VALUE IN HEALTH ▮ (2016) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Development and Validation of Algorithms to Identify Statin Intolerance in a US Administrative Database

Kathy L. Schulman, MA[1,*], *Lois E. Lamerato, PhD*[2], *Mehul R. Dalal, PhD*[3], *Jennifer Sung, PharmD, MS*[3], *Mehul Jhaveri, PharmD*[3], *Andrew Koren, MD*[3], *Usha G. Mallya, PhD*[3], *JoAnne M. Foody, MD*[4]

[1]*Outcomes Research Solutions, Inc., Waltham, MA, USA;* [2]*Henry Ford Health System, Detroit, MI, USA;* [3]*Sanofi US, Bridgewater, NJ, USA;* [4]*Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*

ABSTRACT

Objectives: To develop and validate algorithms to define statin intolerance (SI) in an administrative database using electronic medical records (EMRs) as the reference comparison. **Methods:** One thousand adults with one or more qualifying changes in statin therapy and one or more previous diagnoses of hyperlipidemia, hypercholesterolemia, or mixed dyslipidemia were identified from the Henry Ford Health System administrative database. Data regarding statin utilization, comorbidities, and adverse effects were extracted from the administrative database and corresponding EMR. Patients were stratified by cardiovascular (CV) risk. SI was classified as absolute intolerance or titration intolerance on the basis of changes in statin utilization and/or the occurrence of adverse effects and laboratory testing for creatine kinase. Measures of concordance (Cohen's kappa [κ]) and accuracy (sensitivity, specificity, positive predictive value [PPV], and negative predictive value) were calculated for the administrative database algorithms. **Results:** Half of the sample population was white, 52.9% were women, mean age was 60.6 years, and 35.7% were at high CV risk. SI was identified in 11.5% and 14.0%, absolute intolerance in 2.2% and 3.1%, and titration intolerance in 9.7% and 11.8% of the patients in the EMR and the administrative database, respectively. The algorithm identifying any SI had substantial concordance ($\kappa = 0.66$) and good sensitivity (78.1%), but modest PPV (64.0%). The titration intolerance algorithm performed better ($\kappa = 0.74$; sensitivity 85.4%; PPV 70.1%) than the absolute intolerance algorithm ($\kappa = 0.40$; sensitivity 50%; PPV 35.5%) and performed best in the high CV-risk group (n = 353), with robust concordance ($\kappa = 0.73$) and good sensitivity (80.9%) and PPV (75.3%). **Conclusions:** Conservative but comprehensive algorithms are available to identify SI in administrative databases for application in real-world research. These are the first validated algorithms for use in administrative databases available to decision makers.

Keywords: administrative data, cardiovascular risk, claims data, electronic medical record, hypercholesterolemia, statin intolerance, validation.

Introduction

Research has consistently demonstrated that statins decrease both the risk for cardiovascular (CV) events and mortality rates in patients with hypercholesterolemia [1,2]. Consequently, contemporary lipid management guidelines recommend statin therapy for patients with increased CV risk who are most likely to benefit in terms of atherosclerotic cardiovascular disease (ASCVD) risk reduction. The American College of Cardiology/American Heart Association guidelines identify four major statin benefit groups: 1) individuals with clinical ASCVD, 2) individuals with elevated low-density lipoprotein cholesterol (LDL-C; ≥190 mg/dl), 3) individuals with diabetes and increased LDL-C (70–189 mg/dl), and 4) individuals with estimated 10-year ASCVD risk of 7.5% or higher [3]. The National Lipid Association (NLA) recommends moderate- to high-intensity statin therapy for patients with ASCVD or diabetes mellitus, regardless of baseline lipid levels [4].

Despite the known benefits of statins, many patients, including those at high CV risk, discontinue treatment [5]. Experiencing statin-related adverse effects (AEs) is one of the most common reasons for statin switching or discontinuation [6]. Even among adherent patients, providers may not always be able to prescribe the preferred therapeutic dose as AE frequency increases with dose intensity [7,8]. The most common statin-associated AEs are muscle-related, and these have been documented in 16.0% to 32.9% of patients receiving statins and in 15.4% to 33.2% of

Conflicts of interest: At the time this work was completed, K. L. Schulman was a consultant for Sanofi US. L. E. Lamerato has received research support from Sanofi US. At the time this work was completed, M. R. Dalal was employed by and had ownership interest in Sanofi US. J. Sung, A. Koren, and U. G. Mallya are employed by and have ownership interest in Sanofi US. At the time this work was completed, M. Jhaveri was employed by Sanofi US. At the time this work was completed, J. M. Foody was a consultant for Merck, Pfizer, Bristol-Myers Squibb, Sanofi, and AstraZeneca.

* *Address correspondence to:* Kathy L. Schulman, Outcomes Research Solutions, Inc., 303 Wyman Street, Suite 300, Waltham, MA 02451. E-mail: kathy.schulman@orsolutionsinc.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2016.03.1858

placebo-treated patients in long-term randomized clinical trials [9]. A number of other AEs, however, have been associated with statin therapy, including elevated transaminase, headache, insomnia, fatigue, dyspepsia, nausea, rash, alopecia, constipation, diarrhea, gastrointestinal disturbance, arthritis, and renal disorders [10–12]. Studies have estimated the incidence or prevalence of muscle-related symptoms (5%–25%), but these data are from clinical trials, single-site retrospective cohort studies, and patient surveys and as such may have limited generalizability to the larger population [6,13–18]. Factors associated with increased risk of statin intolerance (SI) include advanced age, clinical or subclinical hypothyroidism, and pre-existing liver or chronic kidney disease [19].

