Published in final edited form as:

J Asthma. 2018 September ; 55(9): 1035–1042. doi:10.1080/02770903.2017.1389952.

A computable phenotype for asthma case identification in adult and pediatric patients: External validation in The Chicago Area Patient-Outcomes Research Network (CAPriCORN)

Majid Afshar^{1,*}, Valerie G. Press^{2,*}, Rachel G. Robison³, Abel N. Kho⁴, Sindhura Bandi⁵, Ashvini Biswas⁵, Pedro C. Avila⁶, Harsha Vardhan Madan Kumar⁷, Byung Yu⁸, Edward T. Naureckas², Sharmilee M. Nyenhuis^{9,⊥}, and Christopher D. Codispoti^{5,⊥}

¹Division of Pulmonary and Critical Care, Loyola University Chicago Stritch School of Medicine, Maywood, IL USA

²Department of Medicine, University of Chicago Medicine, Chicago, IL USA

³Division of Allergy/Immunology, Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL USA

⁴Center for Health Information Partnerships, Northwestern University Feinberg School of Medicine, Chicago, IL USA

⁵Department of Immunology and Microbiology, Allergy/Immunology Section, Rush University Medical Center, Chicago, IL USA

⁶Department of Medicine, Northwestern University, Chicago, IL USA

⁷University of Illinois at Chicago, Chicago, IL USA

⁸Department of Pediatrics, Division of Allergy/Immunology, Cook County Health & Hospital System, Chicago, IL USA

⁹Department of Medicine, University of Illinois at Chicago, Chicago, IL USA

Abstract

Objective: Comprehensive, rapid, and accurate identification of patients with asthma for clinical care and engagement in research efforts is needed. The original development and validation of a computable phenotype for asthma case identification occurred at a single institution in Chicago and demonstrated excellent test characteristics. However, its application in a diverse payer mix, across different health systems and multiple electronic health record vendors, and in both children and adults was not examined. The objective of this study is to externally validate the computable phenotype across diverse Chicago institutions to accurately identify pediatric and adult patients with asthma.

* co-first authors. [⊥] co-senior authors. Corresponding Author: Valerie G. Press, MD, MPH, Department of Medicine, University of Chicago, 5841 South Maryland Avenue, MC 5000, Chicago, IL 60637. vpress@medicine.uchicago.edu. 773-702-1295.

Declaration of Interest:

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Methods: A cohort of 900 asthma and control patients was identified from the electronic health record between January 1, 2012 and November 30, 2014. Two physicians at each site independently reviewed the patient chart to annotate cases.

Results: The inter-observer reliability between the physician reviewers had a κ -coefficient of 0.95 (95% CI 0.93–0.97). The accuracy, sensitivity, specificity, negative predictive value, and positive predictive value of the computable phenotype were all above 94% in the full cohort.

Conclusions: The excellent positive and negative predictive values in this multi-center external validation study establish a useful tool to identify asthma cases in the electronic health record for research and care. This computable phenotype could be used in large-scale comparative-effectiveness trials.

Keywords

asthma; electronic health record; algorithm

Introduction

Asthma affects nearly 18 million adults and over 6 million children costing nearly 60 billion dollars in 2007.(1, 2) Asthma guidelines recommend either spirometry-based bronchodilator response or methacholine challenge to confirm diagnosis of asthma.(3) However, clinicians infrequently employ these methods in clinical practice. Clinicians rely more on medical history of asthma symptoms and response to asthma medications to diagnose asthma. This adds to the challenge of identifying asthma patients from large networks and health systems. Comprehensive, rapid, and accurate identification of patients with asthma for clinical care and engagement in research efforts is needed. In 2010, the Patient Centered Outcomes Research Institute (PCORI) was established for comparative effectiveness research. Funding from PCORI enabled creation of the Patient-Centered Clinical Research Network (PCORnet) to support large data marts across multiple health systems to leverage electronic health records (EHR) for conduct of comparative effectiveness research.(4, 5) The PCORI infrastructure has made it feasible to develop more sophisticated algorithms beyond billing data with incorporation of medication use and physician diagnoses to better discriminate disease phenotypes from the EHR.

