Research | **Open Access**

Performance of an administrative claims algorithm to estimate the incidence of pure red cell aplasia in chronic hepatitis C patients

Scott C. Quinlan[1*], Jennifer C.L. Hawes[1], Daniel Mines[2], Shirin Ahmed[2], Stephan Lanes[1], Vinay Mehta[2], Crystal N. Holick[1], Nancy Santanello[2] and T. Christopher Mast[2]

*Correspondence: squinlan@healthcore.com

CrossMark
← Click for updates

[1]HealthCore, Inc., 800 Delaware Avenue, Fifth Floor, Wilmington, DE 19801, USA.
[2]Department of Epidemiology, Merck Research Laboratories, 351 N. Sumneytown Pike, North Wales, PA 19454, USA.

Abstract

Background: We developed and validated an algorithm to evaluate pure red cell aplasia (PRCA) incidence in chronic hepatitis C (CHC) patients in a large observational database.

Methods: We conducted a retrospective study using the HealthCore Integrated Research Database[SM] (HIRD) in which we identified CHC patients and followed them forward to identify PRCA occurrences. Possible PRCA cases were identified based on a medical claim for aplastic anemia (ICD-9-CM 284.8x or 284.9x) with a claim for bone marrow biopsy in the prior 30 days. Medical records were requested and reviewed by an Adjudication Committee (AC) to confirm PRCA case status. The positive predictive value (PPV) and 95% confidence interval (CI) for the PRCA algorithm were estimated based on adjudication results.

Results: A total of 36,164 CHC patients were identified yielding 25 suspected PRCA cases. Medical records were obtained and reviewed for 17 cases. Of these 17 cases, none were confirmed as PRCA (PPV: 0.0%; 95% CI 0.0-19.5%). Estimated confirmed PRCA incidence was 0.0/1,000 person-years (95% CI 0.00-0.05/1,000 person-years).

Conclusions: Automated case definitions for PRCA performed poorly in identifying PRCA in CHC patients, limiting our ability to estimate PRCA incidence. PRCA in CHC patients is rare and difficult to study using large automated databases.

Keywords: Observational study, hepatitis C, chronic, red-cell aplasia, pure, incidence

Introduction

Pure red cell aplasia (PRCA) is a rare hematological disorder characterized by severe anemia, reticulocytopenia, and almost complete absence of erythroid precursor cells in the bone marrow. All other cell lineages are present and appear morphologically normal [1]. Pure red cell aplasia is an acquired anemia that may be primary or develop secondary to a variety of neoplastic, autoimmune, or infectious diseases [1].

There have been case reports of PRCA developing in patients undergoing treatment for chronic hepatitis C (CHC) [2-8]. Limited data are available on the population-based PRCA incidence in patients treated for CHC. One hospital-based survey conducted

in France of 6,630 treated CHC patients over one year found that PRCA developed in two of 581 patients who received EPO concomitantly (3.4 cases per 1,000 patients) [9]. The two PRCA cases were based on physician reporting and were not otherwise verified but the investigators were confident in their validity because one of the drug manufacturers also reported two cases of PRCA in CHC treated patients during the same time period [9].

The rarity of PRCA has limited the ability to conduct epidemiologic studies, as large cohorts of patients are needed to identify sufficient cases for examination. Administrative claims databases offer large populations and have recently

been used to estimate incidence of PRCA. Collins and colleagues examined PRCA in United States (US) dialysis patients treated with erythropoietin (EPO) using Medicare data [10]. The investigators used a combination of aplastic anemia diagnosis codes and bone marrow biopsy procedure codes to identify possible PRCA cases, some of which were then ruled-out based on other competing diagnoses, complicating factors, or a clinical course inconsistent with PRCA. Collins et al., found PRCA to be rare in the population studied, possibly as low as one case out of 101,782 dialysis patients examined over a total of 70,707 person-years of EPO exposure [10]. A limitation of the study by Collins and colleagues was that PRCA cases identified in Medicare data were not validated through medical record review, so there is uncertainty that the cases identified were PRCA.

