

# Validation of a claims-based algorithm to identify patients with chronic thromboembolic pulmonary hypertension using electronic health record data

Simon Teal<sup>1</sup>, William R. Auger<sup>2</sup>, Rodney J. Hughes<sup>3</sup>, Dena Rosen Ramey<sup>4</sup>, Kelly S. Lewis<sup>5</sup>, Gerald O'Brien<sup>5</sup>, Avin Yaldo<sup>5</sup>, Tanya M. Burton<sup>6</sup>, Tim Bancroft<sup>6</sup>, Jerry Seare<sup>6</sup> and Joerg Fabian<sup>7</sup>

<sup>1</sup>Bayer AG, Berlin, Germany; <sup>2</sup>University of California, San Diego, LA Jolla, CA, USA; <sup>3</sup>Sheffield Teaching Hospitals, Sheffield, UK; <sup>4</sup>Merck & Co, Inc, Kenilworth, NJ, USA; <sup>5</sup>Bayer Healthcare Pharmaceuticals, Whippany, NJ, USA; <sup>6</sup>Optum, Eden Prairie, MN, USA; <sup>7</sup>Bayer Business Services GmbH, Leverkusen, Germany

## Abstract

This study aimed to validate an algorithm developed to identify chronic thromboembolic pulmonary hypertension (CTEPH) among patients with a history of pulmonary embolism. Validation was halted because too few patients had gold-standard evidence of CTEPH in the administrative claims/electronic health records database, suggesting that CTEPH is underdiagnosed.

## Keywords

administrative claims, healthcare, databases, factual, early diagnosis

Date received: 18 January 2018; accepted: 14 October 2018

Pulmonary Circulation 2018; 9(1) 1–4

DOI: 10.1177/2045894018814772

Chronic thromboembolic pulmonary hypertension (CTEPH) is a rare, debilitating, and life-threatening condition; affected patients experience progressive dyspnea on exertion and may also present with fatigue, syncope, hemoptysis, and right heart failure.<sup>1</sup> Approximately 75% of patients with CTEPH have a history of an acute pulmonary embolism (PE),<sup>2</sup> with estimates of post-PE CTEPH incidence in the range of 0.56–3.2%.<sup>3</sup>

Interventional procedures such as pulmonary endarterectomy (PEA) and balloon pulmonary angioplasty (BPA), and medical therapy with the soluble guanylate cyclase stimulator riociguat, can significantly improve clinical outcomes among patients with CTEPH.<sup>4–6</sup> Unfortunately, CTEPH remains highly underdiagnosed worldwide<sup>7</sup> and prognosis is poor without treatment.<sup>8</sup> To help enable earlier diagnosis and treatment, we developed an algorithm to identify patients with a history of PE who are likely to have CTEPH on the basis of diagnoses, procedures, and tests found in an administrative claims database.<sup>9</sup> The present study aimed to validate this algorithm using an external

claims database in conjunction with electronic health records (EHR). However, the validation was halted because too few patients had gold-standard evidence of CTEPH, further highlighting the problem of underdiagnosis of CTEPH in clinical practice.

## Methods

### Algorithm

The CTEPH patient identification algorithm was developed using administrative claims from the MarketScan database.<sup>9</sup> The algorithm categorizes patients with PE as having high, intermediate, low, or no likelihood of developing CTEPH (Fig. 1).

Corresponding author:

Simon Teal, Bayer AG, Müllerstrasse 178, 13353 Berlin, Germany.

Email: [simon.teal@bayer.com](mailto:simon.teal@bayer.com)



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

© The Author(s) 2018.