As a Clinical Data Research Network (CDRN) of PCORI, the Chicago Area Patient-Centered Outcomes Research Network(CAPriCORN) is a collection of ten private, public, state, and county health systems that serve the third most populous city in the United States with an overall strategy to develop a cross-institutional infrastructure for sustainable, patient-centered outcomes research in Chicago and nationally.(6) Since asthma is a major health problem in Chicago, CAPriCORN researchers have collaborated to address the myriad of clinical challenges in asthma diagnosis and management.(7) CAPriCORN asthma and informatics researchers from several institutions have convened to create the infrastructure and processes to identify and recruit subjects with asthma for large-scale asthma research. In the process, researchers have developed a computable phenotype to identify adult patients with asthma using the EHR at one of the CAPriCORN institutions, which demonstrated excellent accuracy and precision.(8)

In this study, we aimed to externally validate the asthma computable phenotype across other CAPriCORN institutions which represent a more diverse payer mix, incorporate different electronic health record systems, include both children and adults, and address the potential limitations of institution-specific practice behaviors. We hypothesize the test characteristics of the computable phenotype will perform similarly across all CAPriCORN institutions.

Methods

Population and Participating Health Systems in CAPriCORN

The CAPriCORN institutions serve the greater Chicago metropolitan with a population of approximately 9.5 million people. The participating institutions for this study include four academic medical centers (Rush University Health, University of Chicago, University of Illinois Hospital and Health Sciences System-UI Health and the Ann and Robert H. Lurie Children's Hospital of Chicago) and a public county hospital (Cook County Health and Hospital System). The mixed payer population ranges from over 70% uninsured at the county hospital to 70% privately insured at the private academic centers.(9) The five CAPriCORN institutions have EHR systems with components supported by major vendor systems, including Epic (Rush University, University of Chicago and Lurie Children's) and Cerner (UI Health and Cook County). The information systems at all CAPriCORN institutions include a relational data warehouse with structured reporting functionality. (9) The Institutional Review Board (IRB) at each site provided full approval for this study.

Cohort Identification

CAPriCORN developed a computable phenotype through iterative collaboration among subject matter expert cohort leads, informaticians, and structured query language (SQL) developers. The collaboration resulted in developing a computable phenotype that is compatible with CAPriCORN's Common Data Model and PCORI's Common Data Model. The internal validation of the computable phenotype occurred at one CAPriCORN site (Northwestern University) using the EHR-derived Enterprise Data Warehouse.(8) EPIC EHR served as the main data source for the rules of development with linked diagnosis codes, encounter data, and medications. The methods and results were previously reported and demonstrated a positive predictive value (PPV) and negative predictive value (NPV) at 95% and 96%, respectively.(8) To remain consistent with the original study, International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) were applied for asthma diagnoses. The computable phenotype used was expanded to enable identification of children with asthma (5–17 years of age) (Figure 1) and was updated to include newly FDA-approved asthma medications (Figure 2). The co-morbidities to be excluded in the cohort identification are listed in Table 1. The CAPriCORN informatics group at each site built the computable phenotype in a SQL-based integrative database from the institution's local data warehouse. A control algorithm similar to the original study was applied. The control cohort had to have a minimal amount of information in the EHR equivalent to the data required for asthma cases. A simple random sample of 50 adult and 50 pediatric patients for both the asthma and control algorithms, except at Lurie Children's where pediatric patients only were sampled, was performed retrospectively between January 1, 2012 and November 30, 2014 which provided an observational cohort of 900 patients.

Reference Standard

To validate the results of the computable phenotype, the CAPriCORN Asthma Cohort Committee, a clinical research workgroup comprised of asthma clinical researchers with expertise in care of asthma patients performed blinded chart reviews at their respective institution. Two physicians independently reviewed the same patient in the EHR to identify cases of asthma using their best clinical judgement and following current guidelines. In cases of discordance, a discussion between the two reviewers ensued or a third physician reviewed the chart as arbitration. All investigating physicians remained blinded to the results from the computable phenotype.

