Received: 24 February 2021 | Revised: 30 June 2021 | Accepted: 16 August 2021

DOI: 10.1002/pds.5348

ORIGINAL ARTICLE

WILEY

Validity of hospital ICD-10-GM codes to identify anaphylaxis

Dominik de Sordi[1] | Sanny Kappen[1] | Fabian Otto-Sobotka[1] | Anke Kulschewski[2] | Andreas Weyland[3] | Lia Gutierrez[4] | Joan Fortuny[4] | Jonas Reinold[5] | Tania Schink[5] | Antje Timmer[1]

[1]Division of Epidemiology and Biometry, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

[2]Section for Kidney Disease and Hypertension, Clinic of Internal Medicine, Klinikum Oldenburg, Oldenburg, Germany

[3]Department of Anaesthesiology, Intensive Care Medicine, Emergency Medicine, Pain Therapy, Klinikum Oldenburg, Oldenburg, Germany

[4]Pharmacoepidemiology and Risk Management, RTI Health Solutions, Barcelona, Spain

[5]Clinical Epidemiology, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

Correspondence
Antje Timmer, Division of Epidemiology and Biometry, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany.
Email: antje.timmer@uni-oldenburg.de

Funding information
European Consortium of Iron Manufacturers

Abstract

Purpose: Anaphylaxis (ANA) is an important adverse drug reaction. We examined positive predictive values (PPV) and other test characteristics of ICD-10-GM code algorithms for detecting ANA as used in a multinational safety study (PASS).

Methods: We performed a cross-sectional study on routine data from a German academic hospital (2004–2019, age ≥ 18). Chart review was used for case verification. Potential cases were identified from the hospital administration system. The main outcome required at least one of the following: any type of specific in-hospital code (T78.2, T88.6, and T80.5) OR specific outpatient code in combination with a symptom code OR in-hospital non-specific code (T78.4, T88.7, and Y57.9) in combination with two symptom codes. PPV were calculated with 95% confidence interval. Sensitivity analyses modified type of codes, unit of analysis, verification criteria and time period. The most specific algorithm used only primary codes for ANA (numbers added in brackets).

Results: Four hundred and sixteen eligible cases were evaluated, and 78 (37) potential ANA cases were identified. PPV were 62.8% (95% CI 51.1–73.5) (main) and 77.4% (58.9–90.4) (most specific). PPV from all modifications ranged from 12.9% to 80.6%. The sensitivity of the main algorithm was 66.2%, specificity 91.5%, and negative predictive value 92.6%. Corresponding figures for the most specific algorithm were 32.4%, 98.0%, and 87.0%.

Conclusions: The PPV of the main algorithm seems of acceptable validity for use in comparative safety research but will underestimate absolute risks by about a third. Restriction to primary discharge codes markedly improves PPV to the expense of reducing sensitivity.

KEYWORDS
anaphylaxis; validation, claims data base, electronic health records, Germany

Key Points

- Anaphylaxis is an important adverse drug reaction
- Complex ICD-10 based algorithms are needed to describe anaphylaxis in safety studies
- The main algorithm as applied to hospital administrative data showed a positive predictive value (PPV) of 0.63

- Restriction of the algorithm to primary discharge codes will substantially improve the PPV, but many cases will be missed.
- The more restricted algorithm is considered helpful for bias analysis in comparative research on relative risks of ANA.

1 | INTRODUCTION

Anaphylaxis (ANA) is defined as a severe and immediate hypersensitivity reaction with rapid onset following exposure to an antigenic trigger.[1] The definition applies irrespective of the underlying pathophysiological mechanism.[2] Reported incidence rates range from $1.5/10^5$ to $7.9/10^5$ person-years with a lifetime prevalence of up to 5%.[3–5] Cases are most likely encountered in emergency departments, where they may account for about 0.03% of visits and show a 0.7% case fatality rate.[5,6] For ANA, there is no single universally applicable code in the International Classification of Diseases systems (ICD-9, ICD-10). Various ANA-specific codes may apply, depending on context, severity and trigger, in addition to codes describing allergic reactions or adverse drug reactions.[7] Various combination algorithms have been used to identify ANA as a safety outcome in pharmacoepidemiological research. Where reported, diagnostic performance varies widely depending on the algorithm used and the context in which it was applied.[7–10] Information on the validity of algorithms used to identify this adverse event in the respective population/setting is therefore crucial when examining the risk of ANA.

Following safety concerns of the European Medicines Agency (EMA) on the risk of ANA from the use of intravenous iron, a multinational European post-authorization safety study was recently performed (IV iron PASS).[11] Data sources included the German Pharmacoepidemiological Research Database (GePaRD), which contains claims data from statutory health insurance providers in Germany.[12] Access to medical charts is not possible in GePaRD due to data protection regulations. Therefore, we applied the algorithms which were developed for the PASS to administrative data from a single hospital within the catchment area of GePaRD to examine the diagnostic accuracy of these algorithms (external indirect validation). Of primary interest was the positive predictive value (PPV) of the main PASS outcome definition as used in GePaRD.

