VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Identifying Psoriasis and Psoriatic Arthritis Patients in Retrospective Databases When Diagnosis Codes Are Not Available: A Validation Study Comparing Medication/Prescriber Visit–Based Algorithms with Diagnosis Codes

Wendy Dobson-Belaire, PhD, MBA[1], *Jason Goodfield, BMSc/HBA*[1], *Richard Borrelli, MBA*[1,*], *Fei Fei Liu, BSc(Pharm), MBA*[2], *Zeba M. Khan, RPh, PhD*[3]

[1]*QuintilesIMS, Mississauga, Ontario, Canada;* [2]*Celgene Inc., Mississauga, Ontario, Canada;* [3]*Celgene Corp., Summit, NJ, USA*

ABSTRACT

Background: Using diagnosis code–based algorithms is the primary method of identifying patient cohorts for retrospective studies; nevertheless, many databases lack reliable diagnosis code information. **Objectives:** To develop precise algorithms based on medication claims/prescriber visits (MCs/PVs) to identify psoriasis (PsO) patients and psoriatic patients with arthritic conditions (PsO-AC), a proxy for psoriatic arthritis, in Canadian databases lacking diagnosis codes. **Methods:** Algorithms were developed using medications with narrow indication profiles in combination with prescriber specialty to define PsO and PsO-AC. For a 3-year study period from July 1, 2009, algorithms were validated using the PharMetrics Plus database, which contains both adjudicated medication claims and diagnosis codes. Positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity of the developed algorithms were assessed using diagnosis code as the reference standard. Chosen algorithms were then applied to Canadian drug databases to profile the algorithm-identified PsO and PsO-AC cohorts. **Results:** In the selected database, 183,328 patients were identified for validation. The highest PPVs for PsO (85%) and PsO-AC (65%) occurred when a predictive algorithm of two or more MCs/PVs was compared with the reference standard of one or more diagnosis codes. NPV and specificity were high (99%–100%), whereas sensitivity was low (≤30%). Reducing the number of MCs/PVs or increasing diagnosis claims decreased the algorithms' PPVs. **Conclusions:** We have developed an MC/PV-based algorithm to identify PsO patients with a high degree of accuracy, but accuracy for PsO-AC requires further investigation. Such methods allow researchers to conduct retrospective studies in databases in which diagnosis codes are absent.

Keywords: administrative databases, diagnosis, psoriasis, psoriatic arthritis.

Introduction

Research using administrative claims databases often serves as the foundation for epidemiologic research, postmarketing drug surveillance, and cost-effectiveness evaluations. These databases include physician billing claims, hospitalization records, and pharmacy data that hold insight into the utilization of health care resources that can impact health care policy and patient care. Working with administrative databases offers many advantages including 1) large, comprehensive study samples, 2) insight into "real-world" data, and 3) accessibility due to relatively low cost [1]. Nevertheless, a challenge in working with administrative data is identifying an accurate and reliable patient cohort. In Canada, although some small or regional administrative databases and electronic health records contain diagnosis codes, the largest national pharmacy claims databases lack integration with complete and accurate diagnosis information. This barrier presents a challenge to studying disease-specific drug use patterns and their impact on patient outcomes.

To overcome an absence of diagnosis code, algorithms developed on the basis of medication claims and prescribing physician specialty can be used to infer disease diagnosis. These algorithms may then be applied to drug claims lacking diagnosis code so as to study drug use in specific disease cohorts. Numerous studies have evaluated the strength of diagnosis code algorithms against "criterion standards" such as electronic medical records, chart reviews, or patient surveys [2–4]. Although such studies have shown that the addition of pharmacy data and physician specialty to diagnosis code can strengthen predictive value [5–7], there is a paucity of research into the accuracy of diagnosis algorithms based on prescribing data and specialty alone.

Psoriasis (PsO) is a chronic autoimmune-mediated skin disease typically requiring lifelong care. Estimates on PsO's global prevalence vary from 0.6% to 4.8% on the basis of the location and studied population characteristics [8], with Canadian prevalence

* *Address correspondence to:* Richard Borrelli, QuintilesIMS, 6700 Century Avenue, Suite 300, Mississauga, Ontario L5N 6A4, Canada.
E-mail: rborrelli@ca.imsbrogan.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.06.012

estimated to range from 1% to 3% [9]. Psoriatic arthritis (PsA) can develop in up to 6% to 42% of chronic PsO patients and is associated with systemic inflammation, often resulting in joint pain that can become debilitating [10–12]. Although PsO diagnosis typically precedes PsA, in approximately 15% of cases, arthritis may precede PsO by many years [11,13].

