Accuracy of Algorithms to Identify Pulmonary Arterial Hypertension in Administrative Data

A Systematic Review

Kari R. Gillmeyer, MD; Ming-Ming Lee, MD; Alissa P. Link, MS (LIS); Elizabeth S. Klings, MD; Seppo T. Rinne, MD, PhD; and Renda Soylemez Wiener, MD, MPH

BACKGROUND: The diagnosis of pulmonary arterial hypertension (PAH) is challenging, and there is significant overlap with the more heterogeneous diagnosis of pulmonary hypertension (PH). Clinical and research efforts that rely on administrative data are limited by current coding systems that do not adequately reflect the clinical classification scheme. The aim of this systematic review is to investigate current algorithms to detect PAH using administrative data and to appraise the diagnostic accuracy of these algorithms against a reference standard.

METHODS: We conducted comprehensive searches of Medline, Embase, and Web of Science from their inception. We included English-language articles that applied an algorithm to an administrative or electronic health record database to identify PAH in adults.

RESULTS: Of 2,669 unique citations identified, 32 studies met all inclusion criteria. Only four of these studies validated their algorithm against a reference standard. Algorithms varied widely, ranging from single *International Classification of Diseases* (ICD) codes to combinations of visit, procedure, and pharmacy codes. ICD codes alone performed poorly, with positive predictive values ranging from 3.3% to 66.7%. The addition of PAH-specific therapy and diagnostic procedures to the algorithm improved the diagnostic accuracy.

CONCLUSIONS: Algorithms to identify PAH in administrative databases vary widely, and few are validated. The sole use of ICD codes performs poorly, potentially leading to biased results. ICD codes should be revised to better discriminate between PH groups, and universally accepted algorithms need to be developed and validated to capture PAH in administrative data, better informing research and clinical efforts. CHEST 2019; 155(4):680-688

KEY WORDS: administrative data; algorithm; *International Classification of Diseases*; pulmonary hypertension

FOR EDITORIAL COMMENT, SEE PAGE 653

ABBREVIATIONS: ECHO = echocardiogram; EHR = electronic health record; ICD = *International Classification of Diseases*; ICD-9 = *International Classification of Diseases, Ninth Revision*; PAH = pulmonary arterial hypertension; PH = pulmonary hypertension; PPV = positive predictive value; RHC = right heart catheterization; WHO = World Health Organization

AFFILIATIONS: From The Pulmonary Center (Drs Gillmeyer, Lee, Klings, Rinne, and Wiener) and Alumni Medical Library (Ms Link), Boston University School of Medicine, Boston, MA; and the Center for Healthcare Organization & Implementation Research (Drs Rinne and Wiener), Edith Nourse Rogers Veterans Hospital, Bedford, MA.

FUNDING/SUPPORT: K. R. G. was funded through a National Institutes of Health Institutional Training Grant [Grant T32 HL007035]. S. T. R. was funded through a Parker B. Francis Fellowship and a VISA 1 Career Development Award. This study was also supported in part by resources from the Edith Nourse Rogers Memorial VA Hospital.

CORRESPONDENCE TO: Kari R. Gillmeyer, MD, The Pulmonary Center, 72 E Concord St, R304, Boston, MA 02118; e-mail: Kari.Gillmeyer@bmc.org

Copyright © 2018 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2018.11.004>

Pulmonary arterial hypertension (PAH), a relatively rare subgroup of pulmonary hypertension (PH), often presents with nonspecific signs and symptoms, requiring a high index of suspicion to make the diagnosis. Adding to the diagnostic challenge, the clinical presentation of PAH often overlaps with other more common forms of PH, such as those related to left heart disease or chronic lung disease. This necessitates an extensive and invasive workup, with high rates of missed and incorrect diagnoses.^{1,2} Given the rarity and underrecognition of the disease, registries of patients with PAH have been instrumental in illuminating much of what we know about PAH today, including demographic characteristics, natural history, risk factors for progression, and response to treatment.³⁻⁶ By nature of their design, however, registries are subject to selection bias and may not represent PAH care in the community outside of referral centers.