Although SI is recognized as a clinical entity, there is no consensus yet on a single definition [19,20], making it difficult to assess the incidence of SI. The Canadian Working Group (CWG) has defined SI as a clinical syndrome

> characterized by inability to use statins for long-term reduction of lipids and/or CV risk because of significant symptoms and/or biomarker abnormalities that can be temporally attributed to the initiation or dose escalation of statins; if appropriate, drug withdrawal and rechallenge can strengthen the association. [21(p1553)]

The NLA defines SI as

> a clinical syndrome characterized by the inability to tolerate at least 2 statins: one statin at the lowest starting daily dose AND another statin at any daily dose, due to either objectionable symptoms (real or perceived) or abnormal lab determinations, which are temporally related to statin treatment and reversible upon statin discontinuation, but reproducible by re-challenge, with other known determinants being excluded. [22(pS78)]

These definitions, although similar in many respects, differ in terms of mandating statin dosage and rechallenge and also with respect to the validity of patient symptoms.

Accurate identification of SI is important in terms of both establishing its incidence and characterizing the associated clinical, economic, and quality-of-life burdens, which are at present unknown. The objective of this study was to develop and validate algorithms to identify patients with SI in an administrative database on the basis of the overlap between CWG and NLA definitions.

Methods

Data Source

The data source selected to develop and validate the SI algorithms was from the Henry Ford Health System (HFHS). The HFHS offers primary, acute, and specialty care services in the Midwest and also includes a wholly owned nonprofit health maintenance organization, the Health Alliance Plan (HAP). HFHS data repositories include both an administrative database, which provides comprehensive medical billing and pharmacy claims data, and electronic medical records (EMRs), including laboratory results, from all sites of service, linkable for each patient using a lifetime patient identifier. This study was approved by the HFHS institutional review board.

Validation Study Sample

The validation study sample was drawn from the HFHS administrative claims database using the following criteria: adults (≥18 years) who 1) had one or more statin qualifying events between December 1, 2005, and November 30, 2010, 2) were continuously enrolled in HAP for 1 year before and 2 years after the qualifying event, and 3) had one or more diagnoses (*International Classification of Diseases, Ninth Revision, Clinical Modification* [*ICD-9-CM*]) of hyperlipidemia (272.4), hypercholesterolemia (272.0), or mixed hyperlipidemia (272.2) before the qualifying event (index) because lipid levels may have been influenced by existing statin therapy.

Qualifying events included statin discontinuation, switch, or a decrease in statin dosage. Statin switch events were eligible only if the patient switched from a high-intensity statin to a lower intensity statin or from a moderate/low-intensity statin to another moderate/low-intensity statin. Definitions of moderate/low-intensity and high-intensity statins were based on LDL-C–lowering capability. A statin was considered of high intensity or high potency if it reduced the LDL-C level from baseline by more than 45% (Table 1) [23,24].

If a patient had multiple qualifying events, the first qualifying event that met the eligibility criteria was selected. If the statin medication pattern for an individual patient met the criteria for both moderate/low-intensity and high-intensity statins, only the high-intensity qualifier was selected, although this did not exclude evaluation of the low-intensity statin if it also occurred during the study observation window. Study index was the date of each qualifying event. All patients were observed for 1 year pre- and 2 years postindex.

A sample of 1000 patients was drawn from the pool of eligible patients. Patients were categorized as being at high CV risk if they had two or more *ICD-9-CM* diagnoses of diabetes, coronary heart disease, or peripheral artery disease in the 12 months before the qualifying event on outpatient claims on different days or a single diagnosis on an inpatient claim. Data on patients' demographic and clinical characteristics, statin utilization, and AEs were extracted electronically from the administrative database. Statin utilization and AEs data were simultaneously abstracted from the EMR by three trained and credentialed research associates, after completion of a pilot study. Patients with unresolvable conflicts in the EMR or data quality issues in pharmacy claims were excluded from the study.

Statin Utilization and AEs

Statin exposure windows were created to characterize statin utilization and to identify associated AEs. Exposure windows were created independently on the basis of the statin regimen (statin plus dose) prescribed in the EMR as well as on statin fill records from pharmacy data in the administrative database. If present during an exposure window, SI was identified and confirmed.

The following information was abstracted from the EMR: statin prescription dates and dosage, AE type and date, whether changes in prescription were linked to an AE, and the primary and secondary reasons for statin discontinuation or dose lowering, if applicable. A prescription was end-dated if there was an absence of documentation for more than 12 months. Prescription records were then collapsed if the statin name and dosage were unchanged.

Statin exposure windows based on the administrative data relied on pharmacy claims. Periods of continuous use for each statin regimen were identified, allowing for a 30-day gap in the daily drug coverage pattern. In addition, a 45-day grace period was appended to the end of the exposure window for each statin regimen to account for verbal changes in physician instruction that were not reflected in the pharmacy claims records.

A list of AEs (primary and secondary) related to statin use was compiled (Table 2) [10–12]. The occurrence of an AE was based on the documentation of either the clinical term or the *ICD-9-CM* diagnosis code for an eligible AE in the EMR and independently based solely on the documentation of an *ICD-9-CM* diagnosis code for one of the eligible AEs in administrative data.

Table 1 – Statin intensity and daily dosage* [23,24].

Statin intensity	Expected reduction in LDL-C level (%)	Daily statin dosage (mg)						
		Atorvastatin	Fluvastatin	Lovastatin	Pravastatin	Rosuvastatin	Simvastatin	Pitavastatin
Moderate to low	≤30	–	20 or 40	10 or 20	10	–	10	1
	31–40	10	80	40	20	–	20	2
	41–45	20	–	80	40 or 80	5	40	4
High	46–54	40	–	–	–	10	80	–
	55–62	80	–	–	–	20	–	–
	≥63	–	–	–	–	40	–	–

LDL-C, low-density lipoprotein cholesterol.
* Statins were identified using national drug codes.

