Identifying patients with asthma in primary care electronic medical record systems

Chart analysis–based electronic algorithm validation study

Nancy Xi MD Rebecca Wallace MD Gina Agarwal MBBS PhD David Chan MD Andrea Gershon MD FRCPC Samir Gupta MD FRCPC

Abstract

Objective To develop and test a variety of electronic medical record (EMR) search algorithms to allow clinicians to accurately identify their patients with asthma in order to enable improved care.

Design A retrospective chart analysis identified 5 relevant unique EMR information fields (electronic disease registry, cumulative patient profile, billing diagnostic code, medications, and chart notes); asthma-related search terms were designated for each field. The accuracy of each term was tested for its ability to identify the asthma patients among all patients whose charts were reviewed. Increasingly sophisticated search algorithms were then designed and evaluated by serially combining individual searches with Boolean operators.

Setting Two large academic primary care clinics in Hamilton, Ont.

Participants Charts for 600 randomly selected patients aged 16 years and older identified in an initial EMR search as likely having asthma (n = 150), chronic obstructive pulmonary disease (n = 150), other respiratory conditions (n = 150), or nonrespiratory conditions (n = 150) were reviewed until 100 patients per category were identified (or until all available names were exhausted). A total of 398 charts were reviewed in full and included.

Main outcome measures Sensitivity and specificity of each search for asthma diagnosis (against the reference standard of a physician chart review–based diagnosis).

Results Two physicians reviewed the charts identified in the initial EMR search using a standardized data collection form and ascribed the following diagnoses in 398 patients: 112 (28.1%) had asthma, 81 (20.4%) had chronic obstructive pulmonary disease, 104 (26.1%) had other respiratory conditions, and 101 (25.4%) had nonrespiratory conditions. Concordance between reviewers in chart abstraction diagnosis was high ($\kappa = 0.89$, 95% CI 0.80 to 0.97). Overall, the algorithm searching for patients who had *asthma* in their cumulative patient profiles or for whom an asthma billing code had been used was the most accurate (sensitivity of 90.2%, 95% CI 87.3% to 93.1%; specificity of 83.9%, 95% CI 80.3% to 87.5%).

Conclusion Usable, practical search algorithms that accurately identify patients with asthma in existing EMRs are presented. Clinicians can apply 1 of these algorithms to generate asthma registries for targeted quality improvement initiatives and outcome measurements. This methodology can be emulated for other diseases.

EDITOR'S KEY POINTS

- Electronic medical record (EMR) systems are increasingly being promoted as tools that enable improved safety and quality of care, particularly for chronic diseases such as asthma.

- Accurate registries of patients with asthma can be built through simple searches that can easily be performed in various primary care EMRs. Clinicians can use the described searches to accurately identify their patients with asthma for outcome and care monitoring or to target quality improvement initiatives.

- The methods used to identify and to test the accuracy of EMR search algorithms can also be used to establish search algorithms for other chronic diseases.

This article has been peer reviewed.
Can Fam Physician 2015;61:e474-83

Identifier les patients asthmatiques dans un système de dossiers médicaux électroniques, et ce, dans un contexte de première ligne

Étude de validation d'un algorithme électronique à partir d'une analyse de dossiers

Nancy Xi MD   Rebecca Wallace MD   Gina Agarwal MBBS PhD   David Chan MD   Andrea Gershon MD FRCPC   Samir Gupta MD FRCPC

Résumé

Objectif Élaborer et tester divers algorithmes de recherche pour des dossiers médicaux électroniques (DME) afin que les cliniciens puissent identifier avec précision leurs patients asthmatiques afin d'améliorer les soins.

Type d'étude Une analyse rétrospective de dossiers a permis d'identifier 5 champs d'information propres aux DME (registre électronique des maladies, profil cumulatif du patient, code de diagnostic pour la facturation, médication et notes au dossier); des termes de recherche en lien avec l'asthme ont été choisis pour chacun des champs. On a testé la précision avec laquelle chaque terme pouvait repérer les asthmatiques parmi tous les patients dont les dossiers ont été révisés. Des algorithmes de recherche de plus en plus sophistiqués ont ensuite été élaborés et évalués en combinant en série les recherches individuelles avec des opérateurs booléens.

Contexte Deux grandes cliniques universitaires de soins primaires à Hamilton, Ontario.

Participants On a révisé les dossiers de 600 patients de 16 ans et plus, choisis au hasard, qui, d'après une première révision des dossiers, souffraient probablement d'asthme (n=150), d'une autre condition respiratoire (n=150) ou d'un problème non respiratoire (n=150), jusqu'à ce qu'on ait identifié 100 patients dans chacune des catégories (ou jusqu'à épuisement de tous les noms). En tout, 308 dossiers ont été entièrement révisés et inclus.

Principaux paramètres à l'étude Sensibilité et spécificité de chaque recherche d'un diagnostic d'asthme (avec comme référence standard le diagnostic du médecin qui avait révisé le dossier).

Résultats Deux médecins ont révisé les dossiers qui avaient été retenus dans la recherche initiale sur les DME à l'aide d'un formulaire standardisé pour la collecte des données et ont attribué les diagnostics suivants à 398 patients: 112 diagnostics d'asthme (28,1%), 81 de maladie pulmonaire obstructive chronique (20,4%), 104 d'autres maladies respiratoires (26,1%) et 101 de conditions non respiratoires (25,4%). Il y avait une excellente concordance entre les réviseurs pour les diagnostics basés sur les dossiers (κ=0,89, IC à 95% 0,80 à 0,97). Dans l'ensemble, la recherche par algorithme chez les patients qui avaient le terme asthme dans leur profile cumulatif ou chez ceux pour lesquels un code de facturation d'asthme avait été utilisé était la méthode la plus sensible (sensibilité=90,2%, IC à 95% 87,3 à 93,1%; spécificité = 83,9%, IC à 95% 80,3 à 87,5%).