We conducted a large retrospective cohort study using the HealthCore Integrated Research Database℠ (HIRD) to further develop, and validate through medical record review, an algorithm for identifying PRCA in administrative claims data and to use this refined algorithm to estimate PRCA incidence in a cohort of CHC patients that included both treated and untreated patients.

Methods
Study design
We conducted a retrospective cohort study using the HIRD in which we identified a cohort of CHC patients. The HIRD consists of longitudinal medical, pharmacy, and enrollment claims data from 14 major US commercial health plans representing approximately 30 million commercially insured lives. Institutional Review Board approval was obtained prior to the initiation of any study activities.

Patient identification
Patients with CHC between January 1, 2006 and August 31, 2012 were identified based on having at least two medical claims with International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes for CHC (ICD-9-CM codes 070.44, 070.54, 070.70, or 070.71) separated by six months or less (CHC Algorithm 1) or at least one medical claim with an ICD-9-CM code for CHC along with at least one medical claim for CHC testing/genotyping or standard CHC treatment (see **Supplementary Appendix** for codes; CHC Algorithm 2). Two algorithms to identify CHC patients were used to improve the sensitivity of our CHC definition and to ensure sufficient sample size to detect cases of PRCA. The index date (cohort entry date) was the earliest date that the patient fulfilled either of the two CHC algorithms. No continuous health plan eligibility before cohort entry was required. Patients diagnosed with human immunodeficiency virus infection were excluded from the cohort due to privacy issues that would have prevented validation of any PRCA cases identified among these patients.

Follow-up for all cohort members began on index date and continued until the earliest of the end of the study period (August 31, 2012), disenrollment from the health plan, first prescription dispensing for a protease inhibitor (boceprevir or telaprevir), death, or first PRCA diagnosis. Follow-up time for each cohort member was stratified by whether the patient was currently receiving both pegylated interferon (PegIFN) and ribavirin (RBV) concurrently ("treated") or not ("untreated"). We censored follow-up at the time of protease inhibitor dispensing to evaluate PRCA incidence prior to any protease inhibitor exposure, as the goal of this study was to evaluate background PRCA incidence in CHC patients receiving the historical standard of care prior to introduction of direct-acting antiviral agents.

Case identification
Pure red cell aplasia was initially identified in the claims data using two algorithms. In the first algorithm, possible PRCA cases were identified by the presence of the following two factors:
- A medically attended healthcare visit with a diagnosis code for aplastic anemia (ICD-9-CM codes 284.8x or 284.9x) AND
- A medical claim for a bone marrow biopsy (ICD-9-CM procedure codes 41.31 or 41.38; Current Procedural Terminology codes 38221 or 85097) in the 30 days prior to and including the aplastic anemia claim date.

The algorithms were based on codes for aplastic anemia since there are currently no ICD-9-CM diagnosis codes for PRCA. A medically attended healthcare visit included a Current Procedural Terminology code starting with "99" as part of the claim, which corresponds to physician services for evaluation and management of a patient in the outpatient or inpatient settings, including consultations. The PRCA diagnosis date was the date of the medical claim for aplastic anemia.

The second PRCA algorithm identified a subset of possible PRCA cases (the first PRCA algorithm). These cases were expected to have an increased likelihood of being PRCA cases following medical record review based on the presence of additional clinical characteristics in the claims data related to the expected process of care of PRCA. The additional clinical characteristics were suggested by clinical hematologists consulted during protocol development. These cases were deemed probable PRCA based on the claims data and were required to meet the criteria for possible PRCA and at least one of the following three additional features:
- A medical claim for computed tomography scan of the chest within 30 days on or after the PRCA diagnosis date (see **Supplementary Appendix** for codes) OR
- A medical claim for a red blood cell transfusion within 30 days on or after the PRCA diagnosis date (see **Supplementary Appendix** for codes) OR
- A medical or pharmacy claim for the initiation of new immunosuppressive therapy within 60 days on or after the PRCA diagnosis date (see **Supplementary Appendix** for codes).