Article reuse guidelines:  
[sagepub.com/journals-permissions](http://sagepub.com/journals-permissions)  
[journals.sagepub.com/home/pul](http://journals.sagepub.com/home/pul)



### Study sample

The validation included all commercial and Medicare Advantage health plan enrollees with at least one PE diagnosis code in any position (International Classification of Diseases, 9th edition, Clinical Modification [ICD-9-CM] 415.1, 415.11, 415.13, 415.19) on a medical claim in an external claims/EHR database (Optum Clinical Database) from 1 July 2008 through 30 June 2015 (identification period); the index date was defined as the first date of the PE medical claim during the identification period. Additional selection criteria included continuous enrollment with medical and pharmacy benefits for at least two years before the index date to four years after the index date (to reflect the time window used for algorithm development); age  $\geq 18$  years as of the index date; and participation in a medical practice contributing clinical notes to the EHR database or evidence of BPA or PEA in claims or EHR. Because no identifiable protected health information was extracted or accessed during the course of the study, Institutional Review Board approval or waiver of authorization was not required.

### CTEPH gold-standard criteria

Gold-standard criteria were established to classify PE patients with “definitive CTEPH” and “suggestive CTEPH.” Evidence for definitive CTEPH was derived from searching the provider notes in EHR for “CTEPH” and other terms related to chronic thromboembolic disease, including pulmonary hypertension (PH) combined with PE or pulmonary location of thrombosis/embolus and terms for BPA or PEA. Procedure codes for BPA or PEA in claims or EHR were also considered to be definitive evidence of CTEPH.

Evidence for suggestive CTEPH was derived from searching the provider notes for “pulmonary hypertension” or “PH” plus terms for PE or other thromboembolic disease. To confirm whether these patients met the gold-standard CTEPH criteria, notes from up to three recent dates of service with mentions of PH were abstracted for a subset of 20 patients and reviewed by the study’s medical experts to check for an explicit PE diagnosis, evidence of a vascular filling defect on pulmonary angiography, evidence of elevated pulmonary artery pressure measurement ( $\geq 25$  mmHg), or an explicit PH diagnosis noted as a current problem plus an explicit statement that the PH was chronic or active for at least six months. Patients with evidence of congestive heart failure were excluded from consideration.

### Statistical analysis

Variables were summarized descriptively using SAS 9.4 (SAS Institute, Cary, NC, USA).

### Results

Of more than 5.7 million health plan enrollees with medical and pharmacy benefits in the claims/EHR database during

the identification period, 36,791 (0.6%) had at least one medical claim with a PE diagnosis code and 2473 (7%) of these met the continuous enrollment criteria. After 12 enrollees were dropped for missing demographic data and 1406 for not being EHR notes-eligible, the final analytic sample contained 1055 patients. Of 63 patients in the final sample who had some evidence of CTEPH, only four met the gold-standard criteria: two with a PEA procedure code and two with suggestive CTEPH terms plus clinical evidence of CTEPH in the EHR notes. As only  $< 0.4\%$  of the study cohort met the gold-standard criteria, the validation was halted.