Conclusion Cet article décrit des algorithmes de recherche pratiques et d'utilisation facile permettant d'identifier avec précision les patients asthmatiques à partir de DME. Les cliniciens peuvent utiliser ces algorithmes pour élaborer des registres de l'asthme permettant d'entreprendre des mesures d'amélioration de la qualité des soins et d'évaluer les issues. Ces méthodes peuvent être utilisées pour d'autres maladies.

POINTS DE REPÈRE DU RÉDACTEUR

- On préconise de plus en plus l'utilisation de systèmes de dossiers médicaux électroniques (DME), qui sont susceptibles d'améliorer la sécurité et la qualité des soins, en particulier pour des maladies chroniques comme l'asthme.

- On peut créer des registres précis pour les patients asthmatiques grâce à des recherches qui peuvent facilement être effectuées dans les DME de divers milieux de soins primaires. Les cliniciens peuvent utiliser les outils de recherche décrits dans cet article pour identifier de façon précise leurs patients asthmatiques dans le but de surveiller les issues et le traitement ou pour envisager des mesures d'amélioration de la qualité.

- Les méthodes utilisées pour identifier les cas et vérifier la précision des algorithmes de recherche dans les DME peuvent aussi servir pour créer des algorithmes de recherche pour d'autres maladies chroniques.

Cet article fait l'objet d'une révision par des pairs.
Can Fam Physician 2015;61:e474-83

Asthma is the third most common chronic disease in adults in Canada. It affects 8.1% of the population, is increasing in prevalence,[1,2] and carries an annual economic burden of $1.8 billion.[3] Although several international bodies have produced evidence-based asthma diagnosis and management guidelines,[4] care gaps in asthma management remain prevalent, with 53% of Canadian patients having poorly controlled disease according to guideline criteria.[5] Key evidence-based care gaps responsible for this poor control are the underrecognition of suboptimal asthma control by both physicians and patients[5-9]; undertreatment of asthma[5-7]; and clinician failure to provide patients with a written asthma action plan.[5,9,10]

Given that most patients with asthma are seen in primary care,[11] primary care–based quality improvement might bridge these gaps. Electronic medical record (EMR) systems play an increasingly vital role in primary care[12,13] and have the potential to improve the safety and quality of care, reduce costs, and facilitate patient engagement. To realize these benefits, clinicians require the ability to compile valid and reliable disease-specific registries of patients to target for quality initiatives.[14-16] Such registries could facilitate EMR-based quality improvement strategies. For example, decision support integrated into the charts of patients with asthma could prompt clinicians about poor asthma control, provide guideline-based medication recommendations, and automatically fill in an electronic asthma action plan. To date, accurate and practical methods for practitioners to identify patients with asthma from their EMRs have not been presented.

We sought to develop and determine the accuracy of EMR-based search algorithms that would enable clinicians to easily and reliably identify patients with asthma within their practices, to optimize their care.

METHODS

This was a retrospective chart analysis conducted at 2 academic primary care clinics in Hamilton, Ont, with a total of 33 staff physicians (with rotating McMaster University family medicine residents) working in a capitated payment model with 27 300 registered patients. Clinics used the open-source Oscar EMR system, which is used by 1500 clinicians in the care of 2 million patients across Canada (**http://oscarcanada.org**).[17]

The study was approved by research ethics boards at McMaster University and St Michael's Hospital in Toronto, Ont. We retrieved charts of relevant patients aged 16 years and older, registered under any of 14 consenting physicians.

Search strategy design

Based on previous literature[18,19] and our clinical expertise, we identified 5 unique EMR information fields that could be used to determine asthma disease status, and search parameters to identify patients with asthma within each field (**Table 1**). We defined additional parameters expected to exclude patients with chronic obstructive pulmonary disease (COPD) (**Table 1**).

To evaluate the accuracy of search algorithms in distinguishing asthma from other respiratory conditions, we identified patients who had a higher likelihood of having respiratory disease than the general population.[20-22] We created approximately equal cohorts of patients likely to have each of the following conditions: asthma, COPD, other respiratory conditions (ie, not asthma or COPD), and nonrespiratory conditions. Patients with COPD and with other respiratory conditions were included to ensure that algorithms could differentiate asthma from these clinically similar conditions, whereas patients with nonrespiratory conditions acted as healthy controls. For simplicity, these patients were limited to those with hypertension or musculoskeletal disorders.[21] We first identified patients with a high likelihood of carrying 1 of these 4 diagnoses. For possible COPD, other respiratory conditions, and nonrespiratory conditions, we identified patients with relevant diagnoses in the electronic disease registry section of the EMR or a relevant corresponding billing diagnostic code billed within the past 3 years. Diagnoses listed in the electronic disease registry and the Ontario Health Insurance Plan billing codes used to identify patients are available from **CFPlus**.* For possible asthma, in addition to the above strategies, we identified patients who had been prescribed an inhaled asthma medication within the past 12 months (available from **CFPlus***), while excluding patients who had been prescribed tiotropium bromide or ipratropium bromide (medications used predominantly for COPD). Any patients who fulfilled criteria to be included in more than 1 category were placed in the category identified by the most recent relevant billing code or prescription.

After identifying all potential patients within each of these 4 diagnostic categories, we used a random number generator to choose 150 patients in each (600 total) for review, stratified by clinic site and by physician (**Figure 1**).

Chart review

Chart analysis was performed in a random order with respect to diagnostic category by 2 physicians (N.X., R.W.) in the family medicine training program. Each reviewer completed a standardized data collection form to determine which of the 4 diagnostic categories each patient actually belonged to (reference standard),

*__Diagnoses and billing codes used to identify patients__ and **inhaled medications used to identify patients with possible asthma** are available at **www.cfp.ca**. Go to the full text of the article online and click on **CFPlus** in the menu at the top right-hand side of the page.

