ORIGINAL RESEARCH

Positive predictive value of diagnosis coding for hemolytic anemias in the Danish National Patient Register

Dennis Lund Hansen[1]
Ulrik Malthe Overgaard[2]
Lars Pedersen[3]
Henrik Frederiksen[1,3]

[1]Department of Haematology, Odense University Hospital, Odense, [2]Department of Haematology, Herlev Hospital, Herlev, [3]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Purpose: The nationwide public health registers in Denmark provide a unique opportunity for evaluation of disease-associated morbidity if the positive predictive values (PPVs) of the primary diagnosis are known. The aim of this study was to evaluate the predictive values of hemolytic anemias registered in the Danish National Patient Register.

Patients and methods: All patients with a first-ever diagnosis of hemolytic anemia from either specialist outpatient clinic contact or inpatient admission at Odense University Hospital from January 1994 through December 2011 were considered for inclusion. Patients with mechanical reason for hemolysis such as an artificial heart valve, and patients with vitamin-B12 or folic acid deficiency were excluded.

Results: We identified 412 eligible patients: 249 with a congenital hemolytic anemia diagnosis and 163 with acquired hemolytic anemia diagnosis. In all, hemolysis was confirmed in 359 patients, yielding an overall PPV of 87.1% (95% confidence interval [CI]: 83.5%–90.2%). A diagnosis could be established in 392 patients of whom 355 patients had a hemolytic diagnosis. Diagnosis was confirmed in 197 of the 249 patients with congenital hemolytic anemia, yielding a PPV of 79.1% (95% CI: 73.5%–84.0%). Diagnosis of acquired hemolytic anemia could be confirmed in 136 of the 163 patients, resulting in a PPV of 83.4% (95% CI: 76.8%–88.8%). For hemoglobinopathy PPV was 84.1% (95% CI: 77.4%–89.4%), for hereditary spherocytosis PPV was 80.6% (95% CI: 69.5%–88.9%), and for autoimmune hemolytic anemia PPV was 78.4% (95% CI: 70.4%–85.0%).

Conclusion: The PPV of hemolytic anemias was moderately high. The PPVs were comparable in the three main categories of overall hemolysis, and congenital and acquired hemolytic anemia.

Keywords: ICD-10 coding, diagnosis validation, hemoglobinopathy, spherocytosis, epidemiology, congenital hemolytic anemia, acquired hemolytic anemia

Introduction

Hemolytic anemias and chronic hemolytic disorders are among the most prevalent congenital diseases with a major impact on health budgets around the world.[1–7] Chronic anemia causes hypoxia and extra stress to cardiac function, and it may, as well, inflict negatively on the function of other organs.[6,8] Individuals with anemia will most often suffer from headache, fatigue, dyspnea, impaired working ability, and impaired cognitive performance.[6,8–10] Anemia has been reported to worsen prognosis and increase the cost of treatment of other diseases.[6,11,12] Chronic hemolysis even without anemia has been reported to predispose to an increased risk of thrombotic disease and might predispose to infirmity and progressive morbidity.[2,13–17] For some diseases with chronic hemolysis such

Correspondence: Dennis Lund Hansen
Department of Haematology, Odense University Hospital, Sdr Boulevard 29, DK-5000, Odense C, Denmark
Tel +45 6541 2923
Email dlh@dadlnet.dk

as paroxysmal nocturnal hemoglobinuria and sickle cell anemia, an increased risk of thrombosis is well established.[14,16,18,19] In thalassemia a range of organ damage is well described, derived from hypoxia, thrombosis, and iron load.[1,2,13]

Denmark is, with its nationwide public health registers, an ideal place to evaluate morbidity associated with a disorder if the positive predictive value (PPV) of the primary diagnosis is known.[20] All hospital-assigned diagnoses are registered in the Danish National Patient Register (DNPR) but the PPV of non-vitamin-deficiency hemolysis diagnoses has not been established prior to this time in the DNPR.[21]

In order to evaluate if registry-based diagnoses can be used to study morbidity in patients with chronic hemolysis, we studied the PPV of diagnosis coding for hemolytic anemias in the DNPR.

Patients and methods
Patient cohort
We included all patients with a specialist outpatient clinic contact or inpatient admission at Odense University Hospital (OUH) registered in the DNPR with a first-ever diagnosis of hemolytic anemia from 1994 through 2012. This cohort was drawn from the DNPR and the Danish Civil Registration System (CRS). Every inhabitant of Denmark has a unique and permanent entry in the CRS – the civil registration number (CRN), allowing individual-level linkage across all public registers.[20,22] The DNPR contains information on virtually all discharges from public hospitals since 1977 and on outpatient clinic visits since 1995.[20,22] Denmark has very few private hospitals, none of which are involved in the diagnostic process of patients with hemolytic disorders. The data registered in the DNPR include the CRN dates of hospital outpatient visits, dates of hospital discharge, and up to 20 diagnoses coded by physicians according to the WHO International Classification of Diseases Tenth Revision (ICD-10) since 1994. Each hospital contact is coded with the primary (first-listed) diagnosis that led to the hospital contact as well as secondary diagnoses, which the physician considers important in the care of the patient. The inclusion period was from January 1, 1994 to December 31, 2011. OUH has 1.1 million patient contacts per year and serves a population of 1.2 million inhabitants.

All patients with a diagnosis coding of a hemolytic disorder according to the ICD-10 in the DNPR (Tables S1 and S2) who had been admitted for inpatient care or treated on an outpatient basis at OUH were considered for inclusion, both when the hemolysis diagnosis was the first-listed diagnosis or any of the subsequent listed diagnoses. Patients with a new first-ever hemolysis diagnosis coding comprising

erythrocyte membrane diseases, hemoglobinopathies, enzyme deficiencies, and acquired hemolytic diseases were included. Included diagnoses and counts can be seen in Table 1. As a first step, we excluded patients with a hemolytic disorder appearing before January 1, 1994 by means of the ICD-8 diagnosis coding (Tables S1 and S2). Since our goal was to study diagnoses with an erythrocyte-specific cause of hemolysis, patients with a mechanical reason for hemolysis such as an artificial heart valve and patients with either vitamin-B12 or folic acid deficiency were also excluded. The ICD-8 and ICD-10 coding used to define our target population is presented in Tables S1 and S2.