Analysis Plan

Inter-observer reliability was measured by the Kappa Coefficient and a difference in agreement between groups was examined by McNemar's test. Discrimination of the computable phenotype was evaluated using the area under the receiver operating characteristic (ROC) curve against the reference standard of retrospective chart review by two physicians. The area under the ROC was estimated in mixed effects logistic regression with random intercepts for the five CAPriCORN institutions for the full cohort as well as stratified by age group (adult and pediatric). The test characteristics including accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated with 95% Confidence Intervals (CI) in classifying asthma as determined by chart review. Test characteristics were also performed by institution and age subgroups. Analysis was performed using SAS Version 9.4 (SAS Institute, Cary, NC).

Results

The SQL-based query in a simple random sample of patients successfully extracted all criteria of the computable phenotype from the institution's clinical data warehouse for chart review at two of the five institutions (N=500). At the remaining three institutions, the SQL-based query successfully extracted data in 96% (N=96), 70.5% (N=141) and 87.5% (N=175) of the patients sampled for chart review. The most common reason for unsuccessful data extraction was due to missing a second meaningful encounter. The final analysis cohort comprised 90.7% (N=816) of the original sampling cohort across the four institutions. No differences were noted between institutions in cases of patients with and without asthma ($p=0.70$) with a range between 43.0% and 54.0% of patients classified with asthma at each institution. In addition, three sites equally sampled pediatric and adult patients (Rush, UI Health, Cook County), one institution sampled 82.5% adults (University of Chicago), and one institution only served pediatric patients (Lurie Children's). There is significant diversity in the patient and hospital characteristics across the sites, details of which are in Supplemental.

In the establishment of the reference standard by retrospective chart review, the interobserver reliability between the physician reviewers was observed with κ -coefficient at 0.95 (95% CI 0.93–0.97) with no differences in agreement between the reviewers ($p=0.65$). Adjudication for discordance between the reviewers was performed in 2.4% (N=20) of cases.

In mixed effects logistic regression, the discrimination between patients with and without asthma by the computable phenotype had an area under the ROC curve at 0.98 (95% Confidence Interval (CI) 0.98–0.99) among all the institutions. The discrimination of the computable phenotype for adult and pediatric patients were similar with area under the ROC curves at 0.97 (95% CI 0.96–0.99) and 0.99 (95% CI 0.98–0.99), respectively (Figures 3a-c). In subgroup analyses by each institution, the discrimination of the computable phenotype had similar area under the ROC curves (data not shown) with the lowest area under the ROC curve at 0.94 (95% CI 0.95–0.97). The accuracy, sensitivity, specificity, NPV, and PPV are displayed in Table 2 for the full cohort and by institution and age subgroups. None of the test characteristics went below 94.2% in the full cohort and 86.0% in the subgroup analyses by institution and age group.

Discussion

In this external validation of the computable phenotype for case identification of patients with asthma, we demonstrate excellent accuracy and precision across multiple health systems that represent diverse payers, multiple EHR systems, and a dense urban population. The tool's performance was not affected by institution and similar test characteristics were shown when evaluated by age group and institution. The computable phenotype can be successfully applied and implemented in EHRs across various health systems for identifying adult and pediatric patients with asthma.

Ease of implementation is an important factor for institutions attempting to use electronic methods for clinical research and practice. Half the institutions in this study applied the SQL code without problems and the other half did not identify a second meaningful case in less than a quarter of cases. The high success rate was likely a reflection in the use of structured data and excluding the need for more complex methods such as natural language processing(10) which makes this tool more applicable at sites with less informatics capabilities and addresses the need for high-throughput clinical phenotyping.(11)

The benefits of this automated method are multifactorial. The elimination of manual abstraction significantly reduces workload without compromising much accuracy or precision.(8) This method is ideal for EHR platform trials and allows screening without the costly resources of traditional clinical trials, a goal of PCORI.(12) In addition, the heterogeneity of health systems in CAPriCORN adds to the external validity of the results. When examining the results by institution, potential biases in institution-specific behaviors around asthma diagnoses were not apparent in subgroup analysis which showed similar test characteristics across institutions.