Table 1. The EMR search fields and parameters

FIELD NAME	DESCRIPTION OF FIELD	INFORMATION ENTRY	PRIMARY SEARCH PARAMETERS USED	ADDITIONAL SEARCH PARAMETERS USED
Electronic disease registry	Displayed through a link that opens a new window. Used to document chronic disease diagnoses (based on ICD-9 codes)	Requires clinician to click on a link, which opens a separate window where chronic diseases can be added from a drop-down menu	All patients with code 493 (asthma or allergic bronchitis)	Exclusion criteria: patients with codes 491 (chronic bronchitis), 492 (emphysema), or 496 (other COPD)
Cumulative patient profile	4 boxed fields found at top of the electronic chart display (ongoing concerns, social history, medical history, and reminders). Used for documentation of previous and active patient psychosocial and medical issues	Requires free-text entry by clinicians	All patients with the word *asthma* anywhere in the cumulative patient profile	None
Billing diagnostic code	Displayed through a link that opens a window. Entry of a billing diagnostic code corresponding with the main reason for each visit is required for service payment (diagnostic codes are based on ICD-9 codes)	Requires clinician to choose a code from a drop-down menu at each clinical visit	All patients with ≥ 1 code 493 (asthma or allergic bronchitis) billed within the past 3 y	Exclusion criteria: patients with ≥ 1 codes 491, 492, or 496 billed within the past 3 y
Medications	Field found on the right side of the electronic chart display containing prescriptions made both through the EMR and by outside providers	Prescriptions made through the EMR are autopopulated from the prescription software (including generic and trade names, and doses); prescriptions made by outside providers require free-text entry by clinicians	All patients prescribed ≥ 1 inhaled asthma medication within the past y	Exclusion criteria: patients prescribed either tiotropium bromide or ipratropium bromide within the past y
Chart notes	Central display in the body of the electronic chart, where all providers enter notes during each patient encounter	Chief concern for each visit is typed in by the clinic receptionist based on the patient's description. The body of each note is typed in by the clinician	All patients seen within the past 3 y having the word *asthma* anywhere in the chart notes (excluding the chief concern area)	None

COPD—chronic obstructive pulmonary disease, EMR—electronic medical record.

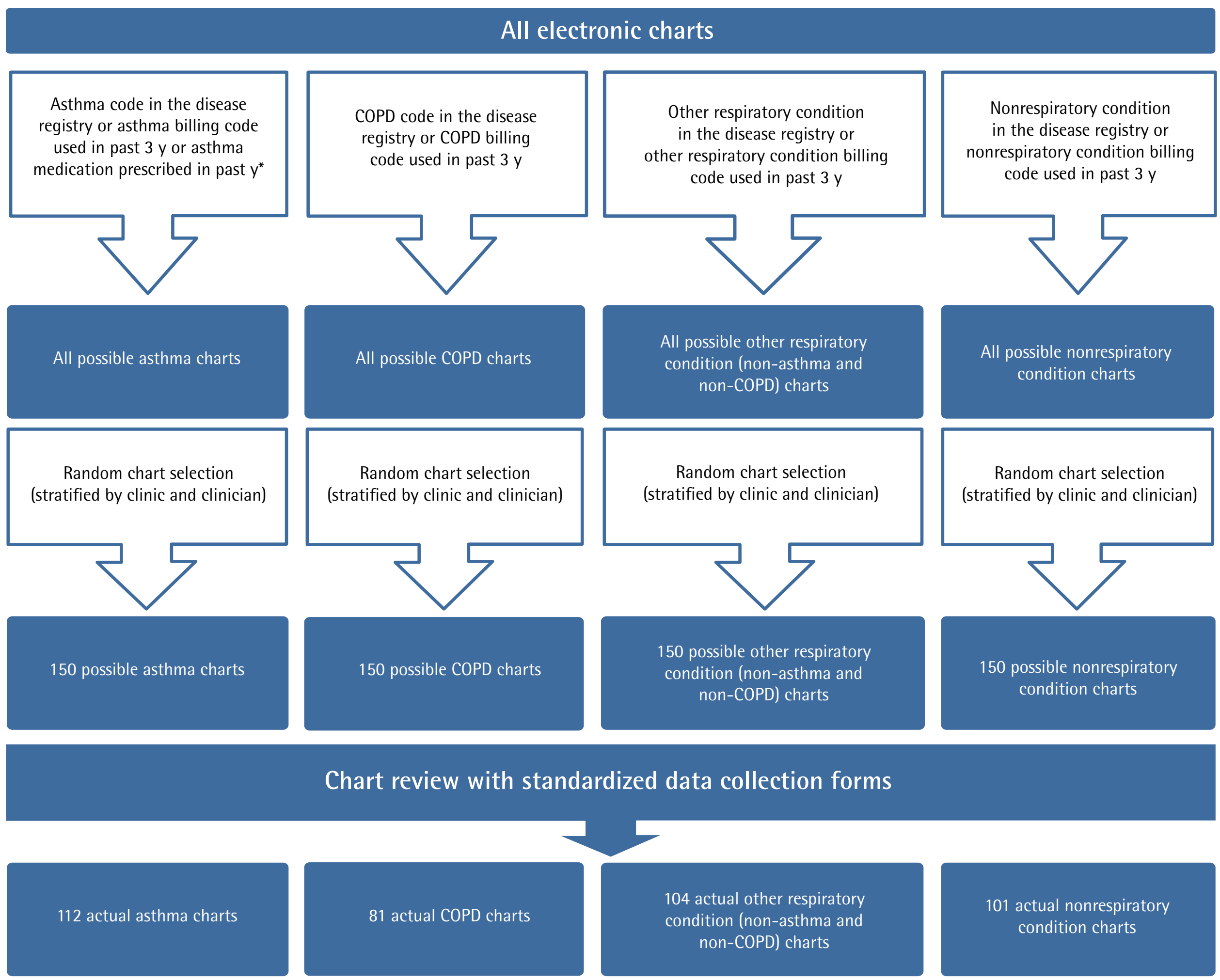
according to (in order of preference) pulmonary function (where available),[23,24] non–family physician specialist opinion (where available), and clinical diagnosis. A random sample of 20.0% of charts was analyzed by both reviewers to determine interrater reliability.