Medical record review
The cohort of patients with a hemolysis diagnosis code was drawn from the DNPR. For the review process, it was divided into two groups based on the registered diagnoses. Patient files with a diagnosis code in the DNPR suggestive of a hemoglobinopathy were reviewed by Ulrik Malthe Overgaard (UMO), the details of which are presented in the following section, while patients assigned diagnosis codes of all other hemolytic disorders were reviewed by Dennis Lund Hansen (DLH). This review included all the medical files and all available laboratory and pathology results. The medical files of the patients who after the initial review still had an unsolved hemolysis diagnosis were reevaluated by DLH and Henrik Frederiksen (HF) together. After this process, a final decision to either accept the original hemolysis diagnosis, or assign a new hemolysis diagnosis, or judge the available material as insufficient for a causal hemolytic diagnosis was made. The workflow and results are illustrated in the flowchart (Figure 1).

Prior to the review of the medical records, an entry module was created in Epidata 3.1 containing ID number, date of birth, date of diagnosis coding for hemolysis, diagnosis code, up to 15 signs and symptoms of hemolysis, ethnicity, gene and flow cytometry test results, a post-review best fitting diagnosis, and the date of the post-review diagnosis, defined as the date where the last test needed to make a causal diagnosis was available.[23]

Review of biochemical data
Laboratory diagnostics of hemoglobinopathies in Denmark is centralized at the Centre for Haemoglobinopathies, Department of Haematology, Herlev Hospital, University of Copenhagen, which performs almost all hemoglobinopathy analyses. The department has filed all laboratory results since May 1995. The database contains information of hemoglobin electrophoresis (currently performed as high-pressure liquid chromatography) and sequencing of alpha- and beta-globin

Table 1 Diagnosis overview

ICD-10	Extracted diagnoses (%)	Herlev database	Primary review	Secondary review	Total verified
D550 – (G6PD) deficiency	–	0	2	0	2
D552 – (PK) deficiency	–	0	1	0	1
D56 – Thalassemia	130 (31.55)	8	20	8	36
D560 – Alpha-thalassemia	–	19	4	0	23
D561 – Beta-thalassemia	–	41	1	1	43
D563 – Thalassemia trait	–	0	5	0	5
D57 – Sickle cell anemia	25 (6.07)	0	3	1	4
D572 – Double heterozygous sickling disorders	–	2	0	0	2
D572a – Hb-SC disease	–	4	1	0	5
D572d – Sickle cell thalassemia	–	5	1	0	6
D573 – Heterozygous hemoglobin S	–	6	0	0	6
D580 – Hereditary spherocytosis (HS)	72 (17.48)	–	–	–	–
D580 – HS, cat 0	–	0	36	8	44
D580 – HS, cat 1a	–	0	0	3	3
D580 – HS, cat 1b	–	0	0	2	2
D580 – HS, cat 1c	–	0	0	9	9
D580 – HS, cat 2	–	0	0	7	7
D581 – Hereditary elliptocytosis	1 (0.24)	0	1	0	1
D582 – Other hemoglobinopathies	2 (0.49)	0	2	1	3
D582e – Hb-E disease	–	3	0	1	4
D589 – Hereditary hemolytic anemia, unspecified	19 (4.61)	0	0	1	1
D590 – Drug-induced AIHA	5 (1.21)	0	1	1	2
D591 – AIHA, endogenous	129 (31.31)	0	80	15	95
D591a – Cold agglutinin	5 (1.21)	0	16	3	19
D592 – Drug-induced non-AIHA	4 (0.97)	0	1	1	2
D594 – Other non-AIHA	14 (3.40)	0	7	1	8
D594c – Toxic hemolytic anemia	1 (0.24)	0	0	0	0
D595 – Paroxysmal nocturnal hemoglobinuria	5 (1.21)	0	4	0	4
D599 – Acquired hemolytic anemia, unspecified	–	0	9	9	18
Uncertain diagnosis	–	0	0	20	20
No hemolytic disorder	–	0	35	2	37
Total	412	88	230	94	412

Note: The lowercase suffix on D572, D582, D591, and D594 is a specific part of the Danish implementation of ICD-10.
Abbreviations: ICD-10, International Classification of Diseases Tenth Revision; G6PD, glucose-6-phosphate dehydrogenase; PK, pyruvate kinase; cat, category; AIHA, autoimmune hemolytic anemia; Hb-SC, sickle cell–haemoglobin C; Hb-E, hemoglobin E.

genes by polymerase chain reaction (PCR) methods as previously described.[24] All patients with a DNPR registration with a diagnosis of hemoglobinopathies were extracted from the complete DNPR cohort and were reviewed by UMO using this database. Patients with a hemoglobinopathy diagnosis code who were not registered in this database or did not have a confirmed hemoglobinopathy diagnosis proceeded to ordinary review of the medical files as described earlier in (Medical record review) section.

Criteria for diagnosis

In all patients, the biochemical verification of hemolysis was assessed first, with a subsequent validation of specific hemolytic diagnoses. Anemia was defined as hemoglobin concentration below 12.88 g/dL for men and 11.27 g/dL for women. For patients <16 years, anemia was defined when

hemoglobin was below the age-adjusted reference intervals. Hemolysis was defined as an increased count of reticulocytes combined with biochemical sign of erythrocyte destruction. The latter was defined when two of the following criteria were present: 1) increased lactate dehydrogenase above upper reference value, 2) decreased haptoglobin, 3) elevated free hemoglobin, or 4) hyperbilirubinemia.[17,25–27] It was a prerequisite that there was not a better explanation to the signs of hemolysis eg, liver disease.[17,25–27] The diagnostic criteria for each of the causal hemolytic disorders were according to the published criteria and are presented in Tables 2 and 3.