Case validation

For all possible PRCA cases identified, all medical claims surrounding the PRCA diagnosis date were reviewed to identify healthcare providers or facilities from which to request medical records to validate the claims-based PRCA diagnosis. Since probable PRCA cases were a subset of possible PRCA cases, all probable cases also underwent medical claims review for possible validation. At least one medical record was requested for each possible PRCA case. For some cases, more than one medical record was requested to capture the complete pattern of care surrounding the diagnosis date. If multiple providers or facilities were associated with the PRCA diagnosis date, records were preferentially requested from physicians with a specialty in hematology or oncology. To be eligible for review, at least one medical record must have been received that included the necessary clinical information to validate the PRCA diagnosis, which included bone marrow biopsy results and complete blood count information. For a small number of possible PRCA cases no medical record was available with bone marrow biopsy results or complete blood count information, but a record was available that included other clinical information that allowed the PRCA diagnosis to be evaluated. These additional possible PRCA cases were also reviewed.

Three physicians specializing in hematology and/or oncology were recruited to serve as an Adjudication Committee (AC) and were responsible for reviewing the medical records obtained for the claims-identified possible PRCA cases to determine whether or not the cases had PRCA. Medical records were provided to the AC for review and were independently adjudicated by all three AC members to determine whether or not the patient had PRCA and, if so, the date of PRCA diagnosis, presence of CHC, and other attributes that confirmed case status or provided additional clinical detail beyond what was found in the claims data. If a case was not confirmed as PRCA, the adjudicators suggested alternative diagnoses. Members of the AC were blinded to each other's reviews. The chair of the AC was tasked with assisting the committee in achieving a consensus for a given case if there were discrepancies among AC members.

Statistical analysis

We calculated descriptive statistics for the full CHC cohort and all possible and probable PRCA cases identified in the claims data. For all possible PRCA cases identified in the claims data, we present descriptive statistics both overall and stratified by whether or not the possible PRCA case underwent adjudication.

We used adjudication results to estimate the positive predictive value (PPV) of the two claims-based algorithms that were developed to identify PRCA. The PPV was estimated by taking the number of PRCA cases confirmed by the AC as PRCA (true positives) divided by the total number of claims-identified possible PRCA cases that underwent adjudication. The PPV is presented as a percentage along with 95% confidence intervals

(CIs) calculated using the exact binomial method. Additional information from the adjudication is also summarized and includes AC comments on the reasons possible PRCA cases identified in the claims data were not confirmed following medical record adjudication.

We estimated the overall PRCA incidence rate by dividing the number of confirmed PRCA cases by the person-time at-risk for the cohort during the study period. We computed exact 95% CIs for the incidence rate using a method designed for rare events [11].

Results

Patient characteristics

We identified 36,164 CHC patients in the HIRD with an average age of 51 years. The majority of CHC patients were male (61.0%), 911 (2.5%) patients were co-infected with hepatitis B and 4,005 (11.1%) had a history of anemia other than aplastic anemia prior to the index date (**Table 1**). A total of 5,798 (16.0%) CHC patients had at least one treatment episode with PegIFN and RBV during follow-up.

Case characteristics

The cohort of CHC patients accrued 75,132 person-years of follow-up during which 25 (<0.1%) possible PRCA cases were identified based on the claims data (**Table 1**). More possible PRCA cases were male (56.0%). Nearly half of all possible PRCA cases had a history of anemia (48.0%) or thrombocytopenia (40.0%) and six (24.0%) had a history of hematologic malignancy in the one year prior to the index date. Five (20.0%) possible PRCA cases had a history of immunosuppressive therapy use in the one year prior to index date (**Table 1**). Compared to the overall CHC cohort, possible PRCA cases were more likely to have a history of hepatitis B virus infection (8.0% versus 2.5%), infectious mononucleosis (4.0% versus 0.1%), chronic renal insufficiency (8.0% versus 3.3%), anemia (48.0% versus 11.1%), thrombocytopenia (40.0% versus 3.8%), hematologic malignancy (24.0% versus 0.7%), and solid tumor (12.0% versus 0.4%) prior to the index date (**Table 1**). Possible PRCA cases were also more likely than those in the overall CHC cohort to have a history of EPO therapy (8.0% versus 1.7%) and immunosuppressive drug therapy (20.0% versus 10.4%) prior to the index date. Four (16.0%) of the 25 possible PRCA cases identified had a claims-based diagnosis of aplastic anemia in the one year prior to the index date. Of the 25 possible PRCA cases identified in the claims data, three (12.0%) had at least one treatment episode with PegIFN and RBV during follow-up, and one (4.0%) of the possible PRCA cases occurred during a PegIFN and RBV treatment episode.