2 | METHODS

2.1 | Study design and setting

The study was performed as a cross-sectional validation study. Potential cases were retrospectively identified from hospital administrative data. Medical chart review was performed to verify the diagnosis. Protocol and procedures were developed in consensus with the IV Iron PASS study group.[11]

The setting was an acute care hospital in Germany (approximately 830 beds). The five departments contributing the most case numbers of eligible discharge codes and those likely to be involved in the treatment of ANA were approached for participation (Gastroenterology, Cardiology, Internal Medicine /Nephrology, General & Visceral Surgery, Dermatology and Emergency Medicine). All but Dermatology agreed to contribute. Data on in-hospital treatment as well as outpatient care delivered on the hospital premises by hospital-employed specialists were available (ICD-10-GM codes). The study period ranged from January 01, 2004, to April 30, 2019. We included patients aged 18 and older.

2.2 | Identification of potential cases from Hospital Information System (HIS)

2.2.1 | Sampling procedure

Sampling from HIS included all cases for which any ANA related diagnostic code as listed in Table 1 was documented during the study period (primary and secondary, discharge and admission). For non-specific codes, only in-hospital care was considered. All sampled cases ("eligible cases") served as a basis (sampling frame) to which the various case-finding algorithms were applied.

2.2.2 | Case finding algorithms

Case-finding algorithms used in the PASS followed published recommendations on the development of ANA algorithms.[9–11] In the PASS, ANA was assumed if at least one of the following criteria was fulfilled:

- A: Inpatient or emergency room encounters: any specific diagnostic code for ANA
- B: Outpatient encounters: any specific diagnostic code for ANA in combination with at least one symptom, procedure and/or treatment code indicating ANA or shock
- C: Inpatient or emergency room encounters: any non-specific diagnostic code combined with at least one symptom code compatible with ANA AND at least one code indicating shock or death.

Symptom and procedure codes as used for criteria B and C are shown in Table 2. For the primary outcome (main algorithm), discharge and admission and primary and secondary codes were considered. Minor modifications were necessitated by characteristics of the contributing databases. For GePARD, this related primarily to using ICD-10-GM (German Modification) codes and missing information on in-hospital medication.

TABLE 1 ICD-10-GM codes used for initial identification of eligible cases

ICD-10-GM codes	Descriptor
"Anaphylaxis-specific codes"	
T78.2	Anaphylactic shock, unspecified
T88.6	Anaphylactic shock owing to adverse effect of correct drug or medicament properly administered
T80.5	Anaphylactic reaction due to serum
Other "non-specific codes for anaphylaxis" or "allergy codes"	
T78.4	Allergy, unspecified
T88.7	Unspecified adverse effect of drug or medicament
Y57.9	Drugs or medicaments causing adverse effects in therapeutic use

TABLE 2 Additional symptoms, procedures and treatments used for case identification

ICD-10-GM code	Description
Symptom codes compatible with ANA	
J98.01	Acute bronchospasm
R06.1	Stridor
T78.3	Angioneurotic oedema
Symptom, procedure and treatment codes suggesting shock or death	
I46.0, I46.1, I46.9	Cardiac arrest with successful resuscitation, sudden cardiac death, cardiac arrest unspecified
I95	Hypotension
T41.5	Oxygen administration
R96, R98, R99	Other sudden death, unattended death, death from unknown cause
Treatments, procedures and symptoms used for simulating algorithm (algorithm 4)	
–	Adrenaline; H1 blocker, steroids, "antiallergic therapy"
–	Transfer to ICU, CPR, artificial ventilation
–	Any symptoms used to define or verify ANA

Abbreviation: ANA: anaphylaxis. ICU: Intensive care unit. CPR: cardiopulmonary resuscitation. H1-blocker: Histamin 1 Receptor Blocker.

For the validation study, we used the main algorithm as used by GePaRD. However, information on medication and procedures other than as coded via ICD-10-GM was unavailable.

In addition, we examined the following modifications:

- Modification 1 (most specific): primary discharge codes only
- Modification 2 (expanded to increase sensitivity): Urticaria (ICD-10-GM L50.0) and any death added as symptoms (applies to criteria B and C only).
- Modification 3 (simulating codes not available in HIS): additional information considered from medical chart review

The expanded algorithm (modification 2) was devised in analogy to a *post hoc* sensitivity analysis in the PASS including GePaRD following the observation of no or hardly any case observed for criteria B, and C. Modifications 1 and 3 do not correspond to outcomes used in the PASS but were added for additional information in the validation study only. Eligible treatments, procedures, and symptoms for modification 3 are listed in Table 2.

2.3 | Case validation

2.3.1 | Data extraction

All eligible cases with specific diagnoses, all cases fitting criterion C, as well as a random sample (target size 300) of cases with non-specific diagnoses, were selected for medical record extraction.