Administrative health data such as insurance databases and electronic health records (EHRs) offer an opportunity to conduct population-level studies with decreased risk of selection or recall bias and improved generalizability. As such, administrative data are increasingly being used to examine practice patterns, disease burden, health-care utilization, and quality of care for a wide variety of diseases, including PH.⁷⁻¹⁰ However, PH codes in the *International Classification of*

Diseases, Ninth Revision (ICD-9) and *International Classification of Diseases, 10th Revision* categorize patients as having either primary or secondary PH, and do not accurately reflect the current World Health Organization (WHO) clinical classification scheme, which sorts PH into individual groups that share similar underlying etiologies, such as left heart disease, chronic lung disease, or chronic thromboembolic disease.¹¹ Potential misclassification of patients into the wrong PH group may bias results obtained from administrative data and bring the validity of those results into question.

Given these potential limitations, it is important to understand how well algorithms using administrative health data perform in identifying patients with PH and distinguishing patients with PAH from those with other types of PH. Although researchers have proposed various algorithms to identify patients with PAH using administrative data, little is known about how well these algorithms perform, how they compare with one another, or what factors researchers might consider in selecting which algorithm to use to capture patients with PAH in administrative databases. To address these questions, we performed a systematic search of the literature, seeking to identify and characterize existing algorithms to detect patients with PAH using administrative data, and to critically appraise the diagnostic accuracy of these algorithms.

Materials and Methods

We conducted comprehensive searches of Embase, Web of Science, and Medline via PubMed from their inception through April 2018, combining both controlled vocabulary and key words, including the core phrases “administrative database” and “pulmonary hypertension” (e-Table 1). We manually searched the reference lists of included articles to identify additional pertinent articles. We also used the Cited-By tool in PubMed to find relevant articles. This search strategy was based on previously published strategies^{12,13} and was further developed in consultation with an experienced medical librarian (A. P. L.). We conducted this literature review according to the standards of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.¹⁴ Because this was a retrospective review of previously published literature without direct involvement of human participants, ethics approval was not required.

Selection of Studies for Review

Two independent reviewers (K. R. G. and M.-M. L.) screened titles and abstracts for eligibility using screening software¹⁵ and resolved discrepancies by discussion. Reviewers were blinded to each other's decision to include or exclude an article. Studies were eligible for full-text review if they (1) addressed the target population of adults aged ≥ 18 years with PAH, (2) used an administrative or EHR database as a data source, (3) clearly reported a case-finding algorithm, and (4) were full-length articles published in English. We decided a priori to include both studies with validated and

unvalidated algorithms given the expected small number of studies on the topic. We quantified the agreement between reviewers on whether or not to include an article using Cohen κ statistic.¹⁶

Data Extraction and Quality Assessment

From the selected studies, we abstracted data using a standardized abstraction form that was piloted on a sample of articles (e-Table 2). For each eligible study, we obtained the following information: study information (title, first author, publication year, and journal), details of the algorithm tested (eg, outpatient or hospitalization code, procedure codes, prescriptions), target population from which the health-care data were collected, outcomes studied, type of health-care database used (eg, hospitalization discharge data, EHR, insurance database), diagnostic accuracy of the algorithm (sensitivity, specificity, positive predictive values [PPVs], and negative predictive values), and reference standard used to validate the algorithm. When possible, measures of diagnostic accuracy that were not reported in the article were calculated from available data.

For articles that used validated algorithms, two reviewers (K. R. G. and M.-M. L.) assessed study quality using a standardized 40-point checklist, which was developed by Benchimol et al¹⁷ and based on the criteria published by the Standards for Reporting of Diagnostic Accuracy initiative.¹⁸ This checklist was designed to assess the quality of the methods of validation studies of administrative data. Disagreements were resolved by discussion. We descriptively reported the presence of any potential biases.