Out of the 25 possible PRCA cases identified, 21 (84.0%) met the automated case definition for probable PRCA (**Table 2**). The most common reason a possible PRCA case qualified as probable was a medical claim for a red blood cell transfusion (56.0%) in the 30 days on or after the PRCA diagnosis date. The most common physician specialty on the medical claim

Table 1. Demographics and baseline characteristics of chronic hepatitis C (CHC) patients and possible pure red cell aplasia (PRCA) identified in the claims data.

N (%) unless otherwise specified	Overall		Possible PRCA*	
	N	(%)	N	(%)
Total	**36,164**	--	**25**	--
Age, years, mean±SD	51.2	±11.6	51.6	±18.1
Gender, female	14,118	(39.0)	11	(44.0)
Year of CHC cohort entry (index date†)				
2006	9,840	(27.2)	15	(60.0)
2007	6,232	(17.2)	4	(16.0)
2008	5,246	(14.5)	4	(16.0)
2009	4,191	(11.6)	--	(0.0)
2010	3,837	(10.6)	1	(4.0)
2011	4,197	(11.6)	--	(0.0)
2012	2,621	(7.2)	1	(4.0)
Co-morbidities prior to index date†				
Cytomegalovirus infection	15	(0.0)	--	(0.0)
Infectious mononucleosis	29	(0.1)	1	(4.0)
Hepatitis A virus infection	176	(0.5)	--	(0.0)
Hepatitis B virus infection	911	(2.5)	2	(8.0)
Mumps	1	(0.0)	--	(0.0)
Parvovirus B19	--	(0.0)	--	(0.0)
Chronic renal insufficiency	1,202	(3.3)	2	(8.0)
End-stage renal disease	497	(1.4)	--	(0.0)
Medical history prior to index date†				
Anemia	4,005	(11.1)	12	(48.0)
Neutropenia	467	(1.3)	7	(28.0)
Pancytopenia	266	(0.7)	1	(4.0)
Thrombocytopenia	1,371	(3.8)	10	(40.0)
Myelodysplastic syndrome	16	(0.0)	--	(0.0)
Possible PRCA	14	(0.0)	4	(16.0)
Autoimmune diseases				
Collagen vascular disease	791	(2.2)	3	(12.0)
Autoimmune disease	319	(0.9)	1	(4.0)
Neoplasm				
Thymoma	1	(0.0)	--	(0.0)
Hematologic malignancy	254	(0.7)	6	(24.0)
Solid tumor	137	(0.4)	3	(12.0)
Co-medication use prior to index date†				
Erythropoietin (EPO)	624	(1.7)	2	(8.0)
Anti-epileptic medications	403	(1.1)	1	(4.0)
Azathioprine	102	(0.3)	--	(0.0)
Chloraphenicol	--	(0.0)	--	(0.0)
Immunosuppressive therapy	3,751	(10.4)	5	(20.0)
Sulfonamides	2,849	(7.9)	--	(0.0)
Isoniazid	20	(0.1)	--	(0.0)

Continuation of Table 1.

Co-medication use prior to index date†				
Procainamide	1	(0.0)	--	(0.0)
Protease inhibitor	56	(0.2)	--	(0.0)

CHC: Chronic hepatitis C; EPO: Erythropoietin; ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification; PRCA: Pure red cell aplasia; SD: Standard deviation
*Possible PRCA cases from claims data are patients with any medically-attended claim for aplastic anemia (ICD-9-CM 284.8x or 284.9x) and any medical claim for a bone marrow biopsy in the 30 days prior to and including the date of the claim for aplastic anemia.
†Earliest date a patient satisfied either of two algorithms for identifying CHC in the claims data.

