Vaccine 31S (2013) K62–K73

Contents lists available at ScienceDirect

Vaccine

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

Review

A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data

Kevin G. Moores [a,*], Nila A. Sathe [b]

[a] Division of Drug Information Service, College of Pharmacy, The University of Iowa, USA
[b] Vanderbilt Evidence-based Practice Center, Vanderbilt University Medical Center, Suite 600, 2525 West End Avenue, Nashville, TN 37203-1738, USA

ARTICLE INFO

Article history:
Received 12 December 2012
Received in revised form 30 April 2013
Accepted 25 June 2013

Keywords:
Systemic lupus erythematosus (SLE)
Administrative database
ICD-9
Positive predictive value (PPV)
Incidence calculation
Algorithm validation

ABSTRACT

Purpose: To examine the validity of billing, procedural, or diagnosis code, or pharmacy claim-based algorithms used to identify patients with systemic lupus erythematosus (SLE) in administrative and claims databases.
Methods: We searched the MEDLINE database from 1991 to September 2012 using controlled vocabulary and key terms related to SLE. We also searched the reference lists of included studies. Two investigators independently assessed the full text of studies against pre-determined inclusion criteria. The two reviewers independently extracted data regarding participant and algorithm characteristics and assessed a study's methodologic rigor using a pre-defined approach.
Results: Twelve studies included validation statistics for the identification of SLE in administrative and claims databases. Seven of these studies used the ICD-9 code of 710.0 in selected populations of patients seen by a rheumatologist or patients who had experienced the complication of SLE-associated nephritis, other kidney disease, or pregnancy. The other studies looked at limited data in general populations. The algorithm in the selected populations had a positive predictive value (PPV) in the range of 70–90% and of the limited data in general populations it was in the range of 50–60%.
Conclusions: Few studies use rigorous methods to validate an algorithm for the identification of SLE in general populations. Algorithms including ICD-9 code of 710.0 in physician billing and hospitalization records have a PPV of approximately 60%. A requirement that the code is obtained from a record based on treatment by a rheumatologist increases the PPV of the algorithm but limits the generalizability in the general population.

Contents

Abbreviations: ACR, American College of Rheumatology; CI, confidence interval; CPT, Current Procedural Terminology; ESRD, end stage renal disease; ICD, International Classification of Diseases; KPNC, Kaiser Permanente Northern California; N, number; NR, not reported; SLE, systemic lupus erythematosus; USRDS, US Renal Data System.

* Corresponding author at: Division of Drug Information Service, College of Pharmacy, The University of Iowa, 100 BVC, Iowa City, IA 52242-1324, USA. Tel.: +1 319 335 4800; fax: +1 319 335 4440.
E-mail addresses: kevin-moores@uiowa.edu (K.G. Moores), nila.sathe@vanderbilt.edu (N.A. Sathe)

0264-410X/$ – see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.vaccine.2013.06.104

1. Introduction

Mini-Sentinel, a pilot project sponsored by the United States Food and Drug Administration (FDA), aims to inform and facilitate the development of an active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products [1]. Mini-Sentinel is one facet of the Sentinel Initiative, an FDA effort to develop a national electronic system that will complement existing methods of safety surveillance.

To support this goal, Mini-Sentinel uses administrative and claims data to examine relationships between medical product exposures and health outcomes [1,2]. A first step in developing the Sentinel system is to understand the validity of algorithms (i.e., combinations of billing, procedural, or diagnosis codes, or pharmacy claims) for identifying health outcomes of interest in administrative data. Mini-Sentinel program collaborators selected health outcomes of interest using an expert elicitation process through which investigators developed a list of candidate outcomes based on input from global vaccine safety experts. A panel of 5 vaccine experts then prioritized the list via an iterative process using criteria including clinical severity, public health importance, incidence, and relevance. Two musculoskeletal conditions, systemic lupus erythematosus (SLE) and rheumatoid arthritis, were included on the list of conditions [3].

Understanding algorithms used to identify health outcomes helps to determine the validity of any safety signals observed in these data. Thus, the goal of this project was to identify algorithms used to detect SLE and describe the performance characteristics of these algorithms as reported by the studies in which they were used.