Results

Search Results

Our initial search strategy identified 2,669 unique citations. No additional articles were identified through manual searching or through the Cited-By tool. After title and abstract screening, we selected 93 articles for full-text review, of which 32 met all inclusion criteria (Fig 1). Primary reasons for exclusion of articles were wrong target population ($n = 1,381$, 52%) and wrong data source ($n = 1,191$, 45%). Agreement between reviewers on inclusion vs. exclusion of abstracts was high with a Cohen κ statistic of 0.80.

Study Characteristics

Details of the included studies are summarized in Table 1¹⁹⁻²² (validated algorithms, $n = 4$) and e-Table 3 (unvalidated algorithms, $n = 28$). Of the

32 included articles, most ($n = 26$) were from the United States, with the remaining studies from Canada ($n = 2$), Europe ($n = 3$), and Taiwan ($n = 1$). Nearly all studies used *International Classification of Diseases* (ICD) diagnosis codes linked to inpatient or outpatient visits, pharmacy claims, procedure claims, or a combination of these to identify PAH, hereafter termed claims-based algorithms. Only one study used physician-entered diagnoses from EHRs,²² hereafter termed EHR-based algorithms. The included studies examined a variety of outcomes, including prevalence, mortality rates, length of hospital stay, and health-care costs associated with PAH. Most studies examined PAH in the general population; however, some identified PAH among specific cohorts, including patients with psoriasis,²³ systemic sclerosis,²⁴ hepatitis C and multiple sclerosis,²⁵ HIV,²⁶ and connective tissue disease.²⁷

Summary of Unvalidated Algorithms

Twenty-eight of the included articles did not validate the applied algorithm (e-Table 3). Among these, the most common case definition ($n = 11$, 39%) identified patients with PAH using only visit-linked ICD diagnosis codes. Some studies required both a visit-linked ICD diagnosis code and a pharmacy claim for a PAH-specific medication,²⁷⁻³³ whereas others required both a visit-linked ICD diagnosis code and a procedure claim for right heart catheterization (RHC) or echocardiogram (ECHO).³⁴⁻³⁶ Several studies attempted to sort patients into distinct WHO groups using codes for associated diagnoses.³⁴⁻⁴⁰ For example, Stein et al³⁹ identified patients with PH using visit-linked ICD diagnosis codes and then further refined the cohort to only PAH by excluding all patients with codes for diagnoses associated with WHO groups II through V PH. Most studies ($n = 15$, 54%) used only ICD-9 code 416.0 (defined as primary PH) to select patients with PAH, whereas others included both ICD-9 codes 416.0 and 416.8 (defined as other chronic pulmonary heart disease) in the case definition.

Summary of Validated Algorithms

Four of the included studies validated their case definitions, all using medical chart review by a PH specialist as the reference standard (Tables 1, 2).¹⁹⁻²¹ For two of these studies, chart review was based on ICD diagnosis codes and procedure codes, rather than granular data, such as RHC results.^{19,20} The case

Figure 1 – Summary of study selection for systematic review. PAH = pulmonary arterial hypertension.

TABLE 1] Characteristics of Studies Using Validated Algorithms

Author, Year, Country	Study Population and Time Period	Study Design	Sample Size	Outcomes Studied	STARD Score ^a
Chang et al, ¹⁹ 2016, Taiwan	Patients with incident PH identified from the single payer Longitudinal Health Insurance Database, 1999-2011	Retrospective cohort	1,092	Mortality, hospitalization within 30 d after PH diagnosis	17
Fox et al, ²⁰ 2014, United Kingdom	Patients with incident PAH identified from a primary care database (Clinical Practice Research Datalink) and a hospitalization database (Hospital Episodes Statistics), 1988-2011	Nested case-control	195	Incident PAH in those who used antidepressants (cases) vs those who did not (control subjects)	21
Link et al, ²¹ 2011, United States	Patients with PH identified from three sources: (1) US National Center for Health Statistics database (1979-2006) for mortality data, (2) National Inpatient Sample (1993-2007) for hospital discharge data, and (3) University of Texas Southwestern Hospital for validation of ICD codes (1999 and 2002)	Retrospective cohort	Not reported	PH-associated mortality, rates of PH at discharge, performance of ICD codes	14
Papani et al, ²² 2018, United States	Patients age ≥ 18 y with primary or secondary PH identified from EHRs from the University of Texas Medical Branch and the University of Virginia (2012-2015)	Retrospective cohort	683	Validation of algorithms identifying PAH	29

EHR = electronic health record; ICD = *International Classification of Diseases*; PAH = pulmonary arterial hypertension; PH = pulmonary hypertension; STARD = Standards for Reporting of Diagnostic Accuracy.