Discharge letters and emergency room notes were collected from the electronic hospital system, complemented by a hard copy search for outpatient notes. Anonymized documents were reviewed in random order by two independent trained researchers (AT, SK). Information was extracted using a standard data form. This included type of stay, physician-reported ANA or allergy-related diagnoses, as well as any information on relevant symptoms, treatments, procedures, suggested trigger and timing (speed of onset, time from exposure to onset of symptoms) (Form available as supplemental material, S1). Free text physician reported ANA or allergy-related diagnoses were grouped as ANA, adverse drug reaction (AE), allergic reaction other than ANA, past ANA/known allergy, and none (no ANA related diagnosis).

2.3.2 | Criteria defining true cases

For case verification based on medical charts we used diagnostic criteria proposed by the National Institute of Allergy and Infectious Disease (NIAID) and the Food Allergy and Anaphylaxis Network (FAAN), as in the PASS.[1,11] Cases were considered confirmed if one of the following three situations applied:

1. At least one major ANA symptom in combination with dermal manifestation, all with acute onset
2. Rapid onset of at least two ANA symptoms (major or minor) following exposure to a likely allergen
3. Hypotension following exposure to a known allergen for this person

Major ANA symptoms comprised severe hypotension as defined by NIAID/FAAN, and respiratory compromise.[1] Minor symptoms could be gastrointestinal (vomiting, severe abdominal cramps), skin or mucosa related (dermal manifestation: general hives/generalized itching, flush, swollen lips, uvula, or tongue), cardiovascular (syncope, incontinence, and collapse) or respiratory (dyspnea, wheeze, and stridor).

2.3.3 | Adjudication procedure

Following independent extraction and preliminary categorization by the clinical reviewers, cases were categorized by reviewer consensus according to the verification criteria as ANA ("true" cases, confirmed ANA), non-ANA and non-evaluable (insufficient information).[5] In addition, verification was performed by applying a computerized algorithm on the extracted data, as shown in Table 3. Consensus and computer-derived based diagnoses were compared, and inconsistencies resolved by re-evaluation.

2.4 | Statistical analysis

2.4.1 | Descriptive and main analyses

All analyses are presented as case-based (per encounter) unless stated otherwise. Additional patient-based analyses were performed by including only the first encounter of a patient within the study period.

PPVs were calculated as the proportion of confirmed ANA cases, relative to all potential cases with EMR available, presented as percentage (%) with 95% Clopper–Pearson confidence intervals (CI). Non-evaluable cases were treated as non-ANA ("worst PPV scenario") for the primary analysis.

2.4.2 | Subgroup and sensitivity analysis, additional analysis using missed cases

Results are also presented separately per criterion if at least ten potential cases were identified for the respective criterion. In addition, patient-based analyses are presented along with case-based analysis. Subgroup analyses were planned for type of stay, sex, and department.

Sensitivity analysis used the following modifications of the case verification algorithm:

- S1 ("best PPV scenario"): Non-evaluable cases are considered true ANA.
- S2 ("clinically sensible diagnosis"): True ANA is assumed despite insufficient information or failure to meet all formal criteria where adrenaline, cardiopulmonary resuscitation (CPR), transfer to intensive care (ICU), intubation, or artificial ventilation had been applied in the context of an allergic event.

Additional sensitivity analysis examined the effect of excluding cases discharged prior to 2008 in order to detect effects from an organizational change in coding practice.

2.4.3 | Additional diagnostic information

Using all cases sampled as eligible for which extraction was performed, sensitivity (SE), specificity (SP) and negative predictive values (NPV) were calculated with 95% Clopper-Pearson CI.

2.5 | Data management and quality control

We used double data entry of chart extraction data, predefined plausibility checks for the full dataset, and double programming for all main analyses. The data collection and management was done using the OpenClinica open source software, vs 3.[13] Main statistical analyses were generated using SAS software, vs 9.4.[14]

2.6 | Ethics and data privacy

Patient consent was not required under §4 German Federal Data Security Law. Documents were blinded prior to extraction and anonymized data used for analysis. The protocol was submitted to the Medical Ethics Committee of the University of Oldenburg and received positive evaluation (No 2018-022, March 14, 2018).

3 | RESULTS

3.1 | Description of HIS data and sampled cases

The evolution of case numbers for the main algorithm is shown in Figure 1. Overall, 5503 cases from HIS were originally identified as eligible cases. Of these, 5330 (96.9%) had non-specific in-hospital codes for adverse drug reactions or other complications of drugs

TABLE 3 Automated case verification algorithm

	Trigger	Timing	Dermal symptoms	Major symptoms (≥ 1)	Minor symptoms (≥ 2)
1		Acute (hours)	Present	a. Respiratory b. Hypotension	
2	„Likely Allergen"	Rapidly after (hours)			a. Dermal b. Respiratory c. Hypotension d. Gastrointestinal
3	„Known allergen for that patient"			Hypotension	

Note: All columns within a row must apply to confirm a case. Any case fulfilling either 1 or 2 or 3 is considered a confirmed case.