Reviewers placed each patient into 1 of the 4 diagnostic categories, until 100 patients per category were identified (or until all available names were exhausted). Patients with asthma and/or COPD who also had other respiratory or nonrespiratory conditions were categorized as *asthma*, *COPD*, or *asthma and COPD*, and patients with both other respiratory and nonrespiratory conditions were categorized as *other respiratory*. Uncertain cases were resolved by a consensus committee consisting of the chart reviewers, a general practitioner (G.A.), and a respirologist (S.G.).

Algorithm testing

We searched for each of the previously identified asthma parameters (**Table 1**) in the EMR information fields of all reviewed charts. We repeated these searches with exclusion criteria designed to eliminate patients with COPD (**Table 1**). Next, we searched 2 fields at a time by connecting any 2 individual searches with an "or" operator, and tested all possible combinations of individual searches. Based on these results, we proceeded to serially combine previous searches, using "or" combinations to increase sensitivity, and/or "and" combinations to increase specificity. Each algorithm was run separately by 2 study personnel to ensure identical results, and a random sample of 20 of the charts in each algorithm was verified to ensure that searches were accurate.

Figure 1. Chart identification and review methodology



COPD—chronic obstructive pulmonary disease.
*Excluding patients prescribed tiotropium bromide or ipratropium bromide.

Analysis

We calculated the interrater reliability of chart abstractor diagnosis using a κ statistic. We compared patient characteristics using the Student *t* test for continuous variables and the χ^2 test for categorical variables. Using chart abstractor diagnosis as the reference standard, we calculated true-positive, true-negative, false-positive, and false-negative rates; the sensitivity and specificity of each search; and the Youden index (J = sensitivity + specificity - 1). We also performed a discordance analysis (an in-depth chart review to identify reasons for misclassification in patients whose test results were falsely positive or falsely negative) in the algorithm with the highest Youden index. We calculated 95% CIs for test characteristics, assuming a binomial distribution. To maximize sensitivity, patients who had both asthma and COPD were categorized as *asthma*. We used SAS, version 9.3, for all analyses.

RESULTS

Study population

Reviewers assessed 460 charts, of which 41 (8.9%) had no available data, 21 (4.6%) were duplicates, and 398 (86.5%) were fully reviewed. These 398 patients consisted of 112 (28.1%) with asthma, 81 (20.4%) with COPD, 104 (26.1%) with other respiratory conditions (neither asthma nor COPD), and 101 (25.4%) with nonrespiratory conditions. Nine of 112 (8.0%) asthma patients had coexisting COPD. Concordance between the 2 reviewers in chart abstraction diagnosis was high (κ=0.89, 95% CI 0.80 to 0.97) (**Table 2**).

Table 2. Patient characteristics

CHARACTERISTIC	PATIENTS WITH ASTHMA (N = 112*)	PATIENTS WITHOUT ASTHMA (N = 286)	P VALUE
Mean (SD) age, y	44.5 (19.1)	57.9 (17.3)	< .001
Sex, n (%)			
• Female	82 (73.2)	172 (60.1)	.02
• Male	30 (26.8)	114 (39.9)	
Smoking status, n (%)			
• Non-smoker	56 (50.0)	96 (33.6)	.005
• Ex-smoker	15 (13.4)	77 (26.9)	
• Smoker	22 (19.6)	69 (24.1)	
• Not documented	19 (17.0)	44 (15.4)	
Electronic disease registry diagnosis, n (%)			
• Asthma	8 (7.1)	1 (0.3)	.04
• COPD	4 (3.6)	10 (3.5)	1.0
Cumulative patient profile diagnosis, n (%)			
• Asthma	71 (63.4)	27 (9.4)	< .001
Atopy, n (%)	54 (48.2)	44 (15.4)	< .001
Comorbidities, n (%)			
• Cardiovascular illness	8 (7.1)	10 (3.5)	.12
• Diabetes mellitus	11 (9.2)	30 (10.5)	.84
• Psychiatric illness	24 (21.4)	66 (23.1)	.72
Relevant notes in chart (past 3 y), n (%)			
• Use of word *asthma*	94 (83.9)	40 (14.0)	< .001
• Physician asthma diagnosis[†]	99 (88.4)	54 (18.9)	< .001
Non–family physician specialist consultation (past 9 y[‡]), n (%)			
• Pulmonologist	27 (24.1)	25 (8.7)	< .001
• Allergist	14 (12.5)	14 (4.9)	.008
• Both	4 (3.6)	0 (0.0)	.006
Previous diagnostic testing (past 9 y[‡]), n (%)			
• Spirometry	50 (44.6)	87 (30.4)	.007
• Bronchodilator challenge performed	27 (24.1)	62 (21.7)	
• Positive bronchodilator response	13 (11.6)	4 (1.4)[§]	
• Methacholine challenge performed	14 (12.5)	10 (3.5)	< .001
• Positive methacholine challenge result	12 (10.7)[‖]	0 (0.0)	
• Chest x-ray scan	58 (51.8)	147 (51.4)	.94
• Skin-prick test	27 (24.1)	20 (7.0)	< .001
Health care use (past 9 y[‡]), n (%)			
• Emergency department visit	22 (19.6)	22 (7.7)	< .001
• Hospitalization	9 (8.0)	14 (4.9)	.23
Asthma medications prescribed in past y, n (%)			
• None	16 (14.3)	199 (69.6)	< .001
• Short-acting bronchodilator only	17 (15.2)	25 (8.7)	
• Inhaled corticosteroid[¶]	37 (33.0)	29 (10.1)	
• Inhaled corticosteroid and long-acting bronchodilator[¶]	30 (26.8)	32 (11.2)	
• Leukotriene receptor antagonist only[¶]	12 (10.7)	1 (0.3)	
Billing codes used (past 3 y), n (%)			
• Asthma	89 (79.5)	30 (10.5)	< .001
• COPD	14 (12.5)	66 (23.1)	.02
• Other respiratory illness	62 (55.4)	200 (69.9)	.006

COPD—chronic obstructive pulmonary disease.
*Includes 9 patients with concurrent COPD.
[†]Based on documented physician impression in chart notes.
[‡]Based on data availability.
[§]All 4 of these patients were seen by non–family physician specialists who diagnosed COPD without asthma.
[‖]The result for 1 patient was not available, and the other missing result was for a patient who had initial negative methacholine challenge results, but subsequently had a meaningful bronchodilator response and spirometric response to a budesonide-formoterol turbo-inhaler, and was given an asthma diagnosis by an allergist.
[¶]With or without a short-acting bronchodilator.