^aMethod of quality assessment of validation studies using administrative data (range, 0-40 possible points), with higher scores indicating higher quality.

definitions for three of these four studies used a single visit-linked ICD diagnosis code.¹⁹⁻²¹ In the fourth study, Papani et al²² developed a series of eight algorithms including both claims-based and EHR-based algorithms. Claims-based algorithms started with visit-linked ICD diagnosis codes and were sequentially built to require an increasing number of classes of PAH-specific therapy from pharmacy claims, whereas EHR-based algorithms paired visit-linked ICD diagnosis codes with EHR encounter diagnoses and successively required claims for RHC or ECHO procedures, PAH-specific therapy, or both. EHR encounter diagnoses were entered by the physician at the time of the visit and were determined to be suggestive of PAH. After development, the algorithms were externally validated in a separate cohort.

Performance Characteristics of Validated Algorithms

Performance characteristics of the validated algorithms are summarized in Table 2. The most commonly reported statistics were PPV (n = 2), sensitivity (n = 2), and specificity (n = 2). In two studies, PPV could be calculated from the provided information.^{19,21} In two studies that solely used visit-linked ICD diagnosis codes in the case definition, the PPV was low, with values of 3.3% and 33.3%, respectively.^{20,21} A similar case definition used by Chang et al¹⁹ performed better, with a PPV of 66.7% and specificity of 96.7%. In the multiple algorithm study by Papani et al,²² algorithms using visit-linked ICD diagnosis codes only performed poorly, with a PPV of 9.3% in the development cohort and 15.8% in the validation cohort. For claims-based

TABLE 2] Performance of Validated Algorithms

Author	Algorithm	Reference Standard	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	C Statistic
Chang et al ¹⁹	Discharge diagnosis with ICD-9 code 416.0	Medical chart review ^a	63.2	96.7	66.7	...	...
Fox et al ²⁰	CPRD: read codes: G41y000, G410.00, 7Q01300, 7Q01200, 7Q01100, 7Q01000, G42.11, G41z.00, G41.00 HES: ICD-10 code I27.0	Medical chart review ^b	...	...	3.3	...	...
Link et al ²¹	Discharge diagnosis with ICD-9 code 416.0 1999 2002	Medical chart review ^c	...	...	6.7 33.3	...	...
Papani et al ²²	Development cohort: claims-based algorithms:	Medical chart review ^d					
	ICD-9 codes 416.0 and 416.8		...	...	9.3	...	...
	ICD codes plus at least one PAH-specific medication		67.4	86.9	34.7	96.3	0.8
	ICD codes plus two or more classes of PAH-specific medications		28.3	98.6	66.9	93.0	0.7
	Development cohort: EHR-based algorithms:						
	ICD codes plus EHR diagnosis		76.9	77.1	25.7	97.0	0.7
	ICD codes plus EHR diagnosis plus ECHO		76.9	78.2	26.6	97.0	0.7
	ICD codes plus EHR diagnosis plus ECHO plus RHC		76.9	91.4	48.0	97.5	0.9
	ICD codes plus EHR diagnosis plus ECHO plus RHC plus medication		67.4	96.9	69.4	96.7	0.9
	ICD codes plus EHR diagnosis plus medication		67.4	96.5	66.2	96.6	0.9
	Validation cohort: claims-based algorithms						
	ICD-9 codes 416.0 and 416.8		...	...	15.8	...	...
	ICD codes plus at least one PAH-specific medication		64.3	81.9	40.0	92.4	...
	ICD codes plus two or more classes of PAH-specific medications		42.7	94.0	57.1	89.7	...
	Validation cohort: EHR-based algorithms						
	ICD codes plus EHR diagnosis		25.0	85.9	25.0	85.9	...