SI Definition

The reference definition of SI was based on EMR data. The SI algorithms were based only on administrative data. Statin-intolerant patients identified by the administrative database algorithms were then compared with those identified by the EMR reference, and performance of the algorithms was evaluated. The EMR was presumed to be a superior source of information about SI, given the richness of its clinical repository (e.g., text-based office notes with symptom information). In the primary analysis, the complete list of potential AEs was included in the definition of SI, whereas the secondary analysis included only musculoskeletal events (Table 2).

All patients, in both data sources, were classified as absolute statin-intolerant (unable to tolerate any exposure to statins) or titration statin-intolerant (maximum tolerated dosage less than the desired therapeutic dosage). Patients in the study could meet the criteria for both absolute and titration intolerance.

The occurrence of absolute intolerance was predicated on the notion that patients must be rechallenged with a second statin unless the documented AE was rhabdomyolysis, a potentially life-threatening syndrome resulting from the breakdown of skeletal muscle fibers and subsequent release of cellular content into the circulation [25]. Accordingly, in the EMR, absolute intolerance required the following three criteria to be satisfied: 1) the physician discontinues two or more prescriptions for different statin agents (only one in the presence of rhabdomyolysis), 2) the patient never receives another statin prescription (minimum of 100 days of observation remaining postdiscontinuation), and 3) the reason for discontinuation (primary or secondary) is an AE, or an explicit link is documented between the statin and an AE, or the patient is documented as statin-intolerant/allergic.

The absolute intolerance algorithm applied to the administrative database also reflected these principles, requiring that 1) the patient stops filling prescriptions for two or more statin agents (only one in the presence of rhabdomyolysis), 2) the patient never fills another statin prescription (minimum of 100 days of observation remaining), and 3) either an elevated creatine kinase test result or an AE is documented during exposure to the last documented statin.

The occurrence of titration SI was predicated on the supposition that, among patients on high-intensity statin therapy, any reduction in intensity reflects the inability of the patient to tolerate the desired therapeutic statin strength. Accordingly, in the EMR, titration intolerance required the following three criteria to be satisfied: 1) the physician discontinues a prescription for a high-intensity statin and initiates a prescription at a lower level of intensity, 2) either the discontinued high-intensity statin is linked to SI or there is an AE documented during the period of high-intensity exposure, and 3) there is no subsequent prescription for a statin at the original intensity level (minimum of 100 days of observation remaining).

The titration intolerance algorithm that was applied to the administrative database also embodied these principles, requiring that 1) the patient discontinues filling a prescription for a high-intensity statin and fills a prescription at a lower level of intensity, 2) no subsequent prescription is filled for a statin at the original intensity level, and 3) there is a minimum of 100 days remaining in the observation window after the lowering of the statin intensity.

Data Analysis

The EMR served as the reference against which the algorithms were evaluated, as is customary in administrative data–based validation studies. Concordance between the reference and the algorithms was measured using Cohen's kappa (κ), a numerical rating of the degree to which the observers (or data sources) agree by chance. Kappa is standardized on a scale of −1 to 1, where

Table 2 – Definition of statin-related AEs.

AE	Clinical term	ICD-9-CM code
Musculoskeletal AEs/side effects		
Myalgia and/or myositis	Muscle pain, spasm, weakness, discomfort, soreness, cramps, or aching	728.85 (spasm of muscle); 729.82 (cramp in limb); 728.87 (muscle weakness-generalized); 729.1x (myalgia and myositis, unspecified)
Arthralgia	Joint pain, joint stiffness	719.4x (pain in joint); 719.5x (stiffness in joint)
Rhabdomyolysis	Rhabdomyolysis	728.88 (rhabdomyolysis)
Myopathy	Myopathy, toxic myopathy	359.4 (toxic myopathy); 359.89 (other myopathy); 359.9 (myopathy, unspecified)
Other muscle toxicity	Limb pain, ligament pain, fascia pain, limb discomfort, other drug poisoning	728.9x (unspecified disorder of muscle, ligament, and fascia); 729.1x (pain in limb); E980.4 (injury from other specified drugs)
Elevated CPK	CPK >5× ULN	
Gastrointestinal-related AEs/side effects		
Nausea	Nausea, vomiting	787.0 (nausea and/or vomiting)
Constipation	Constipation	564.0 (constipation)
Diarrhea	Diarrhea	564.5 (functional diarrhea); 787.91 (diarrhea)
Other gastrointestinal distress	Flatulence, gas, bloating, abdominal pain, gastritis, duodenitis	787.3 (flatulence, eructation, and gas pain); 789.0 (abdominal pain); 535 (gastritis and duodenitis)
Other AEs/side effects		
Anaphylaxis	Anaphylactic shock	995.0 (other anaphylactic shock)
Rash/facial flushing	Urticaria, angioedema, rash, hives, dermatitis, edema, erythema, rosacea, facial flushing, pruritus	782.1 (rash and other nonspecific skin eruption); 708.0 (allergic urticaria); 693.0 (dermatitis due to drugs/medicines); 995.1 (angioneurotic edema); 995.2 (other and unspecified AE of drug, medicinal, and biological substance (due) to correct medicinal substance properly administered); 695.1 (erythema multiforme); 698 (pruritus and related conditions); 695.3 (rosacea)
Cognitive impairment	Memory loss, forgetfulness, confusion	331.83 (mild cognitive impairment, so stated); 780.93 (memory loss); 799.51 (attention or concentration deficit)
Elevated LFT	LFT >3× ULN	

Note. Intolerance definition: The primary analysis included the complete list of potential AEs in the definition of SI, whereas the secondary analysis included only musculoskeletal events.
AEs, adverse effects; CPK, creatine phosphokinase; *ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification*; LFT, liver function test; ULN, upper limit of normal.