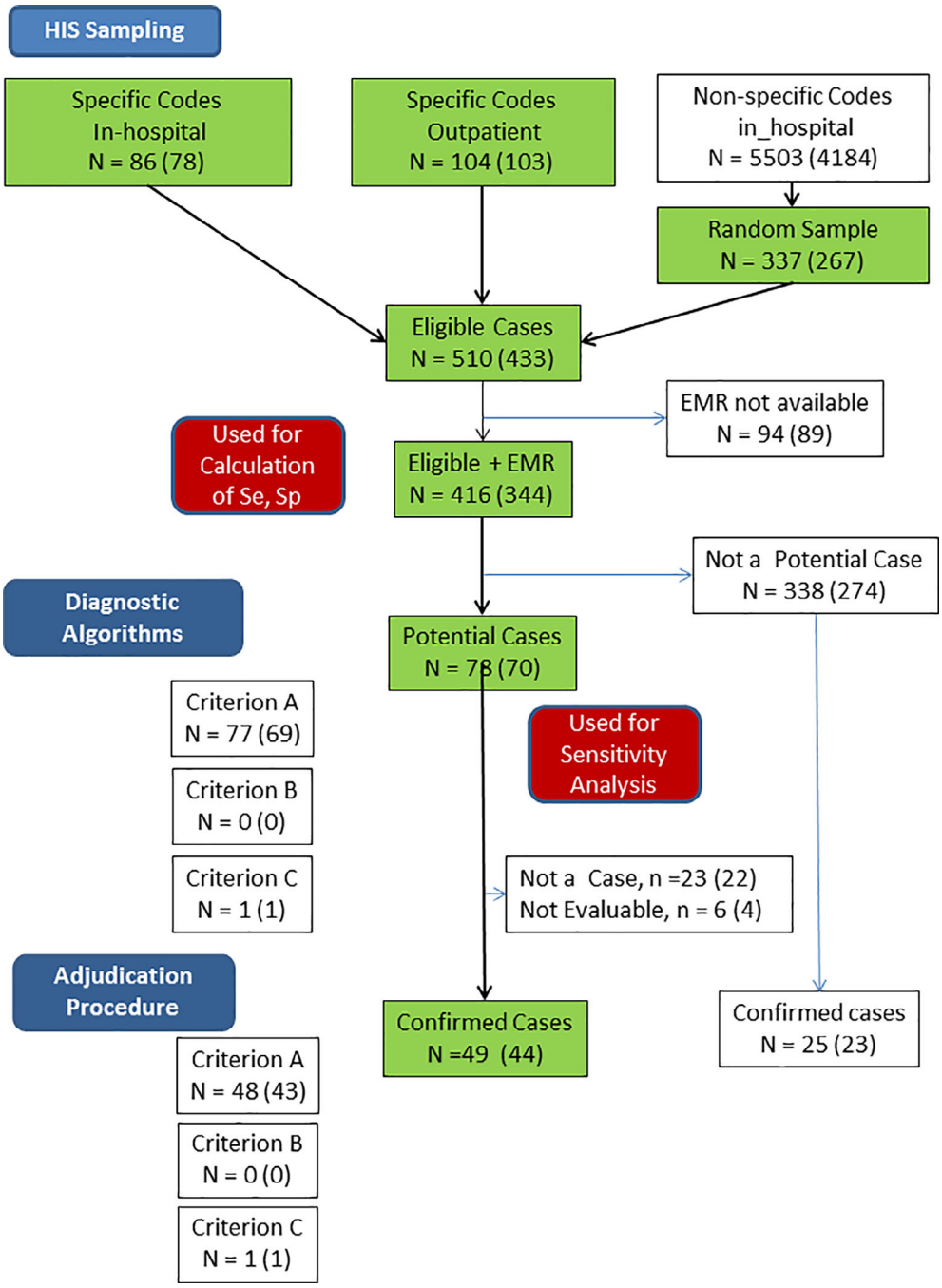

FIGURE 1 Flow of case (patient) numbers through the identification and validation process, main algorithm. HIS: Hospital information system. SE sensitivity. SP specificity. EMR electronic medical records. Codes from HIS and criteria A-C were not mutually exclusive

(Y57.9, T88.7), mostly from oncology and cardiology (Table 4). There was a steep increase in the number of eligible cases from the first 5-year period (2004–2008, 348 cases) to the second time period (2009–2013, 2351 cases). Case numbers further increased by 18.2% from the 2nd time period to the most recent time period.

Nephrology /Internal Medicine emerged as the most important contributing medical specialty once the main algorithm was applied. There were 87 potential cases (79 patients) for the main outcome and 37 (36) for the specific algorithm (modification 1) (Table 4, Figure 1). Almost all potential cases were identified by criterion A, in particular from codes T78.2 ("ANA not specified", 50 cases) and T88.6 ("drug-related ANA," 39 cases). The expanded algorithm (modification 2) increased the overall number of potential cases to 122 (113), mostly due to deceased patients from cardiology or oncology with AE coded as a secondary diagnosis. More detailed numbers per subgroup and algorithm are available as supplementary material (S2).