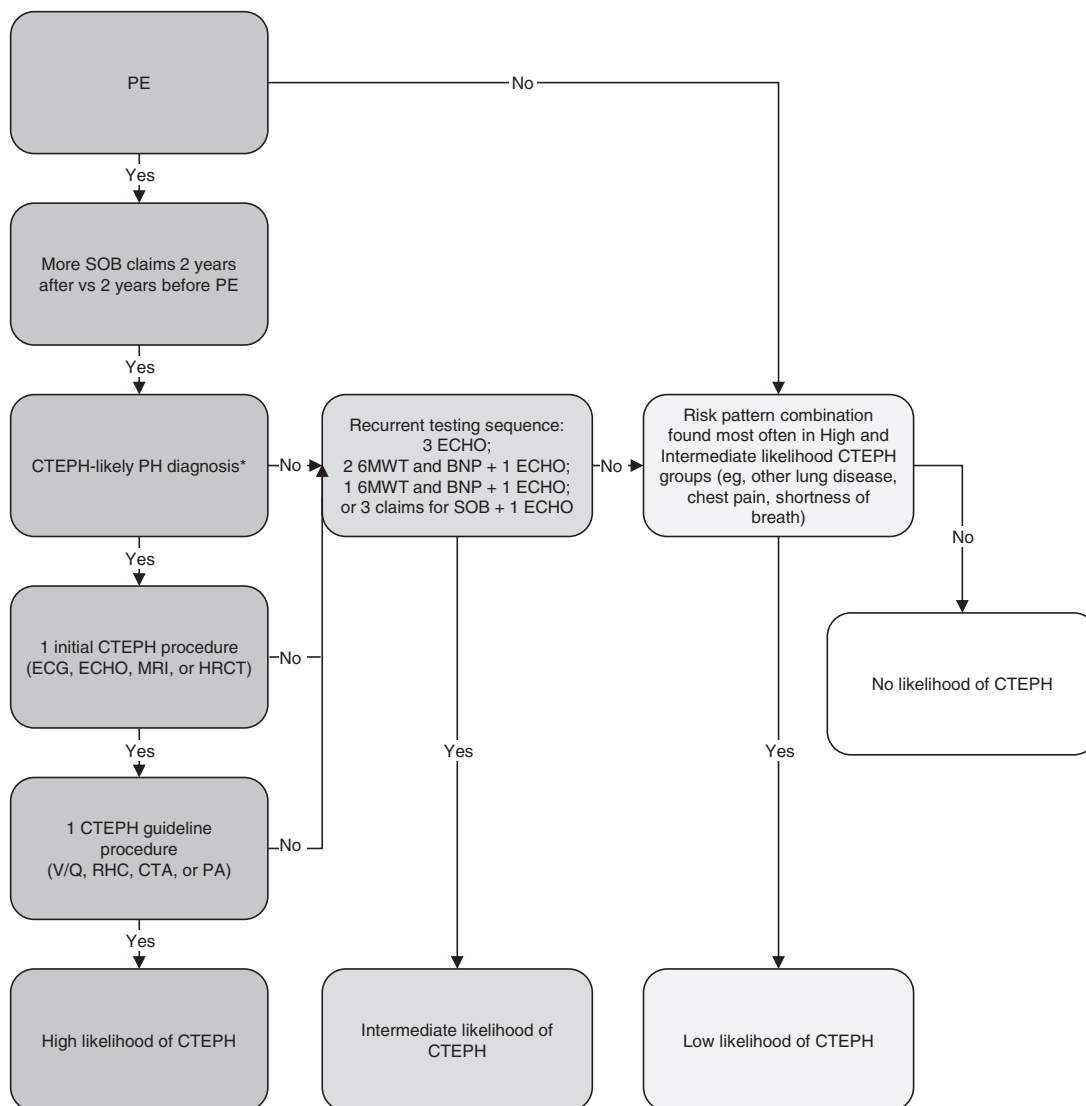
### Discussion

This algorithm validation study could not be completed because only 4/1055 (0.38%) of the final PE sample met the gold-standard criteria for CTEPH. Although the true frequency of CTEPH is unknown, the post-PE incidence has been estimated to be in the range of 0.56–3.2%.<sup>3</sup> The lower than expected proportion of PE patients with gold-standard CTEPH evidence identified here suggests that the proper screening, diagnosis, and/or treatment recommended for CTEPH after PE diagnosis has not been consistently performed.

Some of the challenges we encountered in patient identification were due to limitations of the data sources, as substantial sample attrition was caused by lack of available claims and/or EHR data. For example, the continuous enrollment requirement, which was necessary to reflect the time window used during algorithm development, resulted in elimination of  $> 93\%$  of the 36,791 patients with a PE diagnosis code. In addition, lack of a specific ICD-9-CM code likely made CTEPH less observable in the database and more difficult to track over time.