for PRCA was hematology or oncology (**Table 2**). Nine (36.0%) possible PRCA cases had a history of EPO use in the three months prior to the possible PRCA diagnosis date and 10 (40.0%) had a history of immunosuppressive therapy in the two months prior to the possible PRCA diagnosis date (**Table 2**). Out of 25 possible PRCA cases identified in the claims data, medical records with the necessary clinical information for review were obtained and adjudicated for 17 cases (**Table 2**). Of the 17 cases that underwent adjudication, two (11.8%) had a claims-based diagnosis of aplastic anemia prior to the index date. Compared to cases that underwent adjudication, the eight PRCA cases where medical records were not available for adjudication were older, more likely to be female, and less likely to be diagnosed by a hematologist or oncologist (**Table 2**).

Validation results
Of the 17 possible PRCA and 15 probable PRCA cases that underwent adjudication, none (PPV: 0.0%; 95% CI 0.0-19.5% and PPV: 0.0%; 95% CI 0.0-21.8%) were confirmed as PRCA following medical record review by the AC (**Table 3**). After determining that none of the adjudicated cases identified in the claims data could be confirmed as PRCA, the AC reviewed the medical records in order to determine the actual underlying diagnoses. In most patients, these diagnoses included hematologic malignancies or solid tumors and their associated treatments that had deleterious effects on the bone marrow (**Supplementary Table S1**). Other diagnoses included other types of anemia, as well as an abnormal bone marrow not consistent with the pattern expected for PRCA (**Supplementary Table S1**).

The cohort of CHC patients accrued a total of 75,132 person-years of at-risk follow-up time, of which 3,033 person-years were during treatment with PegIFN and RBV and 72,099 person-years were untreated (**Table 4**). If we estimate the PRCA incidence rate based on the number of possible PRCA cases confirmed following medical record review (N=0), the estimated PRCA incidence rate was 0.00 cases per 1,000 person-years (95% CI: 0.00-0.05 cases per 1,000 person-years). As a sensitivity analysis, we assumed that the eight possible PRCA

Table 2. Clinical characteristics of chronic hepatitis C (CHC) patients with possible and pure red cell aplasia (PRCA) identified by a claims-based algorithm*.

N (%) unless otherwise specified	Possible PRCA*		Adjudicated		Not Adjudicated†	
	N	(%)	N	(%)	N	(%)
Total	**25**	--	**17**	--	**8**	--
Age, years, mean±SD	51.6	±18.1	47.9	±20.2	59.3	±9.8
Gender, female	11	(44.0)	6	(35.3)	5	(62.5)
Probable PRCA definition						
Any of the below criteria	21	(84.0)	15	(88.2)	6	(75.0)
≥1 medical claim for chest CT scan	8	(32.0)	6	(35.3)	2	(25.0)
≥1 medical claim for RBC transfusion	14	(56.0)	10	(58.8)	4	(50.0)
≥1 pharmacy claim for the initiation of a new immunosuppressive therapy	11	(44.0)	8	(47.1)	3	(37.5)
Physician specialty on the first diagnosis of PRCA						
Emergency medicine	2	(8.0)	1	(5.9)	1	(12.5)
Hematology or oncology	10	(40.0)	8	(47.1)	2	(25.0)
Internal medicine	6	(24.0)	3	(17.6)	3	(37.5)
Other (non-physician)	3	(12.0)	2	(11.8)	1	(12.5)
Pediatrics (all types)	1	(4.0)	1	(5.9)	--	(0.0)
Not specified	3	(12.0)	2	(11.8)	1	(12.5)
Co-medication use in the 60 days prior to PRCA diagnosis						
Erythropoietin (EPO) in the three months prior to PRCA diagnosis	9	(36.0)	5	(29.4)	4	(50.0)
Anti-epileptic medications	2	(8.0)	2	(11.8)	--	(0.0)
Azathioprine	--	(0.0)	--	(0.0)	--	(0.0)
Chloraphenicol	--	(0.0)	--	(0.0)	--	(0.0)
Immunosuppressive therapy	10	(40.0)	8	(47.1)	2	(25.0)
Sulfonamides	4	(16.0)	4	(23.5)	--	(0.0)
Isoniazid	--	(0.0)	--	(0.0)	--	(0.0)
Procainamide	--	(0.0)	--	(0.0)	--	(0.0)