3.2 | Chart extraction and case adjudication results

Of 510 cases (433 patients) sampled for data extraction, chart information could be retrieved for 416 cases (344 patients) (81.6%, Figure 1),

TABLE 4 Description of full sample (as drawn from HIS) and potential cases (main algorithm)

	Full HIS sample				Main algorithm			
	A	B	C	any	A	B	C	any
ANA diagnoses								
T78.2 ANA, nos	*50*	*98*	7	**148**	*50*			**50**
T80.5 ANA (serum)								
T88.6 ANA (drug)	39	6	10	**45**	39			**39**
T78.4 Allergy nos		2	116	**118**			1	**1**
T88.7 AE (drug)	11	1	1611	**1612**	11			**11**
Y57.9 Drug Compl	6		3803	**3803**	6			**6**
Department								
Cardiology	20		*1582*	***1600***	20			**20**
Gastroenterology	6	3	456	**465**	6			**6**
Nephrology	40	100	857	**988**	40		1	**41**
Oncology	8		2192	**2197**	8			**8**
Surgery	11	1	234	**243**	11			**11**
Other	1		9	**10**	1			**1**
Time period								
2004–2008	10	77	265	**348**	10			**10**
2009–2013	32	16	2308	**2351**	32			**32**
2014–2018	43	11	2729	**2775**	43		1	**44**
ALL	***86***	***104***	***5330***	***5503***	***86***			***87***

Note: All cases are considered, irrespective of whether medical records (EMR) were available. A: in-hospital or emergency room encounters with specific ANA codes. B: outpatient cases with specific ANA related codes (+ symptom or procedure codes). C: in hospital non-specific cases (+ symptom + procedure codes). Any: any of the criteria A, B or C applies. Criteria A, B, C are not mutually exclusive. Diagnoses are not mutually exclusive. Sum of cases per diagnosis may exceed column total. AE: Adverse effect. Nos: not otherwise specified, not specified. ANA: anaphylactic shock, anaphylactic reaction. Department: Department of discharge. In the case of "other," the case was admitted to one of the participating departments but transferred prior to discharge.

TABLE 5 Main analysis: PPV

	Algorithm	*n*	*N*	PPV	CI lower	CI upper
Case based	Main	49	78	62.8%	51.1	73.5
	Modification 1	24	31	77.4%	58.9	90.4
	Modification 2	52	108	48.1%	38.4	58.0
	Modification 3	57	90	63.3%	52.5	73.2
Patient based	Main	44	70	62.9%	50.5	74.1
	Modification 1	23	30	76.7%	57.7	90.1
	Modification 2	47	99	47.5%	37.3	57.8
	Modification 3	52	82	63.4%	52.0	73.8

Note: *n*: confirmed cases (true ANA); *N*: potential cases, EMR available.

mostly from discharge letters. Loss of cases due to unavailable medical records was frequent in outpatient cases. Only 44 of 104 eligible cases (44.3%) could be evaluated from this group compared to 91.7% completeness for in-hospital cases. The numbers of potential cases (patients) were 78 (70) for the main algorithm and ranged from 31 (30) for the most specific to 108 (99) for the expanded algorithm.

Case verification identified overall 74 true cases (18%) in the screened dataset (76 following expansion for algorithm 3 and 4, 17%).

3.3 | Positive predictive values

For the main algorithm, a PPV of 62.8% (95% CI 51.1%–73.5%) was calculated based on 49 true cases (Table 5). Restriction to primary discharge codes increased the PPV to 77.4% (58.9%–90.4%), but only 24 true cases were identified. Modification 2 increased the number of potential cases to 52, but PPV decreased to 48.1% (95% CI 38.4 to 58.0). This was due to a very low PPV for the non-specific codes

TABLE 6 Clinical information from patient charts on confirmed ANA cases (case based, main algorithm)

	ANA cases	No ANA	Not evaluable	ALL
Diagnosis as reported in chart				
• anaphylactic reaction or shock	36 (73.5%)	6 (26.1%)	5 (83.3%)	47 (60.3%)
• adverse drug reaction	2 (4.1%)			2 (2.6%)
• allergic reaction	9 (18.4%)	9 (39.1%)		18 (23.1%)
• past ANA, known allergy		2 (8.7%)	1 (16.7%)	3 (3.8%)
• no ANA related diagnosis	2 (4.1%)	6 (26.1%)		8 (10.3%)
Suggested Trigger				
• medication	18 (36.7%)			
• serum, plasmapheresis	2 (4.1%)			
• radiocontrast	2 (4.1%)			
• plaster strip, ointment	1 (2.0%)			
• venom (insect sting)	15 (30.6%)			
• food	7 (14.3%)			
• undecipherable	1 (2.0%)			
• not known	7 (14.3%)			
Timing of trigger				
• before admission	39 (79.6%)			
• during current hospital stay	8 (16.3%)			
• no information, n/a	2 (4.1%)			
Total n	49	23	6	78

Note: Diagnoses were treated hierarchically (as ordered), thus were mutually exclusive. Triggers: not mutually exclusive, sum of cases may exceed total.