Search algorithms

Results from each unique search query and the 5 algorithms with the best test characteristics (based on the Youden index) are presented in **Table 3.** All tested search algorithms as well as search characteristics by clinic are available upon request. True-positive and false-negative rates for these searches and for the most sensitive and most specific individual algorithms are represented in **Figure 2**.

Discordance analysis

The algorithm combining *asthma* in the cumulative patient profile (CPP) or use of billing code 493 (asthma or allergic bronchitis) had the highest Youden index (**Table 3**) and was used for the discordance analysis. There were 11 false-negative results and 46 false-positive results. Among the 11 false-negative results, 4 charts (36.4%) simply did not have *asthma* in the CPP and had not been billed for asthma despite clear chart documentation of asthma. Of the remaining 7 charts, 6 (85.7%) had been diagnosed with asthma by an outside specialist. In these cases, we suspect that clinicians might have been less likely to update the CPP because the diagnosis was made elsewhere. Also, 6 of these 7 had COPD in addition to asthma, and clinicians appeared to default to using COPD rather than asthma billing codes for respiratory-related visits in these patients.

Among the 46 false-positive results, 14 (30.4%) had COPD as opposed to asthma, and clinicians might have confused this with asthma when completing the CPP or when billing. Another 13 (28.3%) were initially suspected of having had asthma, but later had negative objective test results for asthma. Five of these patients also saw other specialists and received the following alternate diagnoses: eosinophilic bronchitis (n=2); bronchiectasis (n=2); and gastroesophageal reflux disease (n=1). Some of these charts had falsely positive results because the asthma billing code was used at the time of the initial diagnostic suspicion, and others because the CPP had not been updated in light of objective testing and non–family physician specialist results. Another 16 (34.8%) had an upper or lower respiratory tract infection that resulted in (usually) isolated use of an asthma billing code (with no other evidence of asthma in the chart). Finally, 3 (6.5%) had a single asthma billing code with no plausible explanation and no other evidence of asthma in the chart.

DISCUSSION

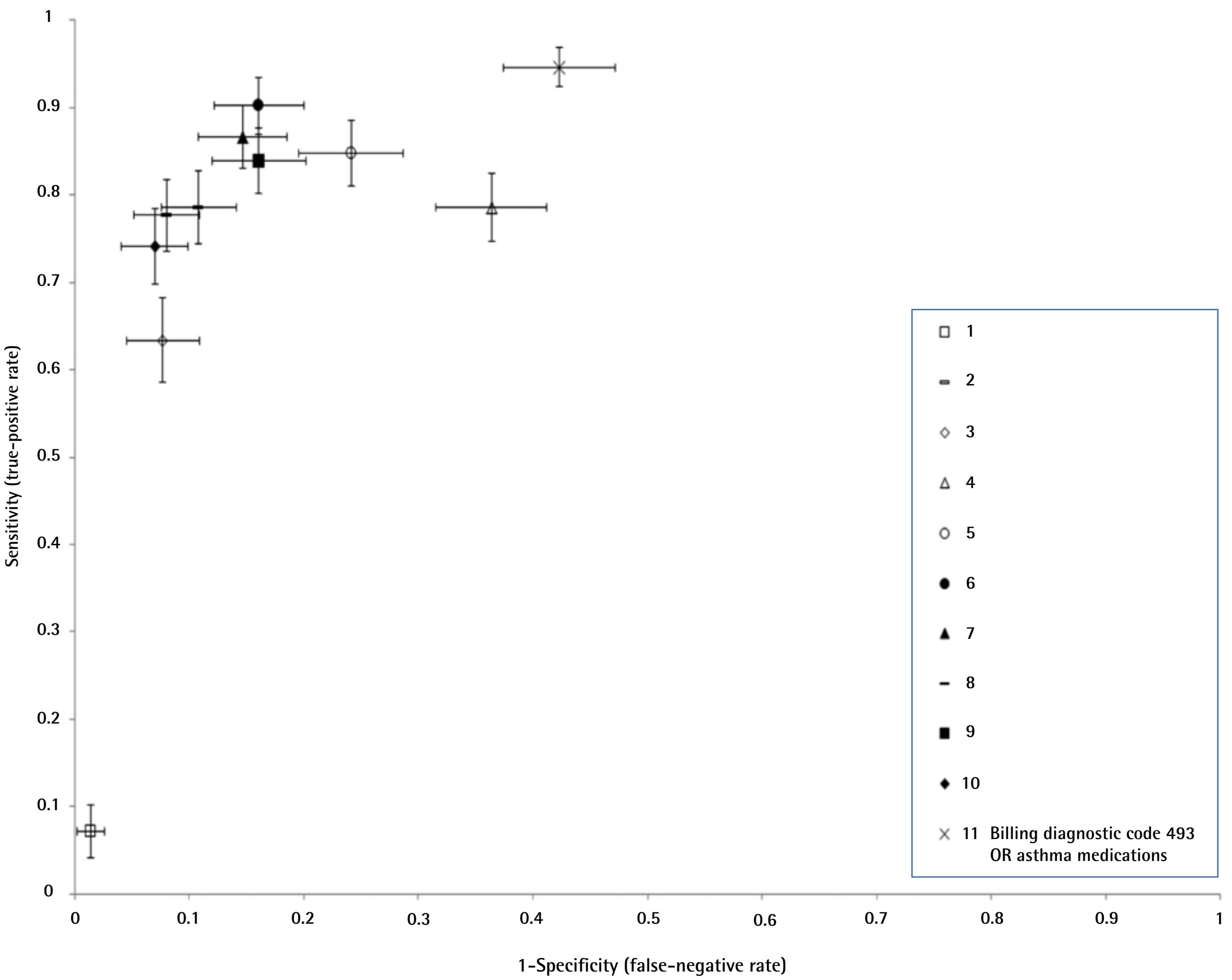
In this article, we present algorithms that will enable clinicians to accurately identify their patients with