Hereditary spherocytosis

Validating the medical files for patients who were coded with, or where a review of the medical files suggested, hereditary spherocytosis (HS) was a special challenge.
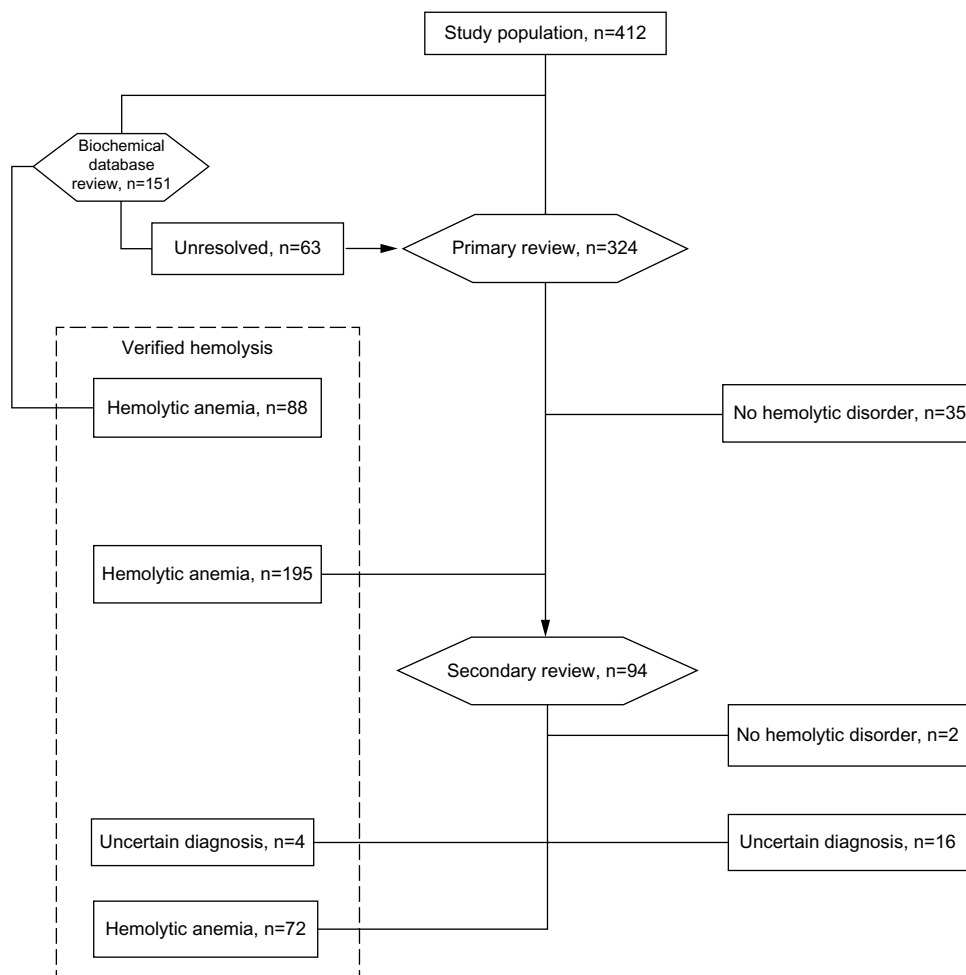
Figure 1 Numbers of patients with extracted and verified diagnosis of hemolytic anemia.
Notes: The hexagons represent reviews. The patients from the cohort with a prior hemoglobinopathy diagnosis were first reviewed in the database for former positive tests for hemoglobinoathy at Herlev Hospital. Unresolved cases proceeded along with the nonhemoglobinopathy patients to the primary and if necessary secondary review.

The disease has a high prevalence in Northern Europe of 2:10,000 to 3:10,000, and up to 1% of the population if milder or subclinical cases are included.[28,29] An HS diagnosis is therefore often made in local hospitals without specialized hematological expertise. Our review was restricted to the medical record from our university hospital including the pathological record. Not all older local medical records

were available. The review of HS diagnostic criteria (Table 4) was made in accordance with literature.[30–33] The diagnosis was based on the presence of hemolysis and additional criteria. Typically, patients were diagnosed based on family history of hemolysis, negative direct antiglobulin test (DAT), and increased osmotic fragility or positive eosin-5′-maleimide (EMA) binding test. The negative DAT was

Table 2 Definition of congenital hemolytic diseases

Disease	ICD-10	Core criteria
Congenital		
Thalassemia	D56	Positive hemoglobin electrophoresis/HPLC and/or gene test
Sickle cell anemia	D57	Positive hemoglobin electrophoresis/HPLC and confirmatory gene test
Hereditary spherocytosis	D588	Hemolysis + reticulocytosis + family history of hemolysis + negative antiglobulin test + increased osmotic fragility or positive EMA test.
Hereditary elliptocytosis	D581	Increased osmotic fragility, elliptocytes, and family history
Other hemoglobinopathies	D582	Positive hemoglobin electrophoresis/HPLC or gene test
Hereditary hemolytic anemia, unspecified	D589	Hemolysis + anemia + family history or congenital lifelong hemolysis and no better explanation available

Abbreviations: ICD-10, International Classification of Diseases Tenth Revision; HPLC, high-pressure liquid chromatography; EMA, eosin-5′-maleimide.

Table 3 Definition of acquired hemolytic diseases

Disease	ICD-10	Core criteria	Supplementary criteria
Acquired			
Drug-induced autoimmune and nonautoimmune hemolytic anemia	D590 + D592	Onset correlated with new medicine, and vanished after discontinuation and no prior history of hemolysis; preferably provocation test	If D590, preferably presence of antibodies
Autoimmune hemolytic anemia	D591	Presence of erythrocyte antigen-specific autoantibodies	Concomitant sensitization with C3d
Chronic cold hemagglutinin disease	D591a	Presence of cold-type autoantibodies	Sensitization only with C3d
Other nonautoimmune hemolytic anemias	D594	Presence of schistocytes and relevant cause (eg, sarcoma) and symptoms abate after removal of potential cause	
Toxic hemolytic anemia	D594a	Presence of potential hemolytic toxin and symptoms abate after removal of potential cause	
Paroxysmal nocturnal hemoglobinuria	D595	Presence of CD59- and CD55-deficient erythrocytes and granulocytes/monocytes by flow cytometry of peripheral blood	
Acquired hemolytic anemia, unspecified	D599	Hemolysis + anemia + known onset and no better explanation available	No family history

Note: The lowercase suffix on D591 and D594 is a specific part of the Danish implementation of ICD-10.
Abbreviation: ICD-10, International Classification of Diseases Tenth Revision.

considered obligatory to the diagnosis, but if the patient was of Scandinavian ethnicity (and thereby "high risk"), negative family history or lack of increased osmotic fragility was accepted if other characteristics were present such as spherocytes and splenomegaly combined with normalization of hemolysis after splenectomy (Table 4). DAT is part of the routine workup in patients with hemolysis, but negative DAT results are not always reported. In the HS category 2, we therefore assume that the DAT was done, and found to be negative, since no treatment applicable to autoimmune hemolytic anemia (AIHA) was used. Also, when the anemia is mild and there are no factors predisposing to AIHA such as malignancies and autoimmune diseases, the probability of AIHA is much smaller compared to the probability of HS.[26,28,32,34–36]

Ethics

The study was approved by the Danish Health and Medicines Authority in accordance with current Danish law (Reference number: 3-3013-307/1). In compliance with Danish law, the project was approved by the Danish Data Protection Agency through the Odense university hospital (Ref no 2008-58-0035). Because this study was based on data extracted from registries, it was exempt from human subjects review, and members of the study population did not have to provide informed consent.