Importantly, though, in the context of the existing literature, our findings suggest a failure of the clinical diagnosis and post-PE referral pathway for CTEPH. Algorithm validation relied on documentation of CTEPH-related diagnoses, tests, and procedures in the integrated claims/EHR database. However, CTEPH is known to be a profoundly underdiagnosed condition, with only an estimated 7–29% of all patients with CTEPH identified across the US and Europe.<sup>7</sup> In addition, even patients who are correctly diagnosed with CTEPH frequently do not receive the recommended treatment. Approximately two-thirds of the estimated 3000 yearly US cases of CTEPH would be candidates for PEA, the current standard of care;<sup>10</sup> yet,  $< 500$  PEA procedures are performed in the US each year.<sup>11</sup>

This widespread failure to diagnose and treat patients with CTEPH has been attributed to a variety of causes, including non-specific symptoms, lack of adherence to international diagnostic guidelines, and delayed referral of patients to expert centers.<sup>12</sup> Multiple studies have revealed a lack of provider understanding that patients with PE are at risk for CTEPH and should be monitored long-term for



**Fig. 1.** CTEPH patient identification algorithm. 6MWT, 6-min walk test; BNP, brain natriuretic peptide test; CTA, computed tomography angiography; CTEPH, chronic thromboembolic pulmonary hypertension; ECHO, echocardiogram; HRCT, high-resolution computed tomography; MRI, magnetic resonance imaging; PA, pulmonary angiography; PE, pulmonary embolism; RHC, right heart catheterization; SOB, shortness of breath; V/Q, ventilation/perfusion scanning. \*Defined as at least one medical claim with a diagnosis code for chronic PE, embolism, and thrombosis of unspecified artery, or other PH without primary PH.

signs of PH.<sup>13,14</sup> In a recent claims database analysis, only 55% of patients with PH symptoms underwent any diagnostic testing over a two-year follow-up period and the most commonly performed test was echocardiography, which is useful in PH screening but cannot accurately diagnose CTEPH.<sup>14</sup> Masking of CTEPH by other conditions is another potential source of missed diagnoses, as many patients who recover from an acute PE have significant cardiopulmonary co-morbidities to which findings of PH may be prematurely attributed. CTEPH may also be missed if ongoing dyspnea and reduced exercise tolerance after PE, known as post-PE syndrome,<sup>15</sup> is not investigated because physicians and patients accept long-term morbidity as an expected outcome. Furthermore, the prevalent misconception that PEA is too risky and/or should be considered only

after medical treatment fails is an additional barrier to appropriate referral for surgical evaluation.<sup>13</sup> Taken together, it is probable that these factors contributed heavily to the absence of relevant provider notes and diagnosis/procedure codes in the claims/EHR data.

### Study limitations

Identification of patients with CTEPH from claims data is complicated because there was no unique ICD-9-CM diagnosis code for CTEPH, and many of the relevant symptoms are non-specific. Diagnosis codes that were used to help classify patients as having CTEPH may also apply to other conditions, which may have led to some patient misclassification. Our findings should also be interpreted with

consideration of certain limitations common to all analyses conducted with administrative claims and EHR. Reported diagnosis codes may not reflect confirmed diagnoses, because codes may have been entered incorrectly, included as rule-out criteria, or chosen for billing optimization. Claims data do not include information about treatments administered during clinical trials, use of over-the-counter medications, samples provided by physicians, or medications filled but not taken as prescribed. Similarly, EHR data do not include the entire medical chart; this raises the possibility of missing data, the extent of which is unknown.

## Conclusions

The pervasive underdiagnosis of CTEPH was likely a substantial contributor to the inability to complete this study, suggesting an urgent need for improved education to reduce the significant gaps that exist between guidelines and real-life clinical practice regarding identification and management of CTEPH. The development and implementation of CTEPH screening tools, such as the algorithm the present study was designed to assess, may help improve CTEPH diagnosis and clinical outcomes.

## Acknowledgments

Medical writing services were provided by Yvette M. Edmonds, an employee of Optum (Eden Prairie, MN, USA). The authors thank Ed Wang (Bayer), Margarita de la Orden (Bayer), Tina Nikolova (Bayer), and Robert Lee Boggs (Merck) for contributions to the study design; and Ed Wang, Sören Hörnig (Bayer), and Carl Baxter (Merck) for contributions to interpretation of the results.