(Continued)

TABLE 2] (Continued)

Author	Algorithm	Reference Standard	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	C Statistic
	ICD codes plus EHR diagnosis plus ECHO		25.0	91.3	35.0	86.6	...
	ICD codes plus EHR diagnosis plus ECHO plus RHC		25.0	96.6	58.3	87.3	...
	ICD codes plus EHR diagnosis plus ECHO plus RHC plus medication		17.9	98.0	62.5	86.4	...
	ICD codes plus EHR diagnosis plus medication		17.9	96.0	45.5	86.1	...

CPRD = Clinical Practice Research Datalink; ECHO = echocardiogram; HES = Hospital Episodes Statistics; ICD-9 = *International Classification of Diseases, Ninth Revision*; ICD-10 = *International Classification of Diseases, 10th Revision*; NPV = negative predictive value; PPV = positive predictive value; RHC = right heart catheterization. See Table 1 legend for expansion of other abbreviations.

^aReview of patient discharge records including prior workup for PH, based on ICD diagnosis codes and procedure codes.

^bReview of the electronic health record including ICD diagnosis codes for PAH-associated diseases, ICD diagnosis codes for cardiac and respiratory diseases associated with PH but not PAH, procedure codes for RHCs, and prescriptions for pulmonary vasodilators.

^cReview of patient discharge records and prior workup for PH including progress notes, ECHO and RHC results, and imaging results.

^dReview of the electronic health record including age, sex, race, comorbidities, ECHO and RHC results, prescribed PAH-specific therapies, and CT angiography or ventilation-perfusion scan results.

algorithms, pairing ICD codes with prescriptions for two or more classes of PAH-specific therapy resulted in a better performance, with a PPV of 66.9% and specificity of 98.6%. The EHR-based algorithm combining visit-linked ICD codes, EHR encounter diagnoses, the performance of both an ECHO and an RHC, and a prescription for a PAH-specific medication achieved the highest PPV (69.4%) and specificity (96.9%), but resulted in a lower sensitivity (67.4%). In the validation cohort, the algorithms had similar specificities but lower sensitivities compared with the development cohort.

Quality Assessment of Validated Studies

The quality of the studies using validated algorithms was variable, with Standards for Reporting of Diagnostic Accuracy scores ranging from 14 to 29 out of 40 maximum points (Table 1, full details in e-Table 4). Only one study revalidated the algorithm in a separate cohort.²² Experts in the field of PH performed the chart review and reported the reference standard in all studies; however, the findings were confirmed by a second reviewer in only two studies.^{20,22}

Discussion

To our knowledge, this is the first systematic review to summarize and critically assess existing algorithms identifying PAH in administrative databases. We

found 32 studies that applied algorithms to specifically select for PAH in administrative data. The case definitions in these studies varied widely, from single ICD diagnosis codes to combinations of visit, procedure, and pharmacy codes. Several studies used nuanced algorithms in an attempt to reflect the current WHO clinical classification scheme of PH and, specifically, PAH, either by including WHO group I-associated diagnoses or by excluding WHO groups II through V diagnoses³⁴⁻⁴⁰; however, none of these algorithms were validated.

Only four of the included studies validated their algorithms against a reference standard, and in two of these studies that reference standard was simply review of ICD diagnosis and procedure codes by a PH specialist.¹⁹⁻²² In general, the sole use of ICD diagnostic codes to select for PAH in administrative data performed poorly with low PPVs. Only one study developed and externally validated algorithms using combinations of visit, procedure and pharmacy codes.²² The authors found that combining ICD codes with EHR encounter diagnoses, procedures, and PAH-specific therapy improved the specificity of the algorithm, but at a cost to sensitivity. Although this study adds valuable insight into this field, it must be noted that many administrative databases only capture billing codes and do not have access to physician-entered EHR diagnoses. Therefore, these algorithms may not be widely applicable to all databases. Moreover, investigators seeking to

examine health-care utilization outcomes, such as procedures performed or treatment provided for patients with PAH, could not use these algorithms because patients who did not receive procedures or PAH-specific therapy would be excluded.