Table 3. Results from each unique search query and the 5 algorithms with the best test characteristics (based on the Youden index)

SEARCH STRATEGY	RESULTS				SENSITIVITY, % (95% CI)	SPECIFICITY, % (95% CI)	YOUDEN INDEX*
	TP	FP	FN	TN			
Individual search queries							
1. *Asthma* in disease registry	8	4	104	282	7.1 (4.6 to 9.7)	98.6 (97.4 to 99.8)	0.057
2. Billing diagnostic code 493	88	31	24	255	78.6 (74.5 to 82.6)	89.2 (86.1 to 92.2)	0.677
3. *Asthma* in CPP	71	22	41	264	63.4 (58.7 to 68.1)	92.3 (89.7 to 94.9)	0.557
4. Asthma medications	88	104	24	182	78.6 (74.5 to 82.6)	63.6 (58.9 to 68.4)	0.422
5. *Asthma* in chart notes	95	69	17	217	84.8 (81.3 to 88.3)	75.9 (71.7 to 80.1)	0.607
Search algorithms							
6. *Asthma* in CPP OR billing diagnostic code 493	101	46	11	240	90.2 (87.3 to 93.1)	83.9 (80.3 to 87.5)	0.741
7. *Asthma* in CPP OR billing diagnostic code 493 (with exclusion of diagnostic codes 491, 492, and 496)	97	42	15	244	86.6 (83.3 to 90.0)	85.3 (81.8 to 88.8)	0.719
8. (*Asthma* in chart notes OR asthma medications) AND billing diagnostic code 493	87	23	25	263	77.7 (73.6 to 81.8)	92.0 (89.3 to 94.6)	0.696
9. (Billing diagnostic code 493 OR asthma medications) AND *asthma* in chart notes	94	46	18	240	83.9 (80.3 to 87.5)	83.9 (80.3 to 87.5)	0.678
10. Billing diagnostic code 493 AND *asthma* in chart notes	83	20	29	266	74.1 (69.8 to 78.4)	93.0 (90.5 to 95.5)	0.671

CPP—cumulative patient profile, FN—false negative, FP—false positive, TN—true negative, TP—true positive.
*J = sensitivity + specificity - 1.

Figure 2. Operating characteristics of individual search queries, and most specific, most sensitive, and most accurate search algorithms: *Operating characteristics are represented for each individual search query and the 5 most accurate queries (as presented in Table 3). The most specific tested algorithm was already represented in Table 3 (*asthma *in the disease registry). The most sensitive tested algorithm was not represented in Table 3, and was added here for purposes of comparison (11). Error bars represent 95% CIs. Algorithms with a high true-positive rate and a low false-negative rate are favoured (upper left-hand corner).*



asthma by searching data routinely recorded in Canadian EMR systems. These will enable clinicians to create asthma registries that can be used for practice audits and to target and evaluate quality improvement measures.

A simple individual search for use of billing code 493 provided a reasonable balance of sensitivity (78.6%) and specificity (89.2%). However, combining individual searches into algorithms further improved their diagnostic yield, with the best overall accuracy achieved by the combination of *asthma* in the CPP or use of billing code 493. As expected, further excluding patients for whom a COPD code had been billed resulted in an algorithm with a higher specificity but a lower sensitivity owing to the fact that patients with a combination of asthma and COPD were no longer counted (**Table 3**).

Observed sensitivities and specificities of individual search strategies and findings of the discordance analysis offer insight into both care and charting patterns. Prescription of asthma medications was neither particularly sensitive (78.6%) nor specific (63.6%). Sensitivity was limited because medications are not always required in mild asthma (14.3% of asthma subjects had not been prescribed a medication within the past year) (**Table 2**). Specificity was reduced both by formulations of these drugs used for non-asthma conditions (eg, nasal steroid sprays) and because asthma medications are also used for COPD. Indeed, exclusion of patients prescribed medications used predominantly for COPD increased the specificity of this algorithm to 73.4%.

The discordance analysis demonstrated that clinicians do not consistently update the CPP in light of other specialist findings, and might confuse COPD and asthma, both of which contributed to false-negative and false-positive results. A tendency to confuse asthma with other conditions that might be associated with wheezing, including respiratory tract infections, was another source of false-positive results.

Previous research has focused on accurate identification of asthma for epidemiologic studies through prescription or health administrative databases. By linking data to a centralized Danish prescription registry, Moth and colleagues determined that at least 1 prescription for any asthma medication (with the exception of β_2-agonists) in a 12-month period had a sensitivity of 96% and a specificity of 43% for a diagnosis of asthma in children.[25] Other authors used a large health administrative database to propose an asthma diagnostic algorithm consisting of 2 or more ambulatory care visits or 1 or more hospitalizations for asthma in 2 years, with a sensitivity of 84% and a specificity of 76%.[21] Afzal and colleagues described a novel machine-learning approach to derive a case-detection algorithm for childhood asthma, with a sensitivity of 96% and specificity of 90%, using both text and coded data in a large primary care database.[26] However, applicability is limited to the specific database that it was developed for, and to researchers with access to the sophisticated search system used. These previous studies report diagnostic algorithms for use in population-based epidemiologic studies rather than practice-level programs, where the required data linkage and special tools would not be available to clinicians. In comparison, our algorithms can be applied by individual practitioners through searches within their own EMRs in order to leverage the unique data elements available in EMRs to yield stronger operating characteristics.

Clinicians and researchers seeking to identify a target population must choose carefully among the 35 algorithms that we tested, depending on their requirements for sensitivity, specificity, and the specific types of patients desired. For example, identifying patients through asthma medication prescriptions yielded 192 patients with asthma, whereas the more specific and less sensitive "*asthma* in CPP" search yielded only 93 patients with asthma. These differences would obviously also affect the results of quality of care assessments in the identified asthma population. For example, 93 of 192 (48.4%) of those identified by the medication prescription search had spirometry performed, whereas 40 of 93 (43.0%) patients identified by the CPP search had spirometry performed (data not shown).