1 represents perfect agreement, 0 represents what would be expected by chance, and negative values indicate agreement less than chance [26]. A kappa statistic between 0.41 and 0.60 was used to indicate moderate agreement, 0.61 to 0.80 substantial agreement, and 0.81 to 1.0 almost perfect agreement [27]. In addition, we used traditional tests of diagnostic accuracy to measure the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the algorithms. The algorithms were refined to incorporate new information and expert opinions after the entire medical record was reviewed for patients identified by the algorithms as false-positive or false-negative. The definitions of SI used in the administrative database were subsequently revised to maximize the agreement with the NLA definition of SI [22] and to maximize the concordance and accuracy of the algorithms. Only the final versions of the algorithms are presented.

Results

Patient Disposition and Characteristics

A total of 46,783 patients enrolled in the HAP were prescribed a statin between December 1, 2004, and November 30, 2012. Figure 1 details accrual into the pool of eligible patients. A total of 1000 patients were selected to be included in the final validation sample on the basis of the observed distribution (80/20) of moderate/low-intensity and high-intensity statins. To adjust for HFHS-specific formulary restrictions, a higher percentage of atorvastatin patients (40%) and a lower percentage of simvastatin patients (40%) than observed in the pool of eligible patients were selected to reflect real-world statin use. The remaining 20% of selected patients represent the use of other statins. Because of irreconcilable differences in the medical records, 10 patients (1%) were excluded from the study after data abstraction.

Table 3 presents the baseline demographic and clinical characteristics of the sample. Half of the sample population was white (49.8%) and half was female (52.9%). Twenty-five percent of the patients resided in an area where the median household income was less than $40,000, whereas 20.6% resided in an area where the median household income was $70,000 or more. Prescriptions for nonstatin lipid-lowering therapy were used infrequently. A total of 353 patients (36%) had baseline evidence of diabetes (27.4%), coronary artery disease (11.7%), or peripheral artery disease (0.6%) and were classified as having high CV risk. Hypertension was the most common comorbidity, identified in 39.9% of the patients at

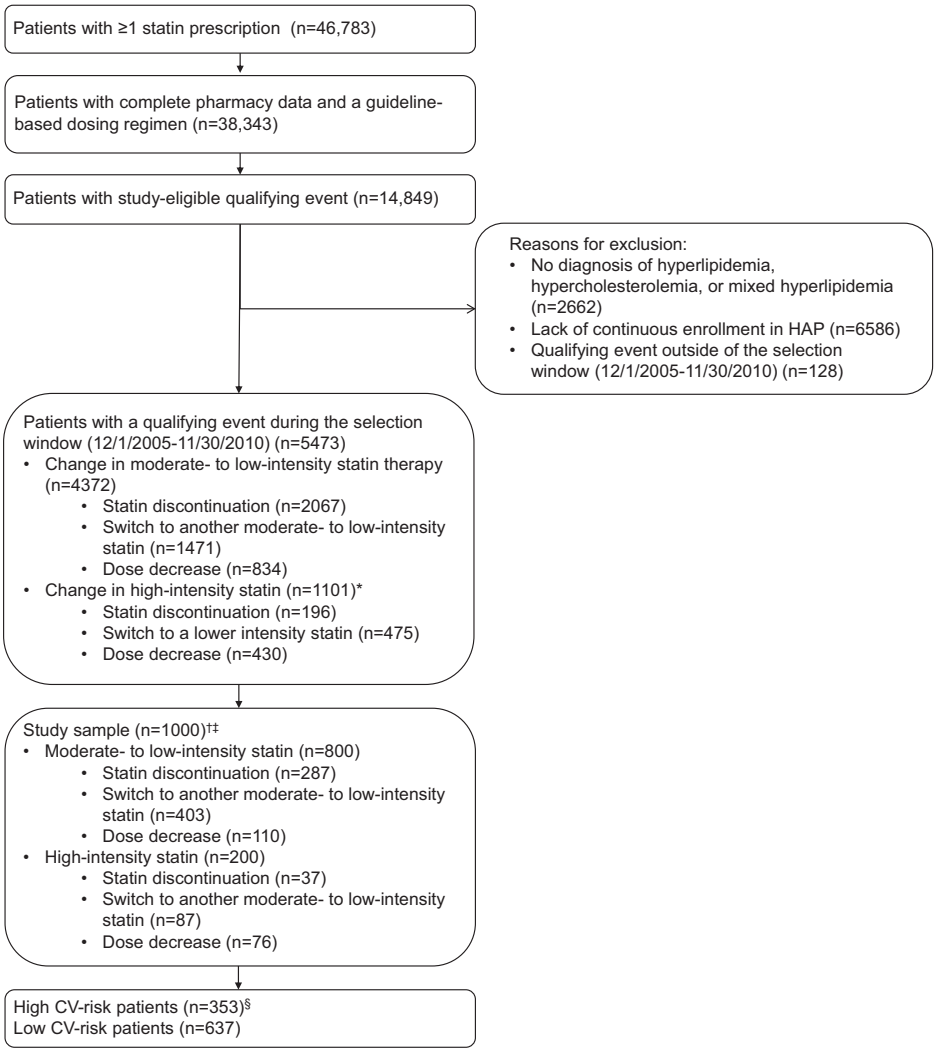

Fig. 1 – Patient disposition. CV, cardiovascular risk; HAP, Health Alliance Plan. *High-intensity statin: atorvastatin 40 or 80 mg; rosuvastatin 10, 20, or 40 mg; and simvastatin 80 mg. †Patients were selected randomly to preserve the 80/20 split between moderate- to low-intensity and high-intensity statin patients observed in the administrative database, with 40% on simvastatin, 40% on atorvastatin, and 20% on other statins. ‡Ten patients were excluded because of data quality issues. §Patients with a diagnosis of diabetes, coronary heart disease, or peripheral artery disease in the 12 mo before the qualifying event.

baseline. In the EMR, the most frequent AEs documented during an episode of statin exposure that was classified as intolerant were other muscle toxicity (19.8%), myalgia or myositis (18.4%), arthralgia (12.8%), and rash (11.3%) (see Table in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2016.03.1858).