ICD codes alone do not adequately differentiate PAH from other groups of PH and are incompatible with the current WHO clinical classification system, a point highlighted by the low diagnostic accuracy of ICD codes found in this review. The instability of these codes to identify PAH has been further illustrated by Link et al,²¹ who showed that sudden shifts in PAH-related mortality and hospital discharges in three different databases were likely simply because of coding changes. There is a need for revision of ICD diagnosis codes to allow discrimination between different WHO PH groups to improve the ability of researchers and administrators to assess the epidemiology and quality of care for patients with PAH.

As administrative databases expand and EHRs become the rule rather than the exception, the wealth of data extracted from these sources has also flourished. Indeed, research publications from administrative data have rapidly increased in recent years.⁴¹ With this abundance of information comes a responsibility to ensure that we have the tools to appropriately interpret this data prior to applying it for clinical and research purposes. Developing these tools, including validated algorithms, is vital to avoid bias, and an internal consortium of researchers has identified this as a priority in health services research.⁴²

Given the complexity of the diagnosis of PAH even at a patient level, nuanced algorithms reflecting the clinical classification scheme and the natural workup of a patient with PAH must be developed and validated. It may be necessary to build many distinct algorithms to answer diverse research questions. For example, including pharmacy claims for PAH-specific

medications in the case definition necessarily prevents analyzing rates of treatment as an outcome. The performance characteristics of these algorithms will also need to be balanced. For surveillance purposes, for example, algorithms should maximize sensitivity without unacceptably compromising specificity. In cases where diagnostic certainty is necessary, however, specificity will need to be prioritized. It will also be important to validate these algorithms in diverse external cohorts to assure their appropriate application in different administrative databases and populations.

A limitation of this review is the restriction of the search strategy to include only articles published in English in peer-reviewed journals, potentially overlooking some studies. To enhance the breadth of the search, however, we used three independent reference databases and manually examined the references of key articles. It is notable that no additional articles were identified through manual searching or through the Cited-By tool, suggesting that our literature search was comprehensive. Additionally, the PAH experts performing medical chart review in two of the four studies that validated their algorithms did not have comprehensive information required to discriminate PAH from other groups of PH, such as results from RHCs or radiology images. Therefore, the reference used in some of these studies may not have been a true gold standard.

Conclusions

The wide-ranging definitions and low diagnostic accuracy of ICD codes for PAH identified by our review highlight the need for universally agreed on, externally validated case definitions to detect PAH in health-care databases. Once created and validated, these algorithms can become vital instruments to increase our understanding of PAH on a population level.

Acknowledgments

Author contributions: K. R. G. had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. K. R. G., S. T. R., R. S. W., and A. P. L. contributed to study concept and design. K. R. G. and M.-M. L. contributed to acquisition of data. K. R. G., M.-M. L., E. S. K., S. T. R., and R. S. W. contributed to analysis and interpretation of data. K. R. G., S. T. R., and R. S. W. drafted the manuscript. All authors contributed to critical revision of the manuscript for important intellectual content.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* the following: E. S. K. receives research grant support from Actelion, Incyte, Bayer, Reata, and Arena; and is a member of the acute chest syndrome adjudication committee for the phase III study of rivipansel for treatment of vaso-occlusive pain in sickle cell disease (Pfizer). None declared (K. R. G., M.-M. L., A. P. L., S. T. R., R. S. W.).

Role of sponsors: The sponsors had no role in the design of the study, the acquisition and analysis of the data, or the drafting of the manuscript.

Other contributions: The views expressed in this article do not necessarily represent the views of the Department of Veterans Affairs or the US Government.

Additional information: The e-Tables can be found in the Supplemental Materials section of the online article.