Strengths and limitations

The strengths of our study include our inclusion of comparison groups consisting of patients with diseases similar to asthma, the exhaustive number of algorithms tested, the strong performance characteristics of leading algorithms, and their practical applicability in primary care. The main limitation of our study is that in intentionally targeting approximately equal proportions of patients in each disease category, we did not use a true population sample, and therefore could not calculate positive and negative predictive values for each algorithm. However, this approach was required to ensure that algorithms could accurately differentiate asthma from clinically similar conditions, and has been applied in previous studies for the same reason.[20-22] A minority of our patients with asthma had undergone objective confirmatory testing, and our reference standard relied on clinical data in most cases. However, this reflects current real-world practice.[27] Accordingly, our algorithms enable clinicians to accurately identify the patients in their practices with a *clinical* diagnosis of asthma. The study was conducted at 2 academic clinics, and we did not include non-teaching community settings. Algorithm operating characteristics were similar between the 2 clinics, despite their unique billing, prescribing, and charting cultures. However, our findings should be validated in a non-academic setting. Similarly, we tested our algorithms in a single EMR system; several different EMR vendors exist, and each system has a unique set of features and usability constraints. However, we limited our searches to data elements that are fundamental features of the patient encounter,[28] are required fields for most EMRs,[29] and have been searched successfully in other EMRs.[19,30,31] Accordingly, our algorithms can be used across different EMR systems. It should be noted that we did not test algorithms based on a billing code, text string, or prescription occurring repeatedly over a period of time, as such searches are not feasible in most EMR systems. Although we did assess an array of algorithms from simple to complex, most are suitable for basic EMR packages with limited search capabilities. If required, individual search results could be exported to database software, where our more complex algorithms can be enacted through data manipulation.

Conclusion

Electronic medical records are a unique tool for primary care clinicians to efficiently measure local disease-specific outcomes and quality of care, and to target corresponding quality improvement. To achieve this, clinicians require the ability to generate accurate disease-specific registries without substantial cost or complexity. We present several EMR search algorithms that can easily be applied to generate accurate asthma registries. These registries can be used by clinicians and researchers alike for audit and feedback initiatives, asthma outcome monitoring, asthma epidemiology monitoring, targeting for asthma-related pay-for-performance incentives, and preventive or active care interventions, including point-of-care interventions.[32] These methods can also be emulated to establish EMR identification algorithms for other chronic diseases.

Dr Xi is Lecturer in the Department of Family Medicine at the University of Toronto in Ontario and Primary Care Program Clinical Teaching Unit Lead with the Trillium Health Partners at the Credit Valley Hospital in Mississauga, Ont. **Dr Wallace** is a practising family physician in Toronto. **Dr Agarwal** is Associate Professor in the Department of Family Medicine at McMaster University in Hamilton, Ont. **Dr Chan** is Professor and Director of Information Technology in the Department of Family Medicine at McMaster University. **Dr Gershon** is Research Director and a physician in the Division of Respirology at Sunnybrook Health Sciences Centre, Lung Health Lead and Scientist in the Institute for Clinical Evaluative Sciences, and Associate Professor in the Department of Medicine at the University of Toronto in Ontario. **Dr Gupta** is a scientist in the Keenan Research Centre at the Li Ka Shing Knowledge Institute and Associate Professor in the Department of Medicine at St Michael's Hospital at the University of Toronto.

Acknowledgment
This work was supported by the Canadian Institutes of Health Research and the Ontario Lung Association.

Contributors
Dr Gupta conceived of the study and performed analyses and manuscript preparation; **Dr Gershon** helped to design the study and reviewed the manuscript; **Drs Xi** and **Wallace** performed data collection, analysis, and manuscript preparation; and **Drs Agarwal** and **Chan** helped with study design, analysis, and manuscript preparation.

