Original article

Agreement between hospital discharge diagnosis codes and medical records to identify metastatic colorectal cancer and associated comorbidities in elderly patients

Concordance entre les codes de diagnostic du PMSI et les données de dossiers médicaux pour identifier le cancer colorectal métastatique et les comorbidités associées chez des patients âgés

A. Gouverneur [a,b,c,d,*], D. Dolatkhani [a,b], M. Rouyer [d,e], A. Grelaud [d,e], F. Francis [a,b],
V. Gilleron [a], A. Fourrier-Réglat [a,b,c,d], P. Noize [a,c,d]

[a] CHU de Bordeaux, 33000 Bordeaux, France
[b] University de Bordeaux, 33076 Bordeaux cedex, France
[c] Inserm U1219, 33076 Bordeaux cedex, France
[d] Inserm CIC1401, 33076 Bordeaux cedex, France
[e] ADERA, 33608 Pessac, France

Abstract

Background. – Quality of coding to identify cancers and comorbidities through the French hospital diagnosis database (*Programme de médicalisation des systèmes d'information*, PMSI) has been little investigated. Agreement between medical records and PMSI database was evaluated regarding metastatic colorectal cancer (mCRC) and comorbidities.

Methods. – From 01/01/2013 to 06/30/2014, 74 patients aged $\geq$ 65 years at mCRC diagnosis were identified in Bordeaux teaching hospital. Data on mCRC and comorbidities were collected from medical records. All diagnosis codes (main, related and associated) registered into the PMSI were extracted. Agreement between sources was evaluated using the percent agreement for mCRC and the kappa (κ) statistic for comorbidities.

Results. – Agreement for primary CRC and mCRC was higher using all types of diagnosis codes instead of the main one exclusively (respectively 95% vs. 53% for primary CRC and 91% vs. 24% for mCRC). Agreement was substantial (κ 0.65) for cardiovascular diseases, notably atrial fibrillation (κ 0.77) and hypertension (κ 0.68). It was moderate for psychiatric disorders (κ 0.49) and respiratory diseases (κ 0.48), although chronic obstructive pulmonary disease had a good agreement (κ 0.75). Within the class of endocrine, nutritional and metabolic diseases (κ 0.55), agreement was substantial for diabetes (κ 0.91), obesity (κ 0.82) and hypothyroidism (κ 0.72) and moderate for hypercholesterolemia (κ 0.51) and malnutrition (κ 0.42).

Conclusion. – These results are reassuring with regard to detection through PMSI of mCRC if all types of diagnosis codes are considered and useful to better choose comorbidities in elderly mCRC patients that could be well identified through hospital diagnosis codes.
© 2017 Elsevier Masson SAS. All rights reserved.

Keywords: Agreement; International Classification of Diseases (ICD-10) codes; Medical records; Colorectal neoplasm; Neoplasm metastasis; Comorbidity

Résumé

Position du problème. – La qualité du codage pour l'identification des cancers et de comorbidités dans la base de données hospitalière française du Programme de médicalisation des systèmes d'information (PMSI) a été peu étudiée. La concordance entre les données de dossiers médicaux et celles de la base de données du PMSI a été évaluée pour le cancer métastatique colorectal (CCRm) et les comorbidités chez des sujets âgés.

Méthodes. – Du 01/01/2013 au 30/06/2014, 74 patients âgés d'au moins 65 ans au diagnostic d'un CCRm ont été identifiés au centre hospitalo-universitaire de Bordeaux. Les données sur le CCRm et les comorbidités ont été recueillies à partir des dossiers médicaux. Tous les codes de

diagnostic (principal, rélié et associés) enregistrés dans le PMSI pour les séjours de ces patients ont été extraits. La concordance entre les sources a été évaluée en utilisant le pourcentage de concordance pour le CCRm et le coefficient kappa (κ) pour les comorbidités.