Statistical analysis

We used the original diagnosis based on the DNPR-registered ICD-10 diagnosis code and the subsequent review diagnosis to compute PPVs for categories of overall hemolysis

Table 4 Categories of hereditary spherocytosis

Category	Probability of hereditary spherocytosis	Criteria
0	Definite	Hemolysis, reticulocytosis, family history of hemolysis (preferably exact diagnosis in first-degree relative), negative antiglobulin test, and increased osmotic fragility or positive EMA test
1	Probable	A. Hemolysis, reticulocytosis, negative antiglobulin test, spherocytes, Scandinavian ethnicity and no better explanation, or B. Highly increased osmotic fragility, reticulocytosis, negative antiglobulin test, Scandinavian ethnicity and no better explanation, or C. Hemolysis splenomegaly and splenectomy, negative antiglobulin test, spherocytosis, Scandinavian ethnicity and no better explanation
2	Likely	Cases fitting category 0, or 1A–C but missing data on one of the following: antiglobulin test, family history, osmotic fragility test, or peripheral smear
Exclusion	Unlikely	Other well-defined hemolytic anemia or negative EMA test (when performed)

Abbreviation: EMA, eosin-5′-maleimide.

congenital hemolysis, acquired hemolysis as well as for the specific hemolytic disorders, which included more than 50 patients. Signs of coding errors were noted in the surgical departments where the hemolysis diagnosis was one character wrong, eg, C569: "Malignant neoplasm of ovary" and G560: "Carpal tunnel syndrome", both mistyped as D56 alpha thalassemia. We therefore also repeated analyses excluding patients who only once were given a hemolysis diagnosis code and solely at a surgical department. For the group of patients with an "overall hemolysis" diagnosis, patients were included as a verified hemolytic disorder if the previously mentioned biochemical criteria for hemolysis were present, regardless of whether the specific causal hemolytic diagnosis code was correct. For the subgroups of specific hemolytic disorders, patients were assigned a causal hemolytic diagnosis if their symptoms and biochemical results matched one of the diagnostic criteria in Tables 2–4. The process was unblinded as the primary extraction diagnosis and the medical file were available.

PPVs and 95% confidence intervals (95% CIs) were calculated as the fraction of confirmed hemolytic diagnoses of the original DNPR-registered ICD-10 diagnosis coded cases. Medical files where the available data were insufficient to judge whether hemolysis and/or a causal hemolytic diagnosis was present were counted as not confirmed. The diagnosis of thalassemia and sickle cell anemia was also accepted if the true diagnosis was sickle cell (beta-)thalassemia (D572, denoted D572d in the Danish implementation of ICD-10). For patients with AIHA, patients with warm type, cold type, or mixed type autoantibodies were grouped together in the calculation of sensitivity and PPV (Table 5). The diagnosis of HS was accepted as confirmed when the post-review score was 0 or 1 (Table 4). The second category was included in some auxiliary calculations of PPV and age-stratified HS.

All statistics computing has been done in STATA version 10.1 using the diagt command (package: sbe36 and sbe36_1) which reports PPV and calculates exact binomial confidence intervals.[37]

Results
Hemolysis
We identified 412 patients with a first-ever hemolysis diagnosis code at OUH during 1994–2011. Medical files from the time of the hemolytic diagnosis could be identified for all patients. Out of the 412 patients, there was no secure evidence of hemolysis in 53 patients and hemolysis was confirmed in 359, yielding a PPV of 87.1% for hemolysis (83.6–90.2).

Thirty-seven patients were excluded as there was no evidence of hemolysis at any time in the record. Among an additional 20 patients, there was insufficient information in the medical file to make an exact diagnosis, however four of whom had sufficient information to confirm the presence of hemolysis. Thirteen of the 20 inconclusive files were from before 1999 when electronic medical records commenced. Some older paper medical records could not be retrieved. The overall proportion of patients with verified hemolysis was constant in the two time periods (1994–2002 and 2003–2011). In both the time periods, the proportion with verified hemolysis was highest in the group below 18 years.

Diagnosis coding
Of the 359 patients who had biochemical verification of hemolysis, the initial diagnosis code was confirmed in 294 patients, yielding an overall PPV of 71.4% (95% CI: 66.7%–75.6%). If patients with clear signs of diagnosis code typing error from surgical departments were disregarded, the PPV increased to 74.2% (95% CI: 69.6%–78.5%), and if patients with a category 2 diagnosis of HS were included, an overall PPV of 76.0% (95% CI: 71.5%–80.1%) was achieved. The PPV values were identical for both sexes, but the PPV was much higher in patients below 19 years (97.9%) (95% CI: 93.9%–99.6%) compared to patients $\geq$19 years (81.6%) (95% CI: 76.4%–86.0%) (Table 6). In 20 patients with hemolysis, not all files necessary for a causal diagnosis could be found and were therefore classified as insufficient for a causal diagnosis. If these patients were disregarded, the PPV in the age group above 18 years was computed to be 92.0% (95% CI: 87.8%–95.1%).

Congenital hemolytic anemia – all subtypes
We defined 249 patients as having congenital hemolytic anemia due to their ICD-10 diagnosis codes (Tables S1 and S2). Hemolysis was confirmed in 216 (86.7%) and 197 patients could after review be confirmed as having congenital hemolytic anemia. The 13 patients for whom the diagnosis was unconfirmed had anemia due to hematological malignancy (n=8) or lacked information (n=5). The overall predictive value of a diagnosis of congenital hemolytic anemia was 79.1% (95% CI: 73.5%–84.0%). If patients with clear signs of typing error during diagnosis from surgical departments were disregarded (ten patients), the PPV increased to 83.8% (95% CI: 78.5–88.3). If category 2 of the HS diagnosis was included and the patients with clear signs of a mistyped one-time diagnosis from surgical departments were disregarded,