SI and Algorithm Concordance

In the primary analysis, 11.5% of the patients (n = 114) were identified in the EMR as statin-intolerant compared with 14.0% of the patients (n = 139) in the administrative database (Table 4). The proportion of patients identified as statin-intolerant was twice as high among patients at high CV risk as it was among those at low CV risk. Titration intolerance occurred in 16.7% and 5.8% of the patients in the EMR and in 18.7% and 8.0% of the patients in the administrative database in high and low CV-risk groups, respectively. Few patients were identified as absolute intolerant in either high or low CV-risk groups. The frequency of SI in the secondary analysis was lower but not notably so (Table 5). It should be noted that these frequencies cannot be used as proxies for SI incidence in the general population because the study was not conducted on a random sample of statin users.

Although rates of SI in the EMR and the administrative database were similar, the algorithms did not always identify the same patients as intolerant. Generally, the algorithm identifying any SI demonstrated high levels of overall agreement (92.4%), substantial concordance ($\kappa = 0.66$), and good sensitivity (78.1%), but modest PPV (64.0%), reflecting differences in the accuracy of the algorithm associated with detecting absolute and titration intolerance. The absolute intolerance algorithm had notably lower concordance ($\kappa =$ 0.30–0.56), sensitivity (46%–55%), and PPV (24%–60%) both overall and in each of the CV-risk groups. Both absolute and titration intolerance algorithms performed best in the high CV-risk group. In the high CV-risk group, the titration intolerance algorithm was most robust, with overall agreement at 93.5%, substantial concordance ($\kappa = 0.78$), and good PPV (77.3%). The titration intolerance algorithm in the low CV-risk cohort also showed substantial concordance ($\kappa = 0.68$) and good sensitivity (83.8%), but modest PPV (60.8%). NPV was uniformly high, ranging from 95% to 99%. Although the frequency of SI in the secondary analysis was similar to that in the primary analysis,

Table 3 – Patients' demographic and clinical characteristics at baseline.

Characteristic	High CV-risk patients (n = 353)	Low CV-risk patients (n = 637)	All patients (N = 990)
Sex: male, n (%)	183 (51.8)	283 (44.4)	466 (47.1)
Age (y), mean ± SD at index	62.6 ± 12.4	59.5 ± 12.2	60.6 ± 12.4
Race, n (%)			
White	154 (43.6)	339 (53.2)	493 (49.8)
Black	136 (38.5)	171 (26.8)	307 (31.0)
Asian	9 (2.5)	16 (2.5)	25 (2.5)
Others	54 (15.3)	111 (17.4)	165 (16.7)
Medical history, n (%)			
Abdominal aortic aneurysm	1 (0.3)	2 (0.3)	3 (0.3)
Carotid artery disease	3 (0.8)	2 (0.3)	5 (0.5)
Chronic kidney disease	21 (5.9)	10 (1.6)	31 (3.1)
CHD/CAD	116 (32.9)	–	116 (11.7)
Diabetes	271 (76.8)	–	271 (27.4)
Hypertension	195 (55.2)	200 (31.4)	395 (39.9)
PAD	6 (1.7)	–	6 (0.6)
Qualifying statin medication, n (%)			
Atorvastatin	122 (34.6)	275 (43.2)	397 (40.1)
Fluvastatin	1 (0.3)	3 (0.5)	4 (0.4)
Lovastatin	39 (11.0)	58 (9.1)	97 (9.8)
Pitavastatin	0 (0.0)	0 (0.0)	0 (0.0)
Pravastatin	19 (5.4)	34 (5.3)	53 (5.4)
Rosuvastatin	23 (6.5)	19 (3.0)	42 (4.2)
Simvastatin	149 (42.2)	248 (38.9)	397 (40.1)
Qualifying statin event, n (%)			
Dose decrease in existing therapy	89 (25.2)	95 (14.9)	184 (18.6)
High-intensity* to moderate- to low-intensity switch	42 (11.9)	45 (7.1)	87 (8.8)
Moderate- to low-intensity to another moderate- to low-intensity switch	124 (35.1)	274 (43.0)	398 (40.2)
Stopped statin therapy altogether	98 (27.8)	223 (35.0)	321 (32.4)
Other prescribed LLT†, n (%)	62 (17.6)	47 (7.4)	109 (11.0)
Lab results (mg/dl), mean ± SD			
LDL-C	102.1 ± 35.8	123.5 ± 39.6	115.5 ± 39.6
HDL-C	52.2 ± 15.7	57.2 ± 15.9	55.3 ± 16.0
Triglycerides	156.4 ± 85.8	156.0 ± 85.5	156.2 ± 85.5
Total cholesterol	184.9 ± 43.8	211.6 ± 45.1	201.7 ± 46.4
AST	29.0 ± 50.4	25.1 ± 11.4	26.5 ± 32.3
ALT	33.9 ± 40.7	31.3 ± 16.9	32.3 ± 28.4
Creatine kinase	483.3 ± 2549.7	260.5 ± 539.2	360.1 ± 1748.3

ALT, alanine aminotransferase; AST, aspartate aminotransferase; CAD, coronary artery disease; CHD, coronary heart disease; CV, cardiovascular; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; LLT, lipid-lowering therapy; PAD, peripheral artery disease.

* High-intensity statin: atorvastatin 40 or 80 mg; rosuvastatin 10, 20, or 40 mg; and simvastatin 80 mg.

† Other prescribed LLTs included cholestyramine, clofibrate, colesevelam, colestipol, ezetimibe, fenofibrate, fenofibric acid, gemfibrozil, niacin, omega-3-acid ethyl esters, and probucol.

performance was diminished for both absolute intolerance and titration intolerance algorithms.