Previously, others have successfully identified patients with asthma using the combination of ICD-9 diagnoses and medications from the EHR.(13) Even certain phenotypes such as aspirin-exacerbated asthma have been examined with high accuracy and precision.(14) However these studies did not examine adults and pediatrics simultaneously and did not perform external validation of their results. Our study was very similar to the original validation that had results at or greater than 95% for NPV and PPV(8). The authors from the first validation study showed that a less inclusive algorithm that allowed the use of billing

codes with fewer criteria had a lower PPV at 70%.(8) Hence, these validation studies performed from the CAPriCORN institutions for identification of adult and pediatric patients with asthma are the most robust to date.

Several limitations did occur in this study. To remain consistent with the original validation study, only ICD-9 codes were used for diagnoses. Most health systems have incorporated and are using ICD-10 codes so the ICD-9 codes will need to be mapped to ICD-10 codes with validation. Sampling bias may have occurred because only patients with asthma that have had an encounter with the health system on more than one occasion will be detected. The computable phenotype will not capture a single encounter for asthma. In addition, chart review was performed retrospectively rather than prospectively which may introduce additional biases. However, the excellent agreement between reviewers demonstrates physicians can reliably phenotype asthma while being solely dependent on review of EHR data. Patients with overlap chronic obstructive pulmonary disease or patients that also had vocal cord dysfunction would have been excluded in the computable phenotype. Furthermore, medications that were not FDA approved for asthma including anticholinergics or should not be solely used for asthma treatment (e.g. long-acting beta2-agonists) were included in the medication list for asthma inclusion. While these may lead to false negative and false positive cases, the high NPV and PPV in external validation demonstrated these occurrences were exceptionally low. Nevertheless, the computable phenotype would not be applicable for studies attempting to address these subphenotypes and would possibly include providers using non-FDA approved medications.

Conclusion

The near perfect PPV and excellent NPV in this large external validation study establish a useful tool for future research and care. The computable phenotype for asthma is not only applicable within the CAPriCORN Common Data Model but also in PCORI Common Data Model for large-scale comparative-effectiveness trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to thank the following individuals for their assistance with the project: Nicole Twu, Edward Kim, Brenda Louise Giles, and Trevor A. Robison.

This work was supported by Patient Centered Outcomes Research Institute (PCORI) CDRN-1306-04737. Valerie G. Press is supported by a K23 (HL1189151) from the National Heart, Lung, and Blood Institute (NHLBI), Sharmilee M. Nyenhuis is supported by a K23 (HL133370) from NHLBI, and Majid Afshar is supported by a K23 (AA024503).

References

1. CDC Vital Signs, Asthma in the US [Internet]. Center for Diseases Control and Prevention. 5 2011 Available from: <http://www.cdc.gov/VitalSigns/asthma/>.