Table 5 Table of results – PPV of specific diagnoses

ICD diagnosis	Extraction	Review	Agreement	Disagreement	PPV
Congenital hemolysis					
Congenital hemolysis	**249**	**200**	**197**	**52**	**79.1 (73.5–84.0)**
Congenital hemolysis, nonsurgical	**235**	**200**	**197**	**38**	**83.8 (78.5–88.3)**
D55x Enzyme deficiencies					
D56x Thalassemia					
D57x Sickle cell anemia					
D582x Other hemoglobinopathies					
D580 Hereditary spherocytosis (0–1)					
D581a Hereditary elliptocytosis					
D582x Other hemoglobinopathies					
D588 Other hereditary hemolytic anemia					
D589 Hereditary hemolytic anemia					
Hemoglobinopathies	**157**	**137**	**132**	**25**	**84.1 (77.4–89.4)**
Hemoglobinopathies, nonsurgical	**144**	**137**	**132**	**12**	**91.7 (85.9–95.6)**
D56x Thalassemia					
D57x Sickle cell anemia					
D582x Other hemoglobinopathies					
D56 Thalassemia	**130**	**113**	**107**	**23**	**82.3 (74.7–88.4)**
D56 Thalassemia, nonsurgical	**122**	**113**	**107**	**15**	**87.7 (80.5–93.0)**
D56 Thalassemia		36			
D560 Alpha-thalassemia		23			
D561 Beta-thalassemia		43			
D563 Thalassemia minor		5			
D572d Sickle cell thalassemia		6			
D57 Sickle cell anemia	**25**	**23**	**18**	**7**	
D57 Sickle cell anemia		4			
D572 Double heterozygous sickling disorders		2			
D572a Hb-SC disease		5			
D572d Sickle cell thalassemia		6			
D573 Heterozygous hemoglobin S		6			
D580 Hereditary spherocytosis (0–1)	**72**	**58**	**51**	**21**	**70.8 (58.9–81.0)**
D580 Hereditary spherocytosis (0–2)	**72**	**65**	**58**	**14**	**80.6 (69.5–88.9)**
D581a Hereditary elliptocytosis	1	1	1	0	
D582 Other hemoglobinopathies	2	7	1	1	
D582 Other hemoglobinopathies		3			
D582e Hb-E disease		4			
D588 Other hereditary hemolytic anemia	2	0	0	2	
D589 Hereditary hemolytic anemia, unspecified	17	1	0	17	
Acquired					
Acquired hemolysis	**163**	**148**	**136**	**27**	**83.4 (76.8–88.8)**
D590 AIHA, drug-induced		2			
D591x Other AIHA, non-drug-induced		13			
D592 Hemolytic nonautoimmune anemia		2			
D594x Other hemolytic nonautoimmune		7			
D595 Paroxysmal nocturnal hemoglobinuria		4			
D598 Other acquired hemolytic anemia		5			
D599 Acquired hemolytic anemia, unspecified		3			
D590 AIHA, drug-induced	5	2	0	5	
D591 AIHA, not drug-induced	**134**	**114**	**105**	**29**	**78.4 (70.4–85.0)**
D591a Cold-agglutinin disease		19			
D592 Hemolytic nonautoimmune anemia	4	2	2	2	
D594 Other hemolytic nonautoimmune	14	8	5	9	
D594c Toxic hemolytic anemia	1	0	0	1	
D595 Paroxysmal nocturnal hemoglobinuria	5	4	4	1	

Notes: The lowercase suffix on D572, D582, D591, and D594 refers to the specific Danish implementation of ICD-10 using lowercase letters to denote subtypes. Bold figures refer to classification lines (congital versus acquired) and major subgroups (congenital all forms, all forms of hemoglobinopathies thalassemia, sickle cell anaemia, and spherocytosis, all forms of acquired hemolysis and AIHA [not drug-induced]).
Abbreviations: ICD-10, International Classification of Diseases Tenth Revision; AIHA, autoimmune hemolytic anemia; PPV, positive predictive value; Hb-SC, sickle cell–hemoglobin C; Hb-E, hemoglobin E.

Table 6 Table of results – overall hemolysis

ICD diagnosis	Extraction	Review	Agreement	Disagreement	PPV
1994–2011					
Overall hemolysis	**412**	**359**	**359**	**53**	**87.1 (83.5–90.2)**
Male (all ages)	178	158	158	20	88.8 (83.2–93.0)
Female (all ages)	234	201	201	33	85.9 (80.8–90.1)
Male, female 0–18 (Y)	141	138	138	3	**97.9 (93.9–99.6)**
Male, female 18+ (Y)	271	221	221	50	**81.6 (76.4–86.0)**
Overall hemolysis, nonsurgical	**396**	**359**	**359**	**37**	**90.7 (87.4–93.3)**
Male, nonsurgical (all ages)	169	158	158	11	93.5 (88.7–96.7)
Female, nonsurgical (all ages)	227	201	201	26	88.6(83.7–92.4)
Male, female 0–18 (Y), nonsurgical	141	138	138	3	97.9 (93.9–99.6)
Male, female 18+ (Y), nonsurgical	255	221	221	34	86.7 (81.9–90.6)
1994–2002					
Overall hemolysis	**145**	**124**	**124**	**21**	**85.5 (78.7–90.8)**
Male (all ages)	69	63	63	6	91.3 (82.0–96.7)
Female (all ages)	76	61	61	12	80.3 (69.5–88.5)
Male, female 0–18 (Y)	59	56	56	3	**94.9 (85.9–98.9)**
Male, female 18+ (Y)	86	68	68	18	**79.1 (69.0–87.1)**
Overall hemolysis nonsurgical	**139**	**124**	**124**	**15**	**89.2 (82.8–93.8)**
Male, nonsurgical (all ages)	66	63	63	3	95.4 (87.3–99.1)
Female, nonsurgical (all ages)	73	61	61	12	83.6 (73.1–91.2)
Male, female 0–18 (Y), nonsurgical	59	56	56	3	94.9 (85.9–98.9)
Male, female 18+ (Y), nonsurgical	80	68	68	12	85.0 (75.3–92.0)
2003–2011					
Overall hemolysis	**267**	**235**	**235**	**32**	**88.0 (83.5–91.7)**
Male (all ages)	109	95	95	14	87.2 (79.4–92.8)
Female (all ages)	158	140	140	18	88.6 (82.6–93.1)
Male, female 0–18 (Y)	82	82	82	0	**100 (95.6–100)**
Male, female 18+ (Y)	185	153	153	32	**82.7 (76.5–87.9)**
Overall hemolysis nonsurgical	**257**	**235**	**235**	**22**	**91.4 (87.3–94.6)**
Male, nonsurgical (all ages)	103	95	95	8	92.2 (85.3–96.6)
Female, nonsurgical (all ages)	154	140	140	14	90.9 (85.2–94.9)
Male, female 0–18 (Y), nonsurgical	82	82	82	0	100 (95.6–100)
Male, female 18+ (Y), nonsurgical	175	153	153	22	87.4 (81.6–92.0)

Notes: Values in parenthesis denote 95% CI. Bold figures refer to subgroups according to time period and diagnosis made in non-surgical departments.
Abbreviations: ICD-10, International Classification of Diseases Tenth Revision; PPV, positive predictive value; CI, confidence interval; Y, year.

204 of the initial 249 patients were confirmed after review as having congenital hemolytic anemia, yielding a PPV of 86.8% (95% CI: 81.8%–90.9%). The predictive value of all hemoglobinopathies was 84.1% (95% CI: 77.4%–89.4%) and for thalassemia, in particular, it was 82.3% (95% CI: 74.7–88.4). Further results are listed in Table 5.