CHC: Chronic hepatitis C; CT: Computed tomography; EPO: Erythropoietin; ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification; PRCA: Pure red cell aplasia; RBC: Red blood cell; SD: Standard deviation
*Possible PRCA cases from claims data are patients with any medically-attended claim for aplastic anemia (ICD-9-CM 284.8x or 284.9x) and any medical claim for a bone marrow biopsy in the 30 days prior to and including the date of the claim for aplastic anemia.
†No medical record received or received medical records that did not contain the required key elements (i.e., bone marrow biopsy result and complete blood count report) for adjudication.

cases that were not adjudicated would have been confirmed as PRCA if they had been adjudicated, and the PRCA incidence rate was re-estimated as 0.11 cases per 1,000 person-years (95% CI: 0.05-0.21 cases per 1,000 person-years).

Discussion

The results of this study indicate that claims-based algorithms performed poorly in identifying PRCA cases. The PRCA algorithm that identified possible PRCA required an aplastic anemia diagnosis accompanied by a claim for bone marrow biopsy in the 30 days on or before the aplastic anemia date, a pattern of clinical care thought to be consistent with PRCA, but one that might also be consistent with the pattern of care expected for a number of other conditions including certain cancers. The requirement for a red blood cell transfusion or computed tomography scan of the chest that was included as part of the second algorithm to identify probable PRCA in the claims data could also be consistent with patterns of care for certain cancers, particularly leukemia and metastatic cancers that have spread to the bone marrow. These cancers occur much more commonly than PRCA.

Difficulties in identifying PRCA using claims data have been encountered before and are likely related to the extreme rarity of the condition [10]. The low prevalence of possible PRCA based on the claims data in our cohort is consistent with previous studies [9,10].

Table 3. Medical record validation of pure red cell aplasia (PRCA) cases identified by claims-based algorithms.

Total	Possible PRCA	Probable PRCA
	17	15
Number of true positives*	0	0
Number of false positives†	17	15
Positive predictive value (PPV; 95% CI)‡	0.0 (0.0-19.5)	0.0 (0.0-21.8)

CI: Confidence interval; PPV: Positive predictive value; PRCA: Pure red cell aplasia
*Evidence of PRCA identified from claims data using the preliminary claims-based algorithm and confirmed from medical record review.
†Evidence of PRCA identified from claims data using the preliminary claims-based algorithm but not confirmed from medical record review.
‡Positive predictive value (PPV) is defined as the number of true positives divided by the sum of the number of true positives and the number of false positives. 95% CIs are exact binomial CIs.

Table 4. Pure red cell aplasia (PRCA) incidence rates.

	Overall
Total number of possible PRCA cases identified from claims data	25
Number of possible PRCA cases identified from claims data with received medical records that underwent medical record review.	17
Number of true positive PRCA cases*	0
Number of possible PRCA cases† identified from claims data for which no medical record was received or received medical records did not contain the required key elements for adjudication	8
At-risk person-time, person-years, total	75,132
PRCA IR‡, true positive events per 1,000 person-years (95% CI)	0.00 (0.00-0.05)

CI: Confidence interval; IR: Incidence rate; PRCA: Pure red cell aplasia
*Evidence of true positive PRCA identified from claims data using the preliminary claims-based algorithm and confirmed from medical record review.
†Evidence of possible PRCA identified from claims data using the preliminary claims-based algorithm and confirmed from medical record review. Eight claims-identified PRCA cases, for which no medical record was received, were assumed to be possible positive PRCA cases.
‡Incidence rate (IR) is defined as the number of new true positive PRCA cases divided by the total at-risk person-time.

One of the challenges in developing a claims-based algorithm for identifying PRCA is the lack of an ICD-9-CM code specific for acquired PRCA. In the International Classification of Diseases, Tenth Revision coding system, there are codes specific to PRCA (starting with D60) which may be more widely adopted and could then be used to develop a more specific claims-based algorithm for identifying PRCA. The rarity of PRCA also poses a challenge when trying to develop a claims-based algorithm.