SLE is an autoimmune disease with diverse clinical manifestations in association with autoantibodies to components of the cell nucleus. The expression of tissue injury and clinical manifestations of SLE are believed to be determined by genetic, epigenetic, environmental, hormonal and immunoregulatory factors [4]. It occurs most commonly in young women with a peak incidence between the ages of 15 and 40 years and a female:male ratio of 6–10:1. In the United States, people of African, Hispanic, or Asian ancestry have a higher prevalence of SLE and greater involvement of vital organs compared to other racial or ethnic groups. The estimates of the prevalence of SLE in the United States vary widely with a reported range of as high as 1,500,000 to as low as 161,000 [5,6]. The annual number of deaths with SLE as the underlying cause was reported as 879–1406 from 1979 to 1998, with the highest number reported among black women 45–64 years of age [7]. Patients with SLE have 80–90% survival at 10 years. The presentation of SLE is highly variable and can include various signs and symptoms involving many organ systems including dermatologic, musculoskeletal, renal, nervous, cardiovascular, and pulmonary systems. Considering the clinical heterogeneity of SLE the American College of Rheumatology established 11 criteria to improve the consistency of the diagnosis and to provide some standardization for entry into clinical trials or outcome studies. A definite diagnosis is considered to be made with 4 or more criteria occurring either simultaneously or in succession [8,9]. The American College of Rheumatology Criteria (ACR) for systemic lupus erythematosus are provided in the appendices. Most patients with SLE have general constitutional symptoms including fatigue, malaise, fever, anorexia, and weight loss. The presence of anti-nuclear antibodies is the hallmark of the disease and is present in over 90% of patients.

2. Materials and methods

A detailed description of the methods for the project can be found in the accompanying paper by McPheeters et al. [10]. Briefly, we searched the MEDLINE database via the PubMed interface using the strategies outlined in Appendix A. We developed the search strategy by building on prior Mini-Sentinel approaches to searching [2]. We assessed the need to assess gray literature, including that located via Google Scholar, by testing prior approaches. We also tested EMBASE and other databases to determine the need to search them in addition to MEDLINE. These test searches did not yield any citations beyond the traditional search, thus our final search was conducted in MEDLINE. We limited searches to the last 21 years (1991 to September 2012) and required that included studies address SLE; use an administrative database reporting data from the United States or Canada; and clearly define an algorithm to identify cases of SLE. We also tracked whether studies reported validation of the algorithm (e.g., via chart review or independent diagnosis). We also searched the reference lists of included studies. Two investigators independently assessed the full text of each study against our inclusion criteria with disagreements resolved via a third investigator or discussion to reach consensus.

One investigator extracted data regarding the study population, outcome studied, algorithms used, validation procedure, and validity statistics. A second investigator independently verified the accuracy of the data extracted. The first author independently extracted data regarding methodologic elements of included studies such as the population sampled and sampling methods, methods for locating cases, and methods for validating the accuracy of diagnoses in cases located in order to the inform the writing of the report. We summarized results of studies qualitatively and report key characteristics below.

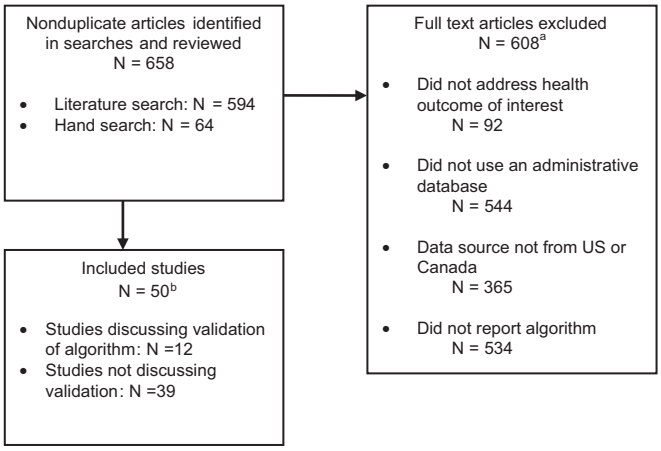

Fig. 1. Disposition of studies located for review.
[a]Numbers do not tally as studies could be excluded for multiple reasons.
[b]9/51 included studies identified via hand search.

3. Results

Our searches identified 658 citations of which 50 met our inclusion criteria (Fig. 1). Among the 50 studies meeting our criteria, 12 described validation of the algorithm used to identify SLE cases. We focus on those studies in this review, and Table 1 summarizes these study characteristics. Characteristics of studies not describing validation of the algorithm are summarized in Table 2. The appendices include our search strategies and a list of studies not meeting our review criteria.

Five studies examined data from general populations and seven studies reported on more selected populations of patients seen by a rheumatologist or patients who had experienced the complication