## Conflict of interest

ST is an employee of Bayer AG (Berlin, Germany), which funded the study. TMB, TB, and JS are employees of Optum (Eden Prairie, MN, USA), which was contracted by Bayer AG to perform a portion of this study. Authors WRA and RJH provided consulting services to Bayer AG. KSL, GO, and AY are employees of Bayer Healthcare Pharmaceuticals (Whippany, NJ, USA). JF is an employee of Bayer Business Services GmbH (Leverkusen, Germany). DRR is an employee of Merck & Co, Inc (Kenilworth, NJ, USA).

## Funding

This study was funded by Bayer AG (Berlin, Germany), which participated in designing the study, analyzing, and interpreting the data, preparing the manuscript, and deciding to submit the manuscript for publication.

## References

1. Hoepfer MM, Madani MM, Nakanishi N, et al. Chronic thromboembolic pulmonary hypertension. *Lancet Respir Med* 2014; 2(7): 573–582.

2. Pepke-Zaba J, Delcroix M, Lang I, et al. Chronic thromboembolic pulmonary hypertension (CTEPH): results from an international prospective registry. *Circulation* 2011; 124(18): 1973–1981.
3. Ende-Verhaar YM, Cannegieter SC, Vonk Noordegraaf A, et al. Incidence of chronic thromboembolic pulmonary hypertension after acute pulmonary embolism: a contemporary view of the published literature. *Eur Respir J* 2017; 49(2): 1601792.
4. Mayer E, Jenkins D, Lindner J, et al. Surgical management and outcome of patients with chronic thromboembolic pulmonary hypertension: results from an international prospective registry. *J Thorac Cardiovasc Surg* 2011; 141(3): 702–710.
5. Madani MM, Auger WR, Pretorius V, et al. Pulmonary endarterectomy: recent changes in a single institution's experience of more than 2,700 patients. *Ann Thorac Surg* 2012; 94(1): 97–103.
6. Ghofrani HA, D'Armini AM, Grimminger F, et al. Riociguat for the treatment of chronic thromboembolic pulmonary hypertension. *N Engl J Med* 2013; 369(4): 319–329.
7. Gall H, Hoepfer MM, Richter MJ, et al. An epidemiological analysis of the burden of chronic thromboembolic pulmonary hypertension in the USA, Europe and Japan. *Eur Respir Rev* 2017; 26(143): 160121.
8. Fedullo P, Kerr KM, Kim NH, et al. Chronic thromboembolic pulmonary hypertension. *Am J Respir Crit Care Med* 2011; 183(12): 1605–1613.
9. Teal S, Auger WR, Hughes AJ, et al. Development of a claims-based algorithm to identify patients with chronic thromboembolic pulmonary hypertension. *Am J Respir Crit Care Med* 2017; 195: A2840.
10. Beckman MG, Hooper WC, Critchley SE, et al. Venous thromboembolism: a public health concern. *Am J Prev Med* 2010; 38(4 Suppl): S495–501.
11. Edward JA and Mandras S. An update on the management of chronic thromboembolic pulmonary hypertension. *Curr Probl Cardiol* 2017; 42(1): 7–38.
12. Gopalan D, Delcroix M and Held M. Diagnosis of chronic thromboembolic pulmonary hypertension. *Eur Respir Rev* 2017; 26(143): 160108.
13. Gall H, Preston IR, Hinzmann B, et al. An international physician survey of chronic thromboembolic pulmonary hypertension management. *Pulm Circ* 2016; 6(4): 472–482.
14. Tapson VF, Platt DM, Xia F, et al. Monitoring for pulmonary hypertension following pulmonary embolism: the INFORM study. *Am J Med* 2016; 129(9): 978–985.
15. Klok FA, van der Hulle T, den Exter PL, et al. The post-PE syndrome: a new concept for chronic complications of pulmonary embolism. *Blood Rev* 2014; 28(6): 221–226.