Previous studies of PRCA in CHC patients have largely been single case reports pointing to the rarity of the disease [2-9], and studies in populations of dialysis patients have indicated that the prevalence of PRCA could be as low as one case per 10,000 or 100,000 patients [10,12], which is consistent with the upper bound of the 95% CI for the PRCA incidence estimates we report. One hospital-based survey in France identified 2 PRCA cases among 6,630 CHC patients, however both cases were receiving EPO therapy which has been linked to PRCA development through the formation of anti-erythropoietin antibodies [6,7]. The formulation of EPO varies by country, and different formulations may confer different PRCA risks leading to variations in PRCA prevalence by country among EPO treated patients [12]. It is possible, therefore, that there were no PRCA cases in our cohort of more than 36,000 CHC patients. A future study might consider looking for possible PRCA cases in a larger cohort not specific to CHC, perhaps including the full HIRD of approximately 30 million lives or all patients undergoing dialysis with the hope of identifying a small number of PRCA cases confirmed by medical record review. To potentially reduce the number of false positive cases identified in this larger sample, researchers might consider excluding from medical record review potential PRCA cases who also have diagnoses for leukemia or other metastatic cancers in the same timeframe, as our results show that bone marrow abnormalities are much more likely to be related to these cancers and their treatment than to PRCA. The shared information in the medical records of these confirmed PRCA cases along with other clinical information in the claims data could be used to further refine the claims-based algorithm for identifying PRCA to improve its performance, which could then be tested in a different cohort of patients. It is possible that by reviewing the medical records and claims data of confirmed PRCA cases researchers might identify additional shared diagnoses or PRCA-specific treatment patterns that could be added to the original algorithm to capture additional PRCA cases not identified using the original algorithm that required an aplastic anemia diagnosis. Future researchers might also consider conducting a study of PRCA using a data source that has access to both claims data and electronic health records, which might allow for more efficient identification and screening of potential PRCA cases due to the additional clinical detail available in electronic health records that is not available in claims data. For example, electronic health records might offer the possibility of conducting key word searches in physician-recorded note fields to identify possible PRCA cases that might not have received an aplastic anemia diagnosis code and therefore would not be identified in the claims data using our original algorithm.

The strengths of this study include access to medical records to review and validate possible PRCA cases identified in the claims data. All possible PRCA cases with available medical records were reviewed by a three-member AC, all members of which specialized in clinical hematology and/or oncology. In

addition, the possible and probable claims-based algorithms were developed following consultation with several clinical hematologists familiar with PRCA.

Limitations of the study include the inability to obtain medical records sufficient for review for eight of the 25 possible PRCA cases identified in the claims data, which prevented these possible PRCA cases from being evaluated by the AC. It is possible that there may have been some PRCA cases among the eight possible PRCA cases that were not adjudicated.

Conclusions

Pure red cell aplasia is a rare hematologic disease that is not easily identified in administrative claims data due in large part to its rarity, absence of an ICD-9-CM code specific to acquired PRCA, and the clinical and diagnostic characteristics of PRCA being shared by other more common conditions. The two claims-based algorithms we developed were unable to identify any PRCA cases in a cohort of more than 36,000 CHC patients. Future studies might consider validating a claims-based algorithm for PRCA in a larger cohort or utilizing the upcoming International Classification of Diseases, Tenth Revision coding system to allow development of a more specific algorithm.

List of abbreviations

AC: Adjudication Committee
CHC: Chronic hepatitis C
CI: Confidence interval
EPO: Erythropoietin
HIRD: HealthCore Integrated Research Database[SM]
ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification
PegIFN: Pegylated interferon
PPV: Positive predictive value
PRCA: Pure red cell aplasia
RBV: Ribavirin
US: United States

Additional files

> **Supplementary Table S1**
> **Supplementary Appendix**

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Authors' contributions	SCQ	JCH	DM	SA	SL	VM	CNH	NS	TCM
Research concept and design	✓	✓	✓	✓	--	✓	✓	✓	✓
Collection and/or assembly of data	✓	✓	--	--	✓	--	--	--	--
Data analysis and interpretation	✓	✓	✓	✓	✓	✓	--	--	✓
Writing the article	✓	✓	--	--	--	--	--	--	--
Critical revision of the article	✓	✓	✓	✓	✓	✓	✓	✓	✓
Final approval of article	✓	✓	✓	✓	✓	✓	✓	✓	✓
Statistical analysis	✓	✓	--	--	✓	--	--	--	--

References

1. Dessypris EN and Lipton JM. **Red Cell Aplasia**. In Greer JP, Foerster J, Rodgers GM, Paraskevas F, Glader B, Arber DA, and Means RT (**Eds.**), **Wintrobe's Clinical Hematology**. 2009; 1194-1209.

2. al-Awami Y, Sears DA, Carrum G, Udden MM, Alter BP and Conlon CL. **Pure red cell aplasia associated with hepatitis C infection**. *Am J Med Sci*. 1997; **314**:113-7. | Article | PubMed

3. Arcasoy MO, Rockey DC and Heneghan MA. **Pure red cell aplasia following pegylated interferon alpha treatment**. *Am J Med*. 2004; **117**:619-20. | Article | PubMed

4. Chang CS, Yan SL, Lin HY, Yu FL and Tsai CY. **Pure red cell aplasia caused by pegylated interferon-alpha-2a plus ribavirin in the treatment of chronic hepatitis C**. *World J Gastroenterol*. 2011; **17**:2155-8. | Article | PubMed Abstract | PubMed Full Text

5. Miura Y, Kami M, Yotsuya R, Toda N and Komatsu T. **Pure red-cell aplasia associated with pegylated interferon-alpha-2b plus ribavirin**. *Am J Hematol*. 2008; **83**:758-9. | Article | PubMed

6. Rossert J, Yue S, Smirnakis K, Mytych DT, Johnson L, Kouchakji E and Casadevall N. **Risk of pure red cell aplasia in patients with hepatitis C receiving antiviral therapy and an erythropoiesis-stimulating agent**. *Clin Gastroenterol Hepatol*. 2014; **12**:341-5. | Article | PubMed

7. Stravitz RT, Chung H, Sterling RK, Luketic VA, Sanyal AJ, Price AS, Purrington A and Shiffman ML. **Antibody-mediated pure red cell aplasia due to epoetin alfa during antiviral therapy of chronic hepatitis C**. *Am J Gastroenterol*. 2005; **100**:1415-9. | Article | PubMed

8. Tanaka N, Ishida F and Tanaka E. **Ribavirin-induced pure red-cell aplasia during treatment of chronic hepatitis C**. *N Engl J Med*. 2004; **350**:1264-5. | Article | PubMed

9. Thevenot T, Cadranel JF, Di Martino V, Pariente A, Causse X, Renou C, Hagege H, Denis J and Lunel-Fabiani F. **A national French survey on the use of growth factors as adjuvant treatment of chronic hepatitis C**. *Hepatology*. 2007; **45**:377-83. | Article | PubMed

10. Collins AJ, Li S, Adamson JW and Gilbertson DT. **Assessment of pure red cell aplasia in US dialysis patients: the limits of the Medicare data**. *Am J Kidney Dis*. 2004; **43**:464-70. | Article | PubMed

11. Rothman KJ and Boice JD: **Epidemiologic analysis with a programmable calculator**. NIH Publication 79-1649. 1979. Washington, DC, US Government Printing Office.

12. Schonholzer C, Keusch G, Nigg L, Robert D and Wauters JP. **High prevalence in Switzerland of pure red-cell aplasia due to anti-erythropoietin antibodies in chronic dialysis patients: report of five cases**. *Nephrol Dial Transplant*. 2004; **19**:2121-5. | Article | PubMed

> **Citation**:
>
> Quinlan SC, Hawes JCL, Mines D, Ahmed S, Lanes S, Mehta V, Holick CN, Santanello N and Mast TC. **Performance of an administrative claims algorithm to estimate the incidence of pure red cell aplasia in chronic hepatitis C patients**. *Epidemiol Rep*. 2015; **3**:1. http://dx.doi.org/10.7243/2054-9911-3-1