Discussion

This study is the first to develop and validate algorithms for the identification of SI in an administrative database. Definitions of SI used in these algorithms align with those endorsed by the NLA Statin Intolerance Panel [22] and the CWG [21] and are an essential ingredient in characterizing the clinical, economic, and quality-of-life burdens associated with SI. The algorithms successfully identified both absolute and titration intolerance in a population of patients at high CV risk. The algorithm for any SI demonstrated substantial concordance (κ = 0.73), sensitivity (80.9%), PPV (75.3%), and NPV (95.4%), despite marginal performance in the absolute intolerance portion of the algorithm. In this study, the titration intolerance algorithm performed best in patients with high CV risk, as evidenced by its substantial concordance (κ = 0.78), high NPV (97.2%) and sensitivity (86.4%), and good PPV (77.3%). There was also a higher incidence of SI in patients with high CV risk than in patients with low CV risk.

On the basis of the results, the algorithms performed comparably to other CV algorithms previously developed for administrative data. Thacker et al. [28] developed an algorithm to identify patients at high risk for coronary heart disease events (PPV 87%; sensitivity 69%; specificity 90%). Hsieh et al. [29] validated a claims-based algorithm to measure the incidence of acute ischemic stroke (PPV 88%; sensitivity 97%; specificity not reported). Numerous algorithms have been developed to identify heart failure (PPV 84%–100%; sensitivity 33%–83%) [30], whereas one study assessing CV outcomes has reported kappa statistics ranging from 0.71 to 0.91 [31].

Both the reference definition and SI algorithms developed in the course of this study reflect SI definitions of the CWG [21] and

Table 4 – Primary analysis of SI (using the complete set of AEs) by data source and risk status.

Patient population	Patients deemed statin–intolerant, n (%), in		Cohen's κ	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Overall agreement (%)
	EMR	Administrative database						
All eligible patients (N = 990)								
Any absolute intolerance	22 (2.2)	31 (3.1)	0.40	50.0	97.9	35.5	98.9	96.9
Any titration intolerance	96 (9.7)	117 (11.8)	0.74	85.4	96.1	70.1	98.4	95.1
Any SI	114 (11.5)	139 (14.0)	0.66	78.1	94.3	64.0	97.1	92.4
Patients at high CV risk (n = 353)								
Any absolute intolerance	11 (3.1)	10 (2.8)	0.56	54.5	98.8	60.0	98.5	97.5
Any titration intolerance	59 (16.7)	66 (18.7)	0.78	86.4	94.9	77.3	97.2	93.5
Any SI	68 (19.3)	73 (20.7)	0.73	80.9	93.7	75.3	95.4	91.2
Patients at low CV risk (n = 637)								
Any absolute intolerance	11 (1.7)	21 (3.3)	0.30	45.5	97.4	23.8	99.0	96.5
Any titration intolerance	37 (5.8)	51 (8.0)	0.68	83.8	96.7	60.8	99.0	95.9
Any SI	46 (7.2)	66 (10.4)	0.57	73.9	94.6	51.5	97.9	93.1

AEs, adverse effects; CV, cardiovascular; EMR, electronic medical record; NPV, negative predictive value; PPV, positive predictive value; SI, statin intolerance.

Table 5 – Secondary analysis of SI (using only musculoskeletal AEs) by data source and risk status.

Patient population	Patients deemed statin–intolerant, n (%), in		Cohen's κ	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Overall agreement (%)
	EMR	Administrative database						
All eligible patients (N = 990)								
Any absolute intolerance	19 (1.9)	29 (2.9)	0.36	47.4	97.9	31.0	99.0	97.0
Any titration intolerance	80 (8.1)	117 (11.8)	0.70	90.0	95.1	61.5	99.1	94.6
Any SI	96 (9.7)	138 (13.9)	0.62	81.3	93.3	56.5	97.9	92.1
Patients at high CV risk (n = 353)								
Any absolute intolerance	10 (2.8)	9 (2.5)	0.51	50.0	98.8	55.6	98.5	97.5
Any titration intolerance	50 (14.2)	66 (18.7)	0.73	90.0	93.1	68.2	98.3	92.6
Any SI	58 (16.4)	72 (20.4)	0.68	82.8	91.9	66.7	96.4	90.4
Patients at low CV risk (n = 637)								
Any absolute intolerance	9 (1.4)	20 (3.1)	0.26	44.4	97.5	20.0	99.2	96.7
Any titration intolerance	30 (4.7)	51 (8.0)	0.65	90.0	96.0	52.9	99.5	95.8
Any SI	38 (6.0)	66 (10.4)	0.54	78.9	94.0	45.5	98.6	93.1

AEs, adverse effects; CV, cardiovascular; EMR, electronic medical record; NPV, negative predictive value; PPV, positive predictive value; SI, statin intolerance.

the NLA Statin Intolerance Panel [22] by requiring statin rechallenge when appropriate clinically [19,21,22] and some evidence of an AE during the period of statin exposure, regardless of whether the presenting symptom was "real or perceived" [22]. In addition, the reference EMR definition required that the documented AE be temporal in nature and attributed to the initiation or dose escalation of statins. Nevertheless, one element of the NLA definition was not included in either the reference definition or in the algorithms, that is, the requirement that one of the two statins administered be at the lowest starting daily dose. Moreover, the SI algorithms did not require exclusion of other known determinants of an AE, such as hypothyroidism, concurrent illness, drug interaction, or muscle disease underlying a change in physical activity [22]. In addition, these algorithms did not require symptom reversal upon statin discontinuation or dose lowering [21,22].