Competing interests
None declared

Correspondence
Dr Samir Gupta; e-mail **guptas@smh.ca**

References

1. Statistics Canada [website]. *Table 105-0501. Health indicator profile, annual estimates, by age group and sex, Canada, provinces, territories, health regions (2013 boundaries) and peer groups.* Ottawa, ON: Statistics Canada; 2015. Available from: **www5.statcan.gc.ca/cansim/a26?lang=eng&id=1050501&p2=17**. Accessed 2014 Jun 20.
2. Public Health Agency of Canada. *Life and breath: respiratory disease in Canada, 2007*. Ottawa, ON: Public Health Agency of Canada; 2007.
3. Ontario Lung Association [website]. *Your lungs, your life: insights and solutions to lung health in Ontario.* Toronto, ON: Ontario Lung Association; 2011. Available from: **www.on.lung.ca/Document.Doc?id=868**. Accessed 2014 Jun 20.
4. Lougheed MD, Lemiere C, Ducharme FM, Licskai C, Dell SD, Rowe BH, et al. Canadian Thoracic Society 2012 guideline update: diagnosis and management of asthma in preschoolers, children and adults. *Can Respir J* 2012;19(2):127-64.
5. FitzGerald JM, Boulet LP, McIvor RA, Zimmerman S, Chapman KR. Asthma control in Canada remains suboptimal: The Reality of Asthma Control (TRAC) study. *Can Respir J* 2006;13(5):253-9.
6. Chapman KR, Boulet LP, Rea RM, Franssen E. Suboptimal asthma control: prevalence, detection and consequences in general practice. *Eur Respir J* 2008;31(2):320-5. Epub 2007 Oct 24.
7. Klomp H, Lawson JA, Cockcroft DW, Chan BT, Cascagnette P, Gander L, et al. Examining asthma quality of care using a population-based approach. *CMAJ* 2008;178(8):1013-21.
8. Boulet LP, Devlin H, O'Donnell DE. The Physicians' Practice Assessment Questionnaire on asthma and COPD. *Respir Med* 2011;105(1):8-14. Epub 2010 Aug 24.
9. Boulet LP, Bourbeau J, Skomro R, Gupta S. Major care gaps in asthma, sleep and chronic obstructive pulmonary disease: a road map for knowledge translation. *Can Respir J* 2013;20(4):265-9.
10. Wiener-Ogilvie S, Pinnock H, Huby G, Sheikh A, Partridge MR, Gillies J. Do practices comply with key recommendations of the British Asthma Guideline? If not, why not? *Prim Care Respir J* 2007;16(6):369-77.
11. Global Asthma Physician and Patient Survey [website]. *Global adult highlights.* Global Asthma Physician and Patient Survey; 2005. Available from: **www.gappsurvey.org/media-key-findings.html**. Accessed 2013 Nov 20.
12. Xierali IM, Hsiao CJ, Puffer JC, Green LA, Rinaldo JC, Bazemore AW, et al. The rise of electronic health record adoption among family physicians. *Ann Fam Med* 2013;11(1):13-9.
13. College of Family Physicians of Canada, Canadian Medical Association, Royal College of Physicians and Surgeons of Canada. *National Physician Survey, 2013. Results by FP/GP or other specialist, sex, age, and all physicians.* Mississauga, ON: College of Family Physicians of Canada; 2013. Available from: **http://nationalphysiciansurvey.ca/wp-content/uploads/2013/08/2013-National-EN-Q12.pdf**. Accessed 2013 Dec 22.
14. Canadian Medical Association [website]. *Health care transformation in Canada*. Ottawa, ON: Canadian Medical Association; 2010. Available from: **http://policybase.cma.ca/dbtw-wpd/PolicyPDF/PD10-05.PDF**. Accessed 2015 Sep 4.
15. Shadd JD, Cejic S, Terry A, Ryan BL, Stewart M, Thind A. You and your EMR: the research perspective. Part 3. Answering practice-level questions. *Can Fam Physician* 2012;58:344-5.
16. Terry AL, Cejic S, Ryan BL, Shadd JD, Stewart M, Fortin M, et al. You and your EMR: the research perspective. Part 4. Optimizing EMRs in primary health care practice and research. *Can Fam Physician* 2012;58:705-6.
17. Oscar Canada Users Society [website]. *About OSCAR*. Vancouver, BC: Oscar Canada Users Society; 2013. Available from: **www.oscarcanada.org/about-oscar**. Accessed 2011 Nov 20.
18. Donahue JG, Weiss ST, Goetsch MA, Livingston JM, Greineder DK, Platt R. Assessment of asthma using automated and full-text medical records. *J Asthma* 1997;34(4):273-81.
19. Pacheco JA, Avila PC, Thompson JA, Law M, Quraishi JA, Greiman AK, et al. A highly specific algorithm for identifying asthma cases and controls for genome-wide association studies. *AMIA Annu Symp Proc* 2009;2009:497-501.
20. Di Lorenzo G, Mansueto P, Esposito-Pellitteri M, Ditta V, Castello F, Lo Bianco C, et al. The characteristics of different diagnostic tests in adult mild asthmatic patients: comparison with patients with asthma-like symptoms by gastro-oesophageal reflux. *Respir Med* 2007;101(7):1455-61. Epub 2007 Mar 13.
21. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009;16(6):183-8.
22. To T, Dell S, Dick PT, Cicutto L, Harris JK, MacLusky IB, et al. Case verification of children with asthma in Ontario. *Pediatr Allergy Immunol* 2006;17(1):69-76.
23. Global Initiative for Asthma. *Global strategy for asthma management and prevention, revised 2014*. Global Initiative for Asthma; 2014. Available from: **www.ginasthma.org/local/uploads/files/GINA_Report_2014_Aug12_1.pdf**. Accessed 2015 Sep 8.
24. Global Initiative for Chronic Obstructive Lung Disease. *Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease, updated 2013*. Global Initiative for Chronic Obstructive Lung Disease; 2013. Available from: **www.goldcopd.org/uploads/users/files/GOLD_Report_2013_Feb20.pdf**. Accessed 2013 Jun 1.
25. Moth G, Vedsted P, Schiøtz PO. Identification of asthmatic children using prescription data and diagnosis. *Eur J Clin Pharmacol* 2007;63(6):605-11. Epub 2007 Mar 27.
26. Afzal Z, Engelkes M, Verhamme KM, Janssens HM, Sturkenboom MC, Kors JA, et al. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases. *Pharmacoepidemiol Drug Saf* 2013;22(8):826-33. Epub 2013 Apr 17.
27. Gershon AS, Victor JC, Guan J, Aaron SD, To T. Pulmonary function testing in the diagnosis of asthma: a population study. *Chest* 2012;141(5):1190-6. Epub 2011 Oct 26.
28. College of Physicians and Surgeons of Ontario [website]. *Medical records*. Toronto, ON: College of Physicians and Surgeons of Ontario; 2012. Available from: **www.cpso.on.ca/policies/policies/default.aspx?ID=1686**. Accessed 2014 Jun 1.
29. OntarioMD. *Electronic medical records. Specification. Appendix A—EMR baseline requirements.* Toronto, ON: eHealth Ontario; 2011. Available from: **www.ontariomd.ca/idc/groups/public/documents/omd_file_content_item/omd011483.pdf**. Accessed 2013 Nov 20.
30. Allen-Ramey C, Nelsen LM, Leader JB, Mercer D, Kirchner HL, Jones JB. Electronic health record-based assessment of oral corticosteroid use in a population of primary care patients with asthma: an observational study. *Allergy Asthma Clin Immunol* 2013;9(1):27.
31. Vollmer WM, O'Connor EA, Heumann M, Frazier EA, Breen V, Villnave J, et al. Searching multiple clinical information systems for longer time periods found more prevalent cases of asthma. *J Clin Epidemiol* 2004;57(4):392-7.
32. Smith JR, Noble MJ, Musgrave S, Murdoch J, Price GM, Barton GR, et al. The At-Risk Registers in Severe Asthma (ARRISA) study: a cluster-randomised controlled trial examining effectiveness and costs in primary care. *Thorax* 2012;67(12):1052-60. Epub 2012 Aug 31.