References

- Deaño RC, Glassner-Kolmin C, Rubenfire M, et al. Referral of patients with pulmonary hypertension diagnoses to tertiary pulmonary hypertension centers: the multicenter RePherral study. *JAMA Intern Med.* 2013;173(10):887-893.
- Maron BA, Choudhary G, Khan UA, et al. Clinical profile and underdiagnosis of pulmonary hypertension in US veteran patients. *Circ Heart Fail.* 2013;6(5):906-912.
- Farber HW, Miller DP, Poms AD, et al. Five-year outcomes of patients enrolled in the REVEAL Registry. *Chest.* 2015;148(4):1043-1054.
- Benza RL, Gomberg-Maitland M, Miller DP, et al. The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension. *Chest.* 2012;141(2):354-362.
- Ling Y, Johnson MK, Kiely DG, et al. Changing demographics, epidemiology, and survival of incident pulmonary arterial hypertension: results from the pulmonary hypertension registry of the United Kingdom and Ireland. *Am J Respir Crit Care Med.* 2012;186(8):790-796.
- Bergot E, Sitbon O, Cottin V, et al. Current epoprostenol use in patients with severe idiopathic, heritable or anorexigen-associated pulmonary arterial hypertension: data from the French pulmonary hypertension registry. *Int J Cardiol.* 2014;172(3):561-567.
- Ford ES. Hospital discharges, readmissions, and ED visits for COPD or bronchiectasis among US adults: findings from the nationwide inpatient sample 2001-2012 and Nationwide Emergency Department Sample 2006-2011. *Chest.* 2015;147(4):989-998.
- Roth GA, Johnson CO, Abate KH, et al. The burden of cardiovascular diseases among US states, 1990-2016. *JAMA Cardiol.* 2018;3(5):375-389.
- Barnett ML, Olenski AR, Jena AB. Opioid-prescribing patterns of emergency physicians and risk of long-term use. *N Engl J Med.* 2017;376(7):663-673.
- Trivedi AN, Nsa W, Hausmann LR, et al. Quality and equity of care in U.S. hospitals. *N Engl J Med.* 2014;371(24):2298-2308.
- Simonneau G, Gatzoulis MA, Adatia I, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol.* 2013;62(25 suppl):D34-D41.
- Carnahan RM, Moores KG. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned. *Pharmacoepidemiol Drug Saf.* 2012;21(suppl 1):82-89.
- McPheeters ML, Sathe NA, Jerome RN, Carnahan RM. Methods for systematic reviews of administrative database studies capturing health outcomes of interest. *Vaccine.* 2013;31(suppl 10):K2-K6.
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535.
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan: a web and mobile app for systematic reviews. *Syst Rev.* 2016;5(1):210.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
- Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttman A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol.* 2011;64(8):821-829.
- Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ.* 2003;326(7379):41-44.
- Chang WT, Weng SF, Hsu CH, et al. Prognostic factors in patients with pulmonary hypertension—a nationwide cohort study. *J Am Heart Assoc.* 2016;5(9):11.
- Fox BD, Azoulay L, Dell'Aniello S, et al. The use of antidepressants and the risk of idiopathic pulmonary arterial hypertension. *Can J Cardiol.* 2014;30(12):1633-1639.
- Link J, Glazer C, Torres F, Chin K. International Classification of Diseases coding changes lead to profound declines in reported idiopathic pulmonary arterial hypertension mortality and hospitalizations implications for database studies. *Chest.* 2011;139(3):497-504.
- Papani R, Sharma G, Agarwal A, et al. Validation of claims-based algorithms for pulmonary arterial hypertension. *Pulm Circ.* 2018;8(2):8.
- Choi YM, Famenini S, Wu JJ. Incidence of pulmonary arterial hypertension in patients with psoriasis: a retrospective cohort study. *Perm J.* 2017;21.
- Fischer A, Kong AM, Swigris JJ, Cole AL, Raimundo K. All-cause healthcare costs and mortality in patients with systemic sclerosis with lung involvement. *J Rheumatol.* 2018;45(2):235-241.
- Papani R, Duarte AG, Lin YL, Kuo YF, Sharma G. Pulmonary arterial hypertension associated with interferon therapy: a population-based study. *Multidiscip Respir Med.* 2017;12:7.
- Henriques-Forsythe M, Annangi S, Farber HW. Prevalence and hospital discharge status of human immunodeficiency virus-associated pulmonary arterial hypertension in the United States. *Pulm Circ.* 2015;5(3):506-512.
- Yang X, Sanders KN, Mardekian J, Mychaskiw MA, Thomas J. Associations between sildenafil use and changes in days of hospitalization in a population with pulmonary arterial hypertension associated with connective tissue disease. *Clin Ther.* 2015;37(5):1055-1063.
- Berger A, Edelsberg J, Teal S, Mychaskiw MA, Oster G. Changes in healthcare utilization and costs associated with sildenafil therapy for pulmonary arterial hypertension: a retrospective cohort study. *BMC Pulm Med.* 2012;12:8.
- Burger CD, Ozbay AB, Lazarus HM, et al. Treatment patterns and associated health care costs before and after treatment initiation among pulmonary arterial hypertension patients in the United States. *J Manag Care Spec Pharm.* 2018;24(8):834-842.
- Burger CD, Pruett JA, Lickert CA, Berger A, Murphy B, Drake W. Prostacyclin use among patients with pulmonary arterial hypertension in the United States: a retrospective analysis of a large health care claims database. *J Manag Care Spec Pharm.* 2018;24(3):291-302.
- Burke JP, Hunsche E, Regulier E, Nagao M, Buzinec P, Drake W. Characterizing pulmonary hypertension-related hospitalization costs among Medicare advantage or commercially insured patients with pulmonary arterial hypertension: a retrospective database study. *Am J Manage Care.* 2015;21(3):S47-S58.
- Copher R, Cerulli A, Watkins A, Laura Monsalvo M. Treatment patterns and healthcare system burden of managed care patients with suspected pulmonary arterial