Hereditary spherocytosis

Of the original 412 patients, 72 were coded with a diagnosis of HS. In general, a more severe disease with symptomatic anemia triggered a more complete diagnostic approach and a reliable diagnosis (data not shown). After review, ten cases did not comply with HS. These patients had iron-deficiency anemia due to bleeding (two patients), anemia due to end-stage liver disease (two patients), showed no signs of hemolysis or had missing information in the medical files (three patients), anemia due to inflammation (two

patients), and one patient had AIHA. With the most restrictive approach with only categories 0 to 1 accepted as HS diagnoses, the PPV was 70.8% (95% CI: 58.9%–81.0%). When the categories 0 to 2 were accepted, the PPV increased to 80.6% (95% CI: 69.5%–88.9%). An additional two patients fulfilled all the criteria for HS, but since their EMA test was negative, we categorized these patients as HS negative, due to the very high reliability of this test.[30] In two patients, medical files were very suggestive of an HS diagnosis. However, this could not be finally confirmed due to lack of sufficient information in one and alcohol abuse and liver disease obscuring the definite hemolytic diagnosis in the other. If these two suspected cases are accepted as HS, a PPV of 83.3% (95% CI: 72.7%–91.1%) was obtained.

The diagnosis coding of HS was most reliable for younger people since the PPV among patients ≤18 years was 84.6% (95% CI: 69.5%–94.1%) for categories 0 to 1 HS diagnosis.

For patients above 18 years, the PPV was 54.6% (95% CI: 36.4%–71.9%) for categories 0 to 1. If category 2 (likely HS) is included as verified HS diagnoses, the PPVs increase to 92.3% (95% CI: 79.1%–98.4%) for patients ≤18 years and 66.7% (95% CI: 48.2%–82.0%) for patients >18 years. Six patients who after review were diagnosed with HS were a priori misclassified as acquired hemolytic anemia (n=2) or "other hereditary hemolytic anemia" (n=4).

Acquired hemolytic anemia

In 163 patients (37.9%), the ICD-10 diagnosis code was acquired hemolytic anemia (Tables S1 and S2). After the review of medical files, there was evidence of hemolysis in 143 (84.7%) patients and the diagnosis was confirmed in 136 patients, yielding an overall predictive value of 83.4% for acquired hemolysis (95% CI: 76.8%–88.8%). The 27 non-confirmed cases consisted of eleven patients without any sign of a hemolytic disorder such as iron deficiency anemia, 13 patients with insufficient medical information for a specific diagnosis, and three patients misclassified as HS. Of the 13 with insufficient information, four had biochemical signs of hemolysis. Among the 111 patients (82%) who had confirmed acquired hemolytic anemia, the causal diagnosis was non-drug-induced AIHA, of whom 19 patients (17%) could be ascribed to cold agglutinin syndrome, 48 patients (43%) had warm-antibody-mediated AIHA, and the remaining 44 cases (40%) were either mixed type or were lacking information on type. One patient had no detectable antibodies or complement binding, but had symptoms, biochemistry, and treatment response identical to AIHA and was classified as AIHA without detectable antibodies (DAT-negative AIHA).[36,38] The PPV for non-drug-induced AIHA was 78.4% (95% CI: 70.4%–85.0%). PPV of other subtypes was not calculated due to low prevalence.

Discussion

The overall PPV of a hemolysis diagnosis varied from ~80% to 98% for some subtypes and categories. The lowest PPV was found in the older age group. This may be due to the fact that all the medical files could not be retrieved for patients who were deceased. Although we excluded patients with an older diagnosis code of hemolytic anemia, our search could not exclude patients with a hemolytic anemia diagnosed prior to the establishment of the DNPR in 1973 if the diagnosis coding was reassigned after 1994. Therefore, our data included at least 6% among the older population where the diagnosis code of hemolytic anemia was not a conclusion of a recent diagnostic process but merely referred information of previous

medical files, and where the original diagnostic workup and result were no longer available. This is probably the main reason for the lower PPV (81.6%) in the older age group, since PPV among older patients with current available medical files was nearly as high as in the younger age groups; 92.0% and 97.9%, respectively. As the prevalence of all hemolytic disorders is unknown, we cannot evaluate the coverage of our data, and we do therefore not know the undiagnosed proportion of patients with hemolytic disorder. The overall PPV of 87.1% for hemolysis was comparable to the proportion of documented hemolysis in the group of congenital hemolysis (86.7%) and the group of acquired hemolysis (84.7%). In the combined disease specific groups with congenital hemolytic anemia and acquired hemolytic anemia the PPVs were slightly lower. This suggest that ongoing hemolysis will be diagnosed as such, but not always be given the correct causal disease or diagnosis coding.

Congenital hemolytic anemia

The prevalence of hemolysis within the group of a priori labeled congenital hemolytic anemia was about 90%, equal to overall prevalence of hemolysis in the entire cohort, supporting the assumption that a diagnosis of hemolysis is not assigned without biochemical evidence of hemolysis. All the 16 patients for whom the correct diagnosis strongly suggested a mistyping of the diagnosis code appeared only once in the DNPR, and all from the surgical departments which would usually not be involved in the diagnostic process of a hemolytic anemia.

The high PPV of the specific diagnosis reflects that the majority of these patients had a hemoglobinopathy often with anemia since childhood, thereby demanding an extensive workup to establish a causal diagnosis, and a high alertness of the clinician, due to prevailing family history. Similarly, 18 of the 20 cases with insufficient information to make a definite review diagnosis were seen among patients above 18 years – where the medical file with the original diagnostic workup could not be retrieved.

Hereditary spherocytosis

Among the patients who were finally defined as cases of HS, 10% were a priori misclassified as other subtypes of hemolytic anemia. In general, a more severe symptomatic anemia triggered a more complete diagnostic approach, and it is therefore likely that the probability of a correct HS diagnosis is increases with the severity of the disease. The problem of incomplete diagnostic approaches in HS triggered a division of patients into different categories. The fourth category was

established on the assumption that one of the four symptoms (increased osmotic fragility, family history, spherocytosis, and negative DAT) could be missing in typical cases, without jeopardizing the definite HS diagnosis. Increased osmotic fragility is not present in all cases of HS, and this information could be missing, if everything else fitted.[30] Spherocytes are reported to be present in nearly all patients, but are regarded as an unspecific sign by the pathologists and are therefore not always reported in the description of the blood smear.[29,30,34] A positive family history for hemolysis is only present in 80% of the HS patients.[29–33]

The most critical aberration from the definitions was a missing DAT, since the most important differential diagnosis is AIHA due to the presence of spherocytosis in both diseases.[26,32,33,35,36,38]