Although the study definitions incorporated the best consensus definition available when the algorithms were developed, the definitions of SI and the guidelines for the management of hypercholesterolemia continue to evolve. Recently, the International Lipid Expert Panel proposed a unified definition of SI, which includes statin rechallenge and a link with confirmed statin-related, muscle-based AEs [19]. Likewise, the treatment guidelines for hypercholesterolemia changed during the course of this study. For example, the 2013 American College of Cardiology/American Heart Association guidelines consider rosuvastatin 10 mg a moderate-intensity statin [3], whereas in this study rosuvastatin is classified as a high-intensity statin. Nevertheless, rosuvastatin 10 mg has demonstrated high efficacy in lowering LDL-C (up to 52% reduction) [23,32].

The algorithms described in this report performed best in patients at high CV risk. These patients are most likely to benefit from lipid-lowering therapy in terms of reduced CV event risk and as such are at the greatest risk for poor outcomes as a result of inadequate LDL-C reduction due to SI. In addition, these patients require close monitoring, especially early in their therapy, because statin-associated muscle AEs occur most frequently within the first 6 months of statin therapy initiation [33]. The diagnosis of SI can, however, be challenging because patients may be seen by clinicians infrequently, may not be clear in describing their symptoms (e.g., joint pain instead of muscular pain), may minimize symptom severity, may present atypically or with symptoms associated with other conditions and/or medications, or may be noncompliant with statins prescribed as part of the rechallenge. In addition, providers may be more focused on treating other urgent health concerns. All these scenarios may contribute to a delay in identifying and remediating SI.

The ability to measure SI in large, real-world data sets can expedite identification of the predictors of SI and enable earlier, more effective intervention. At the practice level, this could result in a targeted intervention for patients at high risk for SI, improving patient outcomes and facilitating a more efficient allocation of valuable primary care resources. At the payer level, an enhanced understanding of the role of SI in patient adherence could inform both care and medication management programs. For researchers, the present algorithm serves as the basis for the development of more sophisticated probabilistic models.

Although standards for what constitutes adequate sensitivity, specificity, PPV, and NPV have not yet been established, the sensitivity associated with the absolute intolerance algorithm is likely too low (50%) to support studies of incidence or prevalence. In contrast, the NPV is sufficiently high (99%) to cull a sample of statin users for whom there is no evidence of absolute intolerance. In the absence of standards establishing minimum thresholds for what constitutes a validated algorithm, researchers must determine whether an algorithm reporting a sensitivity of 86% is adequate to detect disease incidence or whether a PPV of 77% is adequate to cull a cohort of patients with titration intolerance. Until such standards are established, the authors encourage researchers to consider using sensitivity analyses as appropriate.

Researchers should consider a number of other factors before implementing this algorithm. Algorithm performance may not be generalizable across data sources if coding systems differ across populations. The present study was predicated, in part, on the definition of CV risk. Algorithm performance may vary if alternate definitions of high CV risk are used.

The rate of SI observed in the present study cannot be generalized to all statin users because the study was based on a population of patients who had experienced at least one change in statin therapy. Nevertheless, extrapolation to all eligible HFHS patients suggests an overall SI rate between 4% and 6%, an estimate that is at the lower end of the real-world range reported for muscle-related AEs (5%–25%) [6,13–18]. Given the existing literature and the likelihood that symptoms and diagnoses may be underreported if such documentation on a claim will not result in an increase in reimbursement, we believe that the algorithms may produce a conservative estimate of SI.

Zhang et al. [17] used EMR to investigate reasons for statin discontinuation and found that more than half of the study patients discontinued at least one statin but that, among those who were rechallenged, more than 90% remained actively on statin therapy 12 months after a statin-related event. The study, however, focused only on statin discontinuation and analyzed only the first-reported statin-related event. Similarly, Robison et al. [34] used the Intermountain Healthcare EMR database to identify patients with a documented history of SI with and without hypothyroidism. The algorithm, however, was not validated and had numerous other limitations: patients were not required to have a minimum duration of statin therapy; there was no requirement to attempt therapy with more than one statin; the analysis was conducted in a predominantly white population; and the study did not collect information on the specific statins taken by patients or on statin dosage.

The strength of the present study includes the ability to compare data iteratively between a clinically rich EMR and an administrative database. This facilitated the creation and refinement of algorithms to identify SI in a diverse patient population, receiving services in various practice settings. Furthermore, the definition of SI is comprehensive, reflecting the consensus statement of two key organizations and incorporating important clinical elements, such as statin rechallenge, statin intensity, temporal AEs, and a differential threshold for patient cases with rhabdomyolysis. Stratification by CV-risk status revealed the differential performance of the algorithms, which are not without limitations. Definitions of SI in both the administrative database and the EMR relied on sufficient documentation in both data sources. Administrative data were dependent on the documentation of a symptom or a diagnosis on a claim for reimbursement, and our gold standard EMR definition required documentation of the symptom in the medical record. Misclassification may also have occurred if patients filled statin prescriptions outside of insurance programs, if there were numerous verbal changes in prescriptions that were not reflected at the pharmacy, or if clinician concern over low LDL-C levels or patients' financial constraints resulted in dosage lowering.

Conclusions

The validated and comprehensive algorithms presented herein make possible the identification of SI from US administrative databases. These are the first validated algorithms for use in administrative databases to be made available to decision makers.

Acknowledgments

We thank Irfan Khan (Sanofi US, Bridgewater, NJ) for contributing to the initial study concept, data analysis, and manuscript review and Amelito Torres (OR Solutions, Waltham, MA) for SAS programming and analysis. Writing and editorial support for this manuscript was provided by MicroMass Communications, Inc.

Source of financial support: This study was funded by Sanofi US, Bridgewater, NJ, and Regeneron Pharmaceuticals, Inc., Tarrytown, NY.