2. Blackwell DL, Lucas JW, Clarke TC. Summary health statistics for U.S. adults: national health interview survey, 2012. Vital and health statistics Series 10, Data from the National Health Survey. 2014(260):1–161.
3. National Asthma E, Prevention P. Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007. J Allergy Clin Immunol. 2007;120(5 Suppl):S94–138. [PubMed: 17983880]
4. Corley DA, Feigelson HS, Lieu TA, McGlynn EA. Building Data Infrastructure to Evaluate and Improve Quality: PCORnet. J Oncol Pract. 2015;11(3):204–6. [PubMed: 25980016]
5. Hernandez AF, Fleurence RL, Rothman RL. The ADAPTABLE Trial and PCORnet: Shining Light on a New Research Paradigm. Ann Intern Med. 2015;163(8):635–6. [PubMed: 26301537]
6. Kho AN, Hynes DM, Goel S, Solomonides AE, Price R, Hota B, et al. CAPriCORN: Chicago Area Patient-Centered Outcomes Research Network. J Am Med Inform Assoc. 2014;21(4):607–11. [PubMed: 24821736]
7. Shannon JJ, Catrambone CD, Coover L. Targeting improvements in asthma morbidity in Chicago: a 10-year retrospective of community action. Chest. 2007;132(5 Suppl):866S–73S. [PubMed: 17998352]
8. 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. [PubMed: 20351906]
9. Solomonides A, Goel S, Hynes D, Silverstein JC, Hota B, Trick W, et al. Patient-Centered Outcomes Research in Practice: The CAPriCORN Infrastructure. Stud Health Technol Inform. 2015;216:584–8. [PubMed: 26262118]
10. Albright D, Lanfranchi A, Fredriksen A, Styler WF, Warner C, Hwang JD, et al. Towards comprehensive syntactic and semantic annotations of the clinical narrative. J Am Med Inform Assoc. 2013;20(5):922–30. [PubMed: 23355458]
11. Wei WQ, Leibson CL, Ransom JE, Kho AN, Caraballo PJ, Chai HS, et al. Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. J Am Med Inform Assoc. 2012;19(2):219–24. [PubMed: 22249968]
12. Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. JAMA. 2015;313(16):1619–20. [PubMed: 25799162]
13. 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. [PubMed: 9250251]
14. Cahill KN, Johns CB, Cui J, Wickner P, Bates DW, Laidlaw TM, et al. Automated identification of an aspirin-exacerbated respiratory disease cohort. J Allergy Clin Immunol. 2017;139(3):819–25 e6. [PubMed: 27567328]
15. Fact Sheet: The University of Chicago Medicine. Available at: <http://www.uchospitals.edu/about/fact/hospitals-sheet.html>. Accessed 08/30/2017.
16. Executive report. 2015 Community Health Needs Assessment. University of Chicago Medical Center Service Area, Cook County, Illinois Available at: http://www.uchospitals.edu/pdf/uch_047274.pdf. Accessed 8/30/2017.
17. Rush University Medical Center Bondholder Information. Available at: <https://www.rush.edu/about-us/bondholder-information-about-rush>. Accessed 08/30/2017.
18. 2016 Community Benefits Summary. Available at: <https://www.rush.edu/sites/default/files/community-benefits-report-2016.pdf>. Accessed 08/30/2017.
19. Rush University Medical Center Community Health Needs Assessment Report. Available at: <https://www.rush.edu/sites/default/files/rush-chna-august-2016.PDF>. Accessed 8/30/2018.
20. Stoger John H., JR. Hospital of Cook County. About Us. <http://www.cookcountyhhs.org/locations/john-h-stroger-jr-hospital/about-us/>. Accessed 8/30/2017.
21. Illinois Medical District Hospitals. Available at: <http://www.imdc.org/districtpartners/hospitals>. Accessed 8/30/2017. Accessed 8/30/2017.
22. Hospital Profile-CY 2012 John H. Stroger Hospital of Cook County. Available at: <http://app.idph.state.il.us/files/BMI/2012%20Hosp%20Profiles/5272.pdf>. Accessed 8/30/2017.

23. Ann & Robert H. Lurie Children's Hospital of Chicago Facts and Figures. Available at: <https://www.luriechildrens.org/en-us/facts-figures/Pages/index.aspx>. Accessed 8/30/2017.
24. Community Health Needs Assessment 2016 Ann & Robert H. Lurie Children's Hospital of Chicago. Available at: <https://www.luriechildrens.org/en-us/community/community-healthneeds-assessment/Documents/chna-2016.pdf>. Accessed 8/30/2017.
25. Illinois Medical District Hospitals. Available at: <http://www.imdc.org/districtpartners/hospitals>. Accessed 8/30/2017.
26. University of Illinois Community Assessment of needs (UI-CAN) 2016: Toward Health Equity. Available at: https://hospital.uillinois.edu/Documents/about/UI-CAN2016_v2.pdf. Accessed 8/30/2017.