(criterion C), calculated as 12.9% (95% CI 3.6%–29.8%). Modification 3 (simulated codes) did not substantially change PPV estimates as compared to the main algorithm, but this algorithm identified the highest absolute number of true cases (57) (Table 5).

3.3.1 | Subgroup and sensitivity analysis

Subgroup analyses were based on small numbers. However, it seemed that cases admitted via the Emergency Department (main PPV 71.2%, 95% CI 57.9%–82.2%) and those discharged from internal medicine/nephrology (main PPV 75.0%, 95% CI 57.8%–87.9%) were more likely true ANA than those without ED admission (PPV not calculated due to low numbers), or those discharged by all other departments combined (PPV 52.4%, 95% CI 36.4%–68.0%).

Tables of PPV for all algorithms with all case definition modifications, per criterion and for the respective full algorithm, are available as Supplementary Material. Both the best PPV scenario and the "clinically sensitive" case verification slightly improved PPV for all algorithms. This was most marked for the most restrictive algorithm: Case-based PPV increased from 77.4% to 80.6% (61.4%–92.3%) for both variations.

Exclusion of cases presenting prior to 2009 did not substantially nor systematically alter results. Also, patient-based values closely resembled case-based results (Table 5).

3.4 | Extracted clinical information relating to ANA

Of the 49 confirmed cases from the main algorithm, 36 had a diagnosis of ANA reported in the medical chart (73.5%) (Table 6). ANA diagnoses were, in addition, reported in the charts of six of the 23 cases classified as non-ANA (26.1%), as well as in five of the six non-evaluable cases. The majority of confirmed ANA had been triggered before admission. However, there were also eight instances of in-hospital induced ANA. Most triggers were related to medical treatment or diagnostic procedures (18 medication, five other medical agents), followed by venom (15 cases, mostly wasps). Indicated medication included antibiotic therapy, NSAIDs/analgesics, chemotherapy, and other.

The most frequently reported ANA related symptoms were hypotension (29, 59.2%) and dyspnoea (28, 57.1%). There were two instances of ANA related death. The most frequently reported treatments were corticosteroids in 41 cases (83.7%) and H1 blockers in 40 cases (81.6%). Application of adrenaline was specifically mentioned in 23 cases (46.9%) and transfer to ICU in 16 cases (32.7%).

3.5 | Other diagnostic information, results based on missed cases

The main algorithm missed 25 of 74 confirmed cases in the sample of eligible cases (SE 66.2%), but NPV and SP both exceeded 90% (92.6%

TABLE 7 Test characteristics of diagnostic algorithms

	Main algorithm	Modification 1	Modification 2	Modification 3
Case based				
• Missed cases (n)	25 (33.8%)	50 (67.6%)	24 (31.6%)	19 (25.0%)
• PPV (%)	62.8 (51.1–73.5)	77.4 (58.9–90.4)	48.1 (38.4–58.0)	63.3 (52.5–73.2)
• NPV (%)	92.6 (89.3–95.2)	87.0 (83.2–90.2)	92.9 (89.6–95.4)	94.6 (91.7–96.7)
• SE (%)	66.2 (54.3–76.8)	32.4 (22.0–44.3)	68.4 (56.7–78.6)	75.0 (63.7–84.2)
• SP (%)	91.5 (88.0–94.2)	98.0 (95.8–99.2)	84.8 (80.7–88.3)	91.0 (87.6–93.7)
• All true cases	74	74	76[a]	76[a]
• Total evaluated	416	416	444[a]	444[a]
Patient based				
• Missed cases (n)	23 (34.3%)	44 (65.7%)	22 (31.9%)	17 (24.6%)
• PPV (%)	62.9 (50.5–74.1)	76.7 (57.7–90.1)	47.5 (37.3–57.8)	63.4 (52.0–73.8)
• NPV (%)	91.6 (87.7–94.6)	86.0 (81.6–89.6)	91.9 (88.0–94.9)	94.1 (90.7–96.5)
• SE (%)	65.7 (53.1–76.8)	34.3 (23.2–46.9)	68.1 (55.8–78.8)	75.4 (63.5–84.9)
• SP (%)	90.6 (86.5–93.8)	97.5 (94.8–99.0)	82.8 (78.0–86.9)	90.1 (86.1–93.2)
• All true cases	67	67	69[a]	69[a]
• Total evaluated	344	344	371[a]	371[a]

Note: PPV positive predictive value. NPV negative predictive value. SE sensitivity, SP Specificity.
[a]Totals and number of missed cases increased following re-sampling of non-specific codes for application of post hoc modified algorithms (modifications 2 and 3; additional deaths and cases of urticaria included).

and 91.5%, respectively) (Table 7). As expected, for the most restrictive algorithm, SE was worst (32.4%), while SP was excellent (98.0%) (modification 1). SE was best for the simulating algorithm (75.0%, modification 3).

Diagnostic information restricted to cases identified by criterion A was very similar to the main algorithm, with a SE of 64.9% (patient-based: 64.2%), SP of 91.5% (90.6%) and NPV of 92.3% (91.3%). The underlying detailed 2 × 2 diagnostic tables are supplied as Supplementary material (S3).