Acquired hemolytic anemia

Among the patients with a diagnosis code of an acquired hemolytic disorder, 84% were confirmed and 4% had ongoing hemolysis without a definite causal diagnosis. The remaining 12% did not have any signs of a hemolytic disorder. The resulting PPV is comparable to the overall PPV of hemolysis and makes this subgroup a reliable entity in the DNPR. This supports our impression that a diagnosis code of hemolysis is not assigned without biochemical evidence. The high PPV of an AIHA diagnosis can probably be ascribed to the relatively simple diagnostic procedure based on hemolysis and a positive DAT. Approximately, 3%–5% of AIHA patients have been reported to have a negative DAT and this small fraction is likely to be either misclassified or underreported in registries due to this negative test result.[35,36,38]

Strengths and limitations

Our study was based on a review of a large number of patients with hemolytic disorders and with more than 95% completeness of retrieval of medical files. It is a limitation that it is based on patients from one university hospital. In spite of this, a large number of patients had begun the diagnostic process or were already diagnosed at a primary hospital and our study thus reflects the accuracy of the diagnosis in the health care region indirectly. There is a risk, however, that primarily the severely affected and more thoroughly examined patients are referred to the university hospital, and the proportion of misdiagnosed patients could be higher in the entire population. As many of the laboratory tests are centralized at Herlev Hospital, the high PPVs for hemoglobinopathy should account for the entire country. The study relies on both the doctors' ability to make the right diagnosis

and the registration of all Danish patients in the DNPR. If the results are accepted as a proxy for the doctors' ability in making the correct diagnosis, then the results should be at least partly generalizable to comparable hospitals in other countries, but the coverage may vary with different types of health care systems.

Our review process was rigid; all records on patients who did not have a positive electrophoresis or genetic test in the database at Herlev Hospital were read by the same person. All patients with non-classical presentations were evaluated by both DLH and HF, to ensure consistent decision making.

Our study design prohibited inclusion of patients with unregistered hemolysis, and it was not feasible to estimate the prevalence of hemolytic disorders among general patients at OUH. Therefore, the sensitivity and the negative predictive values of the hemolysis diagnosis could not be assessed.

Apart from the most restricted definition of HS where the PPV was 70.8%, all PPVs (overall, congenital, acquired, and larger subgroups) varied from 78.4% to 84.1%. The predictive values are comparable to literature values for PPV in, eg, venous thromboembolism in adults (75%–85%), Monoclonal gammopathy of undetermined significance" (82.3%), and shock (86.1%).[39–41] These findings support the possibility of future studies, but considerations on the diagnosis validity are needed in some subgroups.

There was evidence of typing errors, a problem also seen in other similar works, and a marked increase in PPV when cases with diagnoses from surgical departments were omitted.[40] In future works, this approach should be considered as well as the possible inclusion of some nationwide laboratory records to rule out patients with normal blood tests.

Conclusion

The overall PPV of hemolysis in the cohort was 87.1%, higher in patients below 19 years of age and with a slight improvement over time. The PPVs from subgroups of hemolytic anemias range from 78.4% to 84.1% The PPV was comparable in the three main categories: overall hemolysis, congenital, and acquired hemolytic anemia. The difference in PPV between patients aged >18 and ≤18 years was mainly due to the lack of information in the older medical files and that severe anemia among children triggered a more extensive diagnostic workup.

The diagnosis of hemolytic disorders can be used for further register-based studies, with an acceptable reliability in both older and younger groups. The diagnosis of HS can be used with some precautions given the different definitions of the diagnosis.