Figure 1.
Computable Asthma Phenotype

Asthma Medications [drug name (trade name)]:	
Inhaled Corticosteroids (ICS): Route: Inhaled. <ul style="list-style-type: none"> • Beclomethasone propionate (QVAR) • Budesonide (Pulmicort Turbuhaler or Flexhaler, Pulmicort Respules) • Flunisolide (Aerobid, Aerobid-M, Aerospan 80) • Fluticasone propionate (Flovent HFA, Flovent Diskus) • Mometasone (Asmanex Twisthaler) • Triamcinolone acetonide (Azmecort) • Ciclesonide (Alvesco) 	Short-acting beta agonists (SABA): Route: Inhaled, nebulized, oral (PO). <ul style="list-style-type: none"> • Levalbuterol (Xopenex HFA) • Albuterol, Albuterol Sulfate (Ventolin, Proventil, ProAir, Accuneb) • Metaproterenol (Alupent) • Pirbuterol (Maxair) • Terbutaline (Brethine)
ICS-LABA Combinations: Route: Inhaled. <ul style="list-style-type: none"> • Budesonide-Formoterol (Symbicort) • Fluticasone-salmeterol (Advair HFA, Advair Diskus) • Mometasone-formoterol (Dulera) • Fluticasone furoate + Vilanterol (Breo Ellipta) 	Long-acting beta agonists (LABA): Route: Inhaled, nebulized. <ul style="list-style-type: none"> • Salmeterol (Serevent) • Formoterol (Foradil) • Arformoterol (Brovana) • Indacaterol (Arcapta)
Leukotriene Antagonists (LTA): Route: Oral (PO) <ul style="list-style-type: none"> • Montelukast (Singulair) • Zafirlukast (Accolate) • Zileuton (Zyflo, Zyflo CR) 	Methylxanthines: Route: oral (PO), intravenous (IV). <ul style="list-style-type: none"> • Aminophylline • Theophylline (Slo-Bid, Theo, Theodur, Uniphyll)
Mast-cell stabilizers: Route: Inhaled. <ul style="list-style-type: none"> • Cromolyn sodium, cromoglycate (Intal) • Nedocromil (Tilade) 	Anticholinergic agents: Route: Inhaled. <ul style="list-style-type: none"> • Ipratropium bromide (Atrovent HFA) • Tiotropium bromide (Spiriva HandiHaler, Spiriva Respimat) • Acridinium bromide (Turdoza Pressair) • Umeclidinium bromide (Incruse Ellipta)
Biologicals: Route: Subcutaneous (SC). <ul style="list-style-type: none"> • Omalizumab (Xolair). 	Combination bronchodilators: Route: Inhaled, nebulized. <ul style="list-style-type: none"> • Albuterol + ipratropium (Combivent MDI or Respimat, Duoneb) • Umeclidinium + vilanterol (Anoro Ellipta)

Figure 2.
Medications that were used to identify asthma cases

ROC Curve results for the full cohort in Figure 3A (N=816) as well as by Pediatric group (N=409) in Figure 3B and Adult group (N=407) in Figure 3C demonstrate similar values.

Figure 3.

a-c Area under the Receiver Operating Characteristic Curve for the Computable Asthma Phenotype across CAPriCORN institutions