4 | DISCUSSION

In this indirect external validation study, we determined the validity of ICD-10-GM codes based algorithms describing ANA. For the primary outcome, a PPV of 63% was calculated. This is in accordance with results from direct validation within IV Iron PASS, as well as with results from previous research on comparable algorithms.[9–11] About a third of all true cases in our preselected sample were missed. The more specific algorithm resulted in a PPV of almost 80% but identified substantially fewer cases.

Almost all potential cases had been identified by specific ANA related codes in in-patient cases (criterion A). For outpatient and non-specific ANA-related codes, the algorithm required additional ICD-10 symptom codes (criteria B and C), which were likely to be underreported. We used several approaches to examine the effect this may have had on the validity of the algorithms.

First, the main algorithm was expanded to include all deaths occurring in combination with ANA related codes. This increased numbers of potential ANA, but the resulting PPV was substantially worse. In contrast to a previous report, based on our data, misclassification of ANA related deaths does not seem to play a relevant role when examining ANA in safety studies, at least when case fatality is low.[15]

Second, we examined the frequency with which any of the symptoms, procedures and treatments which were part of the PASS outcome algorithms were reported in the charts, thus would have been available for coding. Using these data to inform case finding ("simulation" algorithm) improved case ascertainment as well as the PPV for criterion C (non-specific codes). The effect on the overall PPV, however, was small.

Lastly, we analyzed the proportion of cases missed by the combination algorithms as compared to using ANA related codes without any further conditions. These results have to be interpreted with caution as cases with non-specific codes were under-sampled for pragmatic reasons. Also, we only evaluated a preselected sample with an increased probability of ANA. Taking these factors into account, SE would be substantially lower, and NPV higher than estimated. This lends evidence to the impression that the inclusion of non-specific codes to identify ANA is probably not efficient, even if more symptom codes were available.

Additional observations from chart review included the use of adrenaline (epinephrine) in less than half of the confirmed ANA cases and a substantial proportion of cases with insufficient information to fulfill formal criteria. Both findings are in accordance with data from the literature.[7,16–19] In particular, substantial guideline non-adherence has been shown with respect to the application of epinephrine in ANA in Germany. Neither underuse of adrenaline nor undercoding seem specific to this particular hospital. The preponderance of patients from nephrology may be explained by the fact that this department also represented internal medicine in general and, possibly, a higher prevalence of multiple drug treatment. A steep increase of case numbers within the first observation period was explained by the establishment of a trained coding team in the hospital around the year 2008.

The strengths of our approach include the application of a number of sensitivity analyses corroborating the robustness of the estimated PPV. Correction for multiple visits (patient-based analysis) did not change the results, nor did the exclusion of cases coded before 2009. The moderate increase of case numbers over time thereafter is in line with consistent reports of an increasing incidence of ANA, in particular drug-induced ANA in older persons.[4]

Overall, the study profited from close collaboration with an international study group with respect to protocol adherence, quality assurance and immediate applicability. Two of our examined algorithms directly compared to the main and the expanded outcome definition used in the PASS.

Limitations include the restriction to a single center, insufficient case numbers for subgroups, and underrepresentation of outpatient cases. The proportion of hospitalization for ANA varies substantially in the literature and may be as low as around 12%.[4,20] However, ANA related codes in ICD-10 as used in the PASS algorithms are reported to be biased towards severe cases.[7] Also, algorithms were formulated with a focus on drug-induced ANA, which seems to be associated with a higher risk of severe ANA as defined by hospitalization, ICU treatment, CPR, or fatal course.[4,20,21] Therefore, we assume that a focus on hospitalized ANA cases is probably appropriate for most comparative drug safety studies using ICD-10 based algorithms. Better documentation by treating physicians will be essential to improve coding, but coding itself is hampered by the unsatisfactory classification of ANA by ICD-10 diagnostic codes. The revised ICD system (ICD-11, expected to be introduced from 2022) and the addition of novel methods to identify potential cases, in particular, natural language processing techniques, may improve the recognition of ANA in the future.[22–25]

5 | CONCLUSIONS

The assessed algorithm seems useful for identifying ANA cases in particular in hospital settings for comparative safety studies. A more restricted modification may be used for sensitivity analysis to examine the effect of including false-positive events on relative estimates. Both algorithms underestimate the absolute risk of ANA. Identifying cases via non-specific or outpatient codes may improve sensitivity, but efficiency is questionable as long as recognition, reporting and coding of diagnoses, symptoms and procedures are insufficient.

ACKNOWLEDGMENTS

The authors acknowledge Constanze Kathan-Selck (Klinikum Oldenburg) for coordination of the in-hospital team; Hans Seifert (Gastroenterology), Claus-Henning Köhne (Oncology), Albrecht Elsässer (Cardiology) and Dalibor Bockelmann (Visceral Surgery) for cooperation/data access; Jan Thies Soller (Epidemiology and Biometry, UOL) for statistical programming and data management; Rainer Röhrig (Medical Informatics, RWTH Aachen) for development of data protection concept and counseling; Lara Disselhoff (project student) for blinding, printing, sorting, and other support work; Dirk Niehaus and Heiko Seemann (KOL administration) for providing for data transfer; and Nuria Saigi-Morgui (Pharmacoepidemiology and Risk Management RTI-HS) for the management of the project and coordination of activities between the Oldenburg and RTI-HS team. Open Access funding enabled and organized by ProjektDEAL.