Résultats. – La concordance pour le CCR primitif et pour le CCRm était plus élevée en utilisant tous les types de codes de diagnostic qu'uniquement le code principal (respectivement 95 % vs 53 % pour le CCR primitif et 91 % vs 24 % pour le CCRm). La concordance était substantielle (κ 0,65) pour les maladies cardiovasculaires, notamment la fibrillation auriculaire (κ 0,77) et l'hypertension (κ 0,68). Elle était modérée pour les troubles psychiatriques (κ 0,49) et les maladies respiratoires dans leur ensemble (κ 0,48), bien que la broncho-pneumopathie chronique obstructive présentait une bonne concordance (κ 0,75). Dans la classe des maladies endocriniennes, nutritionnelles et métaboliques (κ 0,55), la concordance était importante pour le diabète (κ 0,91), l'obésité (κ 0,82) et l'hypothyroïdie (κ 0,72) et modérée pour l'hypercholesté-rolémie (κ 0,51) et la malnutrition (κ 0,42).

Conclusion. – Ces résultats sont rassurants en ce qui concerne la détection des CCRm dans le PMSI, lorsque tous les types de codes de diagnostic sont considérés et utiles pour mieux choisir les comorbidités potentiellement identifiables via le PMSI chez des patients âgés atteints de CCRm.
© 2017 Elsevier Masson SAS. Tous droits réservés.

Mots clés : Concordance ; Codes de la classification internationale des maladies (CIM-10) ; Dossiers médicaux ; Cancer colorectal ; Métastases de cancer ; Comorbidités

1. Introduction

Colorectal cancer (CRC) is among the most common cancers worldwide and the fourth more frequent in France [1,2]. It is more frequent among the elderly; nearly 60% of incident cases are diagnosed in patients aged ≥ 65 years [1]. Moreover, approximately 40% will have metastatic disease (mCRC) [3]. With the population ageing, the burden of mCRC will increase in the future and population-based pharmacoepidemiological evaluations of innovative medications developed in this indication will be required.

Research using French administrative healthcare databases is developing [4,5]. The *Système national d'information inter-régimes de l'assurance maladie* (SNIIR-AM) is covering almost all the French population and records reimbursement data for outpatient care (including prescribed medications); the *Échantillon généraliste des bénéficiaires* (EGB) is a 1/97th permanent and representative sample of the SNIIR-AM. Both databases are now linked to the national hospital discharge diagnosis database (*Programme de médicalisation des systèmes d'information*, PMSI). This record-linkage was successfully used to conduct studies on utilization [6], benefit [7] and risk [8] of medications in the French population.

Data from the PMSI could be useful to identify patients with mCRC and comorbidities that are important to consider in observational studies. Before conducting such studies, the quality of coding of medical events must be investigated. This was performed for specific events, like acute coronary syndrome [9] or infections [10] but, to our knowledge, not for mCRC or for various comorbidities in a population. As part of an observational study on treatment practices in elderly patients with mCRC [11], agreement between data collected from medical records and recorded in the PMSI was evaluated regarding diagnosis of mCRC and comorbidities.

2. Methods

2.1. Study population

The original study has been described elsewhere [11]. Briefly, this was an observational field study conducted in the Bordeaux teaching hospital (Southwest France, around 2800 beds). From January 2013 to June 2014, 78 patients were identified from an ad hoc registry of multidisciplinary meetings held for cancer patients according to the following criteria: CRC, metastatic disease, age ≥ 65 years at mCRC diagnosis, first presentation in multidisciplinary meetings for mCRC during the recruitment period and follow-up at the Bordeaux teaching hospital.

Only patients hospitalized at least once during the study period, thus susceptible to have data recorded in the PMSI were concerned by the present work.

2.2. Data sources

The first source was the medical records of patients [11]. All available records in medical charts were reviewed to confirm the diagnosis of mCRC and collect data concerning the disease (e.g. date of diagnosis, localization of primary tumour and metastases) and the presence of comorbidities at mCRC diagnosis using a standardized questionnaire.

The second source was the PMSI database. In France, for each hospitalization, a summary of clinical data (including comorbidities) is produced and coded by physicians or specialized coding associates using the International Classification of Diseases 10th revision (ICD-10). Data can be encoded in three classes: one main diagnosis which motivates the hospitalization, one diagnosis related to the main diagnosis and associated diagnoses corresponding to existing comorbidities or medical events occurring during the hospitalization. All types of diagnosis codes (main, related and associated) registered at the first hospitalization for mCRC into the PMSI (including hospitalization for chemotherapy) were extracted for each included patient. The first hospitalization was defined as the hospitalization that corresponded to the diagnosis date identified in medical records or the first hospitalization further to the diagnosis date.

2.3. Analyses

To allow the comparison between both sources, each comorbidity collected in medical records was associated with