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Exclusionary ICD9 Codes

ICD-9 Codes	Diagnosis
Cancers & HIV	
160.xx to 165.xx	Cancers of airways and thorax
197.0 to 197.3	Secondary malignant neoplasm of respiratory system
200.xx to 208.xx	Cancers of lymphatic and hematopoietic systems
212.3	Benign neoplasm of the bronchus and lung
042	Human immunodeficiency virus [HIV] disease
079.5×	Retrovirus infections
V08 and 795.71	Asymptomatic & serologic HIV infection.
Airway & Lung Diseases	
ICD-9 Codes	Diagnosis
273.4	Alpha 1- antitrypsin deficiency
277.xx, V77.6, V83.81	Cystic Fibrosis and other metabolic disorders
415.x	Acute Pulmonary Heart Disease (cor pulmonale, embolism)
416.x	Chronic Pulmonary Heart Disease
478.3×	Paralysis of vocal cords or larynx
478.5	Other diseases of vocal cords (Vocal Cord Dysfunction)
491.xx	Chronic obstructive pulmonary disease [COPD]
492.x	Emphysema
494.xx	Bronchiectasis
495.x	Extrinsic allergic alveolitis
496	Chronic obstructive pulmonary disease
500.x – 508.x	Pneumoconioses and post-radiation lung disease
511.81	Malignant pleural effusion
515	Pulmonary fibrosis
516.x	Other alveolar and parietoalveolar pneumonopathy
517.x	Lung involvement in conditions classified elsewhere
518.6	Allergic bronchopulmonary aspergillosis
519.4	Disorders of diaphragm
748.2 to 748.9	Congenital abnormalities of lower airways and lungs
769.xx	Respiratory Distress Syndrome
V81.3	Chronic bronchitis and emphysema
770.7	Bronchopulmonary Dysplasia (BPD)
765.21 to 765.27	Prematurity (Born < 35 weeks). 765.28 and 765.29 are allowed.
934.8	Aspiration syndromes
Diseases that are often treated with systemic steroids	
ICD-9 Codes	Diagnosis
238.77	Post-transplant lymphoproliferative disorder (PTLD)

ICD-9 Codes	Diagnosis
710.x	Systemic lupus erythematosus, scleroderma, dermatomyositis
714.x	Rheumatoid Arthritis and other inflammatory polyarthropathies
996.8x	Complications of transplanted organ
E878.0	Surgical operation with transplant of whole organ
V42.x	Organ or tissue replaced by transplant
V49.83	Awaiting organ transplant status
V58.44	Aftercare following organ transplant
Neuromuscular diseases affecting respiratory system	
ICD-9 Codes	Diagnosis
343.x	Cerebral palsy
335.x	Spinal muscular Atrophy and other motor neuron diseases
359.x	Muscular dystrophies and other myopathies
Other Diseases	
ICD-9 Codes	Diagnosis
428.x	Heart failure

Table 2.

Test Characteristics of Computable Asthma Phenotype across CAPriCORN institutions

Characteristic	% Accuracy (95% CI)	% Sensitivity (95% CI)	% Specificity (95% CI)	% NPV (95% CI)	% PPV (95% CI)
Full Cohort (N=816)	96.7 (95.5–97.9)	98.7 (97.5–99.8)	95.0 (93.0–97.0)	98.2 (97.0–99.4)	94.2 (92.0–96.4)
Institution					
Institution 1 (N=200)	97.5 (95.3–99.7)	99.0 (97.0–100)	96.0 (92.4–99.9)	99.0 (97.1–100)	96.0 (92.2–99.8)
Institution 2 (N=141)	97.2 (94.4–99.9)	100 (100–100)	95.0 (90.2–99.8)	100 (100–100)	93.9 (88.0–99.7)
Institution 3 (N=175)	100 (100–100)	100 (100–100)	100 (100–100)	100 (100–100)	100 (100–100)
Institution 4 (N=200)	93.0 (89.5–96.5)	100 (100–100)	87.7 (80.3–93.1)	100 (100–100)	86.0 (79.2–92.8)
Institution 5 [*] (N=100)	96.0 (92.2–99.8)	92.6 (85.6–99.6)	100 (100–100)	92.0 (84.5–99.5)	100 (100–100)
Pediatric Patients (N=409)	97.8 (96.4–99.2)	97.6 (95.5–99.7)	98.0 (96.1–99.9)	97.6 (95.6–99.7)	97.7 (95.6–99.7)
Adult Patients (N=407)	95.6 (93.6–97.6)	100 (100–100)	92.4 (89.0–95.8)	98.7 (97.3–100)	90.7 (86.9–94.6)

^{*} Pediatric only (ages <18 years); Adult: ages 18 years and older; CI: Confidence

Interval: NPV: negative predictive value; PPV: positive predictive value.