There are several therapeutic drug classes used to treat PsO/PsA; nevertheless, the variability in the disease progression and treatment sequencing requires better investigation. In addition, it is known that there is widespread treatment dissatisfaction, underutilization of systemic treatments, and poor adherence to PsO treatments, especially during prebiologic stages of therapy [14–16]. The ability to accurately identify and observe PsO/PsA patient treatment in the real world is essential to understanding these patients' unmet needs and to better devise strategies to ensure patients are receiving appropriate and cost-effective treatment.

The goal of this project was to develop and validate medication claim/prescriber visit (MC/PV)–based algorithms to identify PsO and PsA patients in health care databases that lack diagnosis codes, such as those found in Canada, to study the drug use behavior in these cohorts. In this study, we describe the development of the algorithms and their subsequent validity testing on the basis of positive predictive value (PPV) using the PharMetrics Plus database, a US claims-based database with diagnosis code and medication claim transactions available. A sample application of the algorithm in Canadian drug databases is also presented and evaluated.

Methods

Data Source

PharMetrics

Validation of the diagnosis algorithm was conducted using the PharMetrics Plus (PharMetrics) database. PharMetrics is an integrated claims database that includes information from fully adjudicated medical and pharmaceutical claims for more than 150 million unique patients across the United States, with 42 million active enrollees with both pharmacy and medical coverage available in 2011. In compliance with the Health Insurance Portability and Accountability Act guidelines, patients are de-identified and represented in the database by unique encrypted patient identification codes. The database includes inpatient and outpatient diagnoses in the form of the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes and procedures in the form of Current Procedural Terminology or Healthcare Common Procedure Coding System for both retail and mail order prescription records. Up to 12 diagnoses can be recorded for a single transaction. Prescription claims also include the National Drug Code, provider specialty, dispensing date, days' supply, quantity dispensed, patients' demographic information, and payment details.

The PharMetrics patient pool is representative of the national, commercially insured US population across various demographic measures including geography, age, and sex [17]. The data are longitudinal, with approximately 22 million patients who have both medical and pharmacy coverage with 4 or more years of continuous enrollment. PharMetrics data have been used for various studies in pharmacoepidemiology, including in PsO [17–20].

Canadian Drug Databases

Private and public prescription claims databases from the Canadian Private Drug Plans, the Ontario Public Drug Plan, and the Régie de l'assurance maladie du Québec, as provided by QuintilesIMS Brogan, Canada, were used to evaluate the application of the proposed algorithms in databases lacking diagnosis code. These databases contain medication claims records with anonymized patient identification, patient age and sex, date of claim, drug identification number, quantity dispensed, and days' supply of prescription. Although the Ontario Public Drug Plan provides census coverage, the Canadian Private Drug Plans and the Régie de l'assurance maladie du Québec data are samples with approximately 70% and approximately 15% of claims coverage, respectively.

Patient Selection

Patients in the PharMetrics database who were actively enrolled in their health, medical, and drug benefit plan from July 1, 2009, to June 30, 2013, were selected. Of those selected, the patient pool was further narrowed on the basis of individuals having *at least* one claim for any PsO-defining molecule as presented in Table 1 or a diagnosis of PsO (*ICD-9-CM* code 696.1) or PsA (*ICD-9-CM* code 696.0). Patients were then profiled on the basis of their demographic characteristics and PsO/PsA medication treatment history at any point in the study period. The patient selection and validation process is shown in Figure 1. Patients included in the Canadian database case study were active in their drug plan for at least 3 years between October 2007 and September 2013 and were selected according to this study's highest precision algorithm of having had two or more claims for a PsO-defining molecule.

Algorithm Development

Algorithms to infer diagnosis of PsO and PsA were developed on the basis of a review of published treatment guidelines for PsO and PsA in Canada and the United States [21–25]. PsO drugs identified through the literature review were subsequently screened to remove molecules with multiple indications in addition to PsO or PsA, such as topical corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs), which would negatively impact algorithm precision. Patients were first classified as PsO on the basis of the presence of a PsO-specific medication in their claims history. These medications are listed in the first column of Table 1. These treatments are generally used for most PsO types, including, but not limited to, scalp, genital, palmopustular, and generalized pustular PsO. Patients identified as

Table 1 – List of molecules included in the PsO and PsO-AC medication claim or prescriber-specialty identification algorithms.

PsO-defining molecules	PsO-AC defining molecules
Acitretin	Abatacept
Anthralin	Anakinra
Calcipotriene	Azathioprine
Calcipotriene + betamethasone	Certolizumab pegol
Calcitriol	Golimumab
Methoxsalen	Hydroxychloroquine
Trioxsalen	Leflunomide
	Rituximab
	Sulfasalazine
	Tocilizumab

PsO, psoriasis; PsO-AC, psoriasis with arthritic conditions.