CONFLICT OF INTEREST

This study was funded by a European Consortium of Iron Manufacturers and was conducted under a contract including the ENCePP Seal granting independent publication rights to the research team. Antje Timmer, Dominik de Sordi, Sanny Kappen and Fabian Otto-Sobotka are employed at the University of Oldenburg. The unit has previously performed a similar study for the pharmaceutical industry. Anke Kulschewski and Andreas Weyland do not report a conflict of interest.

Tania Schink and Jonas Reinold are working at the Leibniz Institute for Prevention Research and Epidemiology – BIPS. Unrelated to this study, BIPS occasionally conducts studies financed by the pharmaceutical industry. Almost exclusively, these are post-authorization safety studies (PASS) requested by health authorities. The studies and the resulting publications are not influenced by the pharmaceutical industry. Lia Gutierrez and Joan Fortuny are full-time employees of RTI Health Solutions, a unit of RTI International, a non-profit research institution that conducts work for government, public, and private organizations, including pharmaceutical companies.

ETHICS STATEMENT

This study received positive evaluation from the Oldenburg University Medical Ethics review board (No 2018-022, March 14, 2018).

ORCID

Fabian Otto-Sobotka https://orcid.org/0000-0002-9874-1311
Jonas Reinold https://orcid.org/0000-0001-8266-2574
Tania Schink https://orcid.org/0000-0002-0224-1866
Antje Timmer https://orcid.org/0000-0001-9579-0516

REFERENCES

1. Sampson HA, Munoz-Furlong A, Campbell RL, et al. Second symposium on the definition and management of anaphylaxis: summary report-second National Institute of allergy and infectious disease/Food Allergy

and Anaphylaxis Network symposium. *J Allergy Clin Immunol*. Feb 2006;117(2):391-397. https://doi.org/10.1016/j.jaci.2005.12.1303
2. Cardona V, Ansotegui IJ, Ebisawa M, et al. World allergy organization anaphylaxis guidance 2020. *World Allergy Org J*. Oct 2020;13(10):100472. https://doi.org/10.1016/j.waojou.2020.100472
3. Panesar SS, Javad S, de Silva D, et al. The epidemiology of anaphylaxis in Europe: a systematic review. *Allergy*. Nov 2013;68(11):1353-1361. https://doi.org/10.1111/all.12272
4. Regateiro FS, Marques ML, Gomes ER. Drug-induced anaphylaxis: an update on epidemiology and risk factors. *Int Arch Allergy Immunol*. 2020;181(7):481-487. https://doi.org/10.1159/000507445
5. Altman AM, Camargo CA Jr, Simons FE, et al. Anaphylaxis in America: a national physician survey. *Allergy*. Mar 2015;135(3):830-833. https://doi.org/10.1016/j.jaci.2014.10.049
6. Montanez MI, Mayorga C, Bogas G, et al. Epidemiology, mechanisms, and diagnosis of drug-induced anaphylaxis. *Front Immunol*. 2017;8:614. https://doi.org/10.3389/fimmu.2017.00614
7. Tuttle KL, Wickner P. Capturing anaphylaxis through medical records: are ICD and CPT codes sufficient? *Ann Allergy Asthma Immunol*. Feb 2020;124(2):150-155. https://doi.org/10.1016/j.anai.2019.11.026
8. de Silva D, Singh C, Muraro A, et al. Diagnosing, managing and preventing anaphylaxis: systematic review. *Allergy*. Sep 2021;76(5):1493-1565. https://doi.org/10.1111/all.14580
9. Schneider G, Kachroo S, Jones N, et al. A systematic review of validated methods for identifying anaphylaxis, including anaphylactic shock and angioneurotic edema, using administrative and claims data. *Pharmacoepidemiol Drug Saf*. Jan 2012;21(Suppl 1):240-247. https://doi.org/10.1002/pds.2327
10. Walsh KE, Cutrona SL, Foy S, et al. Validation of anaphylaxis in the Food and Drug Administration's mini-sentinel. *Pharmacoepidemiol Drug Saf*. Nov 2013;22(11):1205-1213. https://doi.org/10.1002/pds.3505
11. Fortuny J, von Gersdorff G, Lassalle R, et al. Use of IV iron and risk of anaphylaxis: a multinational observational post-authorisation safety study in Europe. *Pharmacoepidemiol drug Saf*. 2021. https://doi.org/10.1002/pds.5319
12. Pigeot I, Ahrens W. Establishment of a pharmacoepidemiological database in Germany: methodological potential, scientific value and practical limitations. *Pharmacoepidemiol Drug Saf*. