDS enables the implementation of algorithms to detect complex diseases
- The validation of these algorithms via the reconstruction of pseudonymized medical charts based on SNDS data
- Here, the validation study shows good performances of the algorithm for mCRC identification
- Allows to adjust some parameters to optimize the algorithm performances
- Provides a validation algorithm generating accurate estimation of the number of mCRC cases in France, as well as a description of their characteristics and therapeutic changes