Author contributions

All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Cappellini M-D, Cohen A, Eleftheriou A, Piga A, Porter J, Taher A. *Guidelines for the Clinical Management of Thalassaemia*. 2nd ed. Nicosia, Cyprus: Thalassaemia International Federation (TIF); 2008.
2. Taher A, Vichinsky E, Musallam K, Cappellini MD, Viprakasit V. *Guidelines for the Management of Non Transfusion Dependent Thalassaemia (NTDT)*. Nicosia, Cyprus: Thalassaemia International Federation (TIF); 2013.
3. Inati-Khoriaty A. *Sicle Cell Disease*. Nicosia, Cyprus: Thalassaemia International Federation (TIF); 2008.
4. Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anaemia in children under five, 2010–2050: modelling based on demographics, excess mortality, and interventions. *PLoS Med*. 2003;10:1–14.
5. Weatherall D. The inherited disorder of haemoglobin: an increasingly neglected global health burden. *Indian J Med Res*. 2011;134: 493–497.
6. Smith RE. The clinical and economic burden of anemia. *Am J Manag Care*. 2010;16(Suppl Issues):S59–S66.
7. Amato A, Grisanti P, Mastropietro F, et al. Epidemiology and screening of sickle cell anemia in the Mediterranean area and in developing countries. *Ig Sanita Pubbl*. 2014;70(1):41–52.
8. Schaffalitzky de Muckadell OB, Haunsø S, Vilstrup H. *Medicinsk kompendium* [Medical Compendium]. 17th ed. Kbh.: Nyt Nordisk Forlag Arnold Busck; 2009.
9. Lozoff B, Beard J, Connor J, Felt B, Georgieff M, Schallert T. Long-lasting neural and behavioral effects of iron deficiency in infancy. *Nutr Rev*. 2006;64(5 Pt 2):S34–S91.
10. Andro M, Le Squere P, Estivin S, Gentric A. Anaemia and cognitive performances in the elderly: a systematic review. *Eur J Neurol Off J Eur Fed Neurol Soc*. 2013;20(9):1234–1240.
11. Clevenger B, Richards T. Pre-operative anaemia. *Anaesthesia*. 2015;70(Suppl 1):20–28.
12. Kansagara D, Dyer E, Englander H, Fu R, Freeman M, Kagen D. Treatment of anemia in patients with heart disease: a systematic review. *Ann Intern Med*. 2013;159(11):746–757.
13. Musallam KM, Rivella S, Vichinsky E, Rachmilewitz E. Non-transfusion-dependent thalassemias. *Haematologica*. 2013;98:833–844.
14. Hill A, Kelly RJ, Hillmen P. Thrombosis in paroxysmal nocturnal hemoglobinuria. *Blood*. 2013;121:4985–4996.
15. Rother RP, Bell L, Hillmen P, Gladwin MT. The clinical sequelae of intravascular hemolysis and extracellular plasma hemoglobin: a novel mechanism of human disease. *JAMA*. 2005;293(13):1653–1662.
16. De FL, Cappellini M-D, Olivieri O. Thrombosis and sickle cell disease. *Semin Thromb Hemost*. 2011;37:226–236.
17. L'Acqua C, Hod E. New perspectives on the thrombotic complications of haemolysis. *Br J Haematol*. 2015;168(2):175–185.
18. Van Bijnen ST, Van Heerde WL, Muus P. Mechanisms and clinical implications of thrombosis in paroxysmal nocturnal hemoglobinuria. *J Thromb Haemost*. 2012;10:1–10.
19. Switzer JA, Hess DC, Nichols FT, Adams RJ. Pathophysiology and treatment of stroke in sickle-cell disease: present and future. *Lancet Neurol*. 2006;5:501–512.
20. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(Suppl 7):30–33.
21. Ghezala IB, Arendt JFB, Erichsen R, et al. Positive predictive value of the diagnosis coding for vitamin B12 deficiency anemia in the Danish National Patient Register. *Clin Epidemiol*. 2012;4:333–338.
22. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health*. 2011;39(Suppl 7):22–25.
23. Lauritsen JM, ed. EpiData Data Entry, Data Management and Basic Statistical Analysis System. *EpiData Assoc*. 2000–2008. Available from: http://www.epidata.dk. Accessed on March 21, 2015.
24. Kornblit B, Hagve TA, Taaning P, Birgens H. Phenotypic presentation and underlying mutations in carriers of beta-thalassaemia and alpha-thalassaemia in the Danish immigrant population. *Scand J Clin Lab Invest*. 2007;67(1):97–104.
25. Marchand A, Galen RS, Van Lente F. The predictive value of serum haptoglobin in hemolytic disease. *JAMA*. 1980;243(19):1909–1911.
26. Schrier S. Approach to the diagnosis of hemolytic anemia in the adult. Available from: http://www.uptodate.com/contents/approach-to-the-diagnosis-of-hemolytic-anemia-in-the-adult?source=search_result &search=hemolysis&selectedTitle=1%7E150. Accessed March 21, 2015.
27. Ballas SK. Lactate dehydrogenase and hemolysis in sickle cell disease. *Blood*. 2013;121(1):243–244.
28. Eber SW, Pekrun A, Schröter W. Prevalence of increased osmotic fragility of erythrocytes in German blood donors: screening using a modified glycerol lysis test. *Ann Hematol*. 1992;68:88–92.
29. Eber S, Lux SE. Hereditary spherocytosis—defects in proteins that connect the membrane skeleton to the lipid bilayer. *Semin Hematol*. 2004;41:118–141.
30. Bianchi P, Fermo E, Vercellati C, et al. Diagnostic power of laboratory tests for hereditary spherocytosis: a comparison study in 150 patients grouped according to molecular and clinical characteristics. *Haematologica*. 2012;97(4):516–523.
31. King M-J, Garçon L, Hoyer JD, et al. ICSH guidelines for the laboratory diagnosis of nonimmune hereditary red cell membrane disorders. *Int J Lab Hematol*. 2015;37(3):304–325.
32. Bolton-Maggs PHB, Langer JC, Iolascon A, Tittensor P, King MJ, General Haematology Task Force of the British Committee for Standards in Haematology. Guidelines for the diagnosis and management of hereditary spherocytosis – 2011 update. *Br J Haematol*. 2012;156(1):37–49.
33. King MJ, Zanella A. Hereditary red cell membrane disorders and laboratory diagnostic testing. *Int J Lab Hematol*. 2013;35: 237–243.
34. Baandrup U. *Klinisk patologi* [Clinical Pathology]. Copenhagen, Denmark: FADL; 2002.
35. Zantek ND, Koepsell SA, Tharp DR, Cohn CS. The direct antiglobulin test: a critical step in the evaluation of hemolysis. *Am J Hematol*. 2012;87(7):707–709.
36. Barcellini W, Fattizzo B, Zaninoni A, et al. Clinical heterogeneity and predictors of outcome in primary autoimmune hemolytic anemia: a GIMEMA study of 308 patients. *Blood*. 2014;124(19): 2930–2936.
37. StataCorp. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP; 2007.
38. Zanella A, Barcellini W. Treatment of autoimmune hemolytic anemias. *Haematologica*. 2014;99(10):1547–1554.
39. Gregersen H, Larsen CB, Haglund A, Mortensen R, Andersen NF, Nørgaard M. Data quality of the monoclonal gammopathy of undetermined significance diagnosis in a hospital registry. *Clin Epidemiol*. 2013;5:321–326.
40. Tuckuviene R, Kristensen SR, Helgestad J, Christensen AL, Johnsen SP. Predictive value of pediatric thrombosis diagnoses in the Danish National Patient Registry. *Clin Epidemiol*. 2010;2:107–122.
41. Lauridsen MD, Gammelager H, Schmidt M, Nielsen H, Christiansen CF. Positive predictive value of International Classification of Diseases, 10th revision, diagnosis codes for cardiogenic, hypovolemic, and septic shock in the Danish National Patient Registry. *BMC Med Res Methodol*. 2015;15:23.

Supplementary materials

Table S1 Included diagnoses

Included diagnoses	ICD-10 code
Hereditary hemolysis	
Enzyme deficiencies	D55
(G6PD) deficiency, (PK) deficiency and others	
Hemoglobinopathies	
Thalassemia	D56
Alpha-/Beta-thalassemia, thalassemia trait	
Sickle cell anemia	D57, M140c, N082c, H368c, M904
Other hemoglobinopathies	D582
Erythrocyte membrane defects	
Hereditary spherocytosis	D580
Hereditary elliptocytosis	D581
Hereditary stomatocytosis	D588a
Other hereditary hemolytic anemia	D588, D589
Acquired hemolysis	
Autoimmune hemolytic anemia	D591
Drug-induced immune and nonimmune hemolysis	D590, D592
Cold-agglutinin disease	D591a
Cold hemoglobinuria	D596a
Other nonimmune hemolysis	D594, D594c
Paroxysmal nocturnal hemoglobinuria	D595

Note: The lowercase suffix is a specific Danish implementation of ICD-10 using lowercase letters to denote subtypes.
Abbreviations: G6PD, glucose-6-phosphate dehydrogenase; PK, Pyruvate kinase; ICD-10, International Classification of Diseases Tenth Revision.

Table S2 Excluded diagnoses

Excluded diagnoses	ICD-8/ICD-10 code
ICD-8 diagnosis code of hemolysis	28209, 28219, 28229, 28230, 28239, 28249, 28250, 28258, 28259, 28299, 28390, 28391, 28392, 28393, 28395, 28399, 28649, 28659, 28690, 28699, 44649, 99751, 390448
Vitamin-B12 deficiency	D51
Folic acid deficiency	D52
Cardiovascular disease including heart valve disease before and up to 1 year after index disease diagnosis date	
Condition with artificial heart valve	Z952, Z953, Z954
Mechanical hemolysis	D594a

Note: The lowercase suffix is a specific Danish implementation of ICD-10 using lowercase letters to denote subtypes.
Abbreviation: ICD, International Classification of Diseases.