- hypertension in the United States. *J Med Econ.* 2012;15(5):947-955.
33. Dufour R, Pruett J, Hu N, et al. Healthcare resource utilization and costs for patients with pulmonary arterial hypertension: real-world documentation of functional class. *J Med Econ.* 2017;20(11):1178-1186.
 34. Kirson NY, Birnbaum HG, Ivanova JL, Waldman T, Joish V, Williamson T. Excess costs associated with patients with pulmonary arterial hypertension in a US privately insured population. *Appl Health Econ Health Policy.* 2011;9(5): 293-303.
 35. Kirson NY, Birnbaum HG, Ivanova JL, Waldman T, Joish V, Williamson T. Prevalence of pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension in the United States. *Curr Med Res Opin.* 2011;27(9): 1763-1768.
 36. Said Q, Martin BC, Joish VN, Kreilick C, Mathai SC. The cost to managed care of managing pulmonary hypertension. *J Med Econ.* 2012;15(3): 500-508.
 37. Carrington M, Murphy NF, Strange G, Peacock A, McMurray JJ, Stewart S. Prognostic impact of pulmonary arterial hypertension: a population-based analysis. *Int J Cardiol.* 2008;124(2): 183-187.
 38. Peacock AJ, Murphy NF, McMurray JJ, Caballero L, Stewart S. An epidemiological study of pulmonary arterial hypertension. *Eur Respir J.* 2007;30(1):104-109.
 39. Stein PD, Matta F, Hughes PG. Scope of problem of pulmonary arterial hypertension. *Am J Med.* 2015;128(8): 844-851.
 40. Wijeratne DT, Lajkosz K, Brogly SB, et al. Increasing incidence and prevalence of world health organization groups 1 to 4 pulmonary hypertension a population-based cohort study in Ontario, Canada. *Circ Cardiovasc Qual Outcomes.* 2018;11(2):13.
 41. Khera R, Krumholz HM. With great power comes great responsibility: big data research from the national inpatient sample. *Circ Cardiovasc Qual Outcomes.* 2017;10(7).
 42. De Coster C, Quan H, Finlayson A, et al. Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium. *BMC Health Serv Res.* 2006;6:77.