Supplementary Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2016.03.1858 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[1] Taylor F, Huffman MD, Macedo AF, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013;1:CD004816.
[2] Cholesterol Treatment Trialists' (CTT) Collaboration. Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174,000 participants in 27 randomised trials. Lancet 2015;385:1397–405.
[3] Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2889–934.
[4] Jacobson TA, Ito MK, Maki KC, et al. National Lipid Association recommendations for patient-centered management of dyslipidemia, part 1: full report. J Clin Lipidol 2015;9:129–69.
[5] Lardizabal JA, Deedwania PC. Benefits of statin therapy and compliance in high risk cardiovascular patients. Vasc Health Risk Manag 2010;6:843–53.
[6] Cohen JD, Brinton EA, Ito MK, Jacobson TA. Understanding Statin Use in America and Gaps in Patient Education (USAGE): an internet-based survey of 10,138 current and former statin users. J Clin Lipidol 2012;6:208–15.
[7] Naci H, Brugts J, Ades T. Comparative tolerability and harms of individual statins: a study-level network meta-analysis of 246,955 participants from 135 randomized, controlled trials. Circ Cardiovasc Qual Outcomes 2013;6:390–9.
[8] Silva M, Matthews ML, Jarvis C, et al. Meta-analysis of drug-induced adverse events associated with intensive-dose statin therapy. Clin Ther 2007;29:253–60.
[9] Newman CB, Tobert JA. Statin intolerance: reconciling clinical trials and clinical experience [editorial]. JAMA 2015;313:1011–2.
[10] Finegold JA, Manisty CH, Goldacre B, et al. What proportion of symptomatic side effects in patients taking statins are genuinely caused by the drug? Systematic review of randomized placebo-controlled trials to aid individual patient choice. Eur J Prev Cardiol 2014;21:464–74.
[11] de Jong HJI, Klungel OH, van Dijk L, et al. Use of statins is associated with an increased risk of rheumatoid arthritis. Ann Rheum Dis 2012;71:648–54.
[12] Maji D, Shaikh S, Solanki D, Gaurav K. Safety of statins. Indian J Endocrinol Metab 2013;17:636–46.
[13] Bruckert E, Hayem G, Dejager S, et al. Mild to moderate muscular symptoms with high-dosage statin therapy in hyperlipidemic patients—the PRIMO study. Cardiovasc Drugs Ther 2005;19:403–14.
[14] Buettner C, Davis RB, Leveille SG, et al. Prevalence of musculoskeletal pain and statin use. J Gen Intern Med 2008;23:1182–6.
[15] El-Salem K, Ababneh B, Rudnicki S, et al. Prevalence and risk factors of muscle complications secondary to statins. Muscle Nerve 2011;44:877–81.
[16] Nichols GA, Koro CE. Does statin therapy initiation increase the risk for myopathy? An observational study of 32,225 diabetic and nondiabetic patients. Clin Ther 2007;29:1761–70.
[17] Zhang H, Plutzky J, Skentzos S, et al. Discontinuation of statins in routine care settings: a cohort study. Ann Intern Med 2013;158:526–34.
[18] Parker BA, Capizzi JA, Grimaldi AS, et al. Effect of statins on skeletal muscle function. Circulation 2013;127:96–103.
[19] Banach M, Rizzo M, Toth PP, et al. Statin intolerance—an attempt at a unified definition. Position paper from an International Lipid Expert Panel. Arch Med Sci 2015;11:1–23.
[20] Ahmad Z. Statin intolerance. Am J Cardiol 2014;113:1765–71.
[21] Mancini GJ, Tashakkor AY, Baker S, et al. Diagnosis, prevention, and management of statin adverse effects and intolerance: Canadian Working Group Consensus update. Can J Cardiol 2013;29:1553–68.
[22] Guyton JR, Bays HE, Grundy SM, Jacobson TA. An assessment by the Statin Intolerance Panel: 2014 update. J Clin Lipidol 2014;8:S72–81.
[23] Nicholls SJ, Brandrup-Wognsen G, Palmer M, Barter PJ. Meta-analysis of comparative efficacy of increasing dose of atorvastatin versus rosuvastatin versus simvastatin on lowering levels of atherogenic lipids (from VOYAGER). Am J Cardiol 2010;105:69–76.
[24] Weng T-C, Yang Y-K, Lin S-J, Tai S-H. A systematic review and meta-analysis on the therapeutic equivalence of statins. J Clin Pharm Ther 2010;35:139–51.
[25] Zutt R, van der Kooi AJ, Linthorst GE, et al. Rhabdomyolysis: review of the literature. Neuromuscul Disord 2014;24:651–9.
[26] Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37:360–3.
[27] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[28] Thacker EL, Muntner P, Zhao H, et al. Claims-based algorithms for identifying Medicare beneficiaries at high estimated risk for coronary heart disease events: a cross-sectional study. BMC Health Serv Res 2014;14:195.
[29] Hsieh C-Y, Chen C-H, Li C-Y, Lai M-L. Validating the diagnosis of acute ischemic stroke in a National Health Insurance claims database. J Formos Med Assoc 2015;114:254–9.
[30] Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. Pharmacoepidemiol Drug Saf 2012;21:129–40.
[31] Hlatky MA, Ray RM, Burwen DR, et al. Use of Medicare data to identify coronary heart disease outcomes in the Women's Health Initiative. Circ Cardiovasc Qual Outcomes 2014;7:157–62.
[32] Crestor® [package insert]. Wilmington, DE: AstraZeneca Pharmaceuticals, 2014.
[33] Huddy K, Dhesi P, Thompson PD. Do the frequencies of adverse events increase, decrease, or stay the same with long-term use of statins? Curr Atheroscler Rep 2013;15:301.
[34] Robison CD, Bair TL, Horne BD, et al. Hypothyroidism as a risk factor for statin intolerance. J Clin Lipidol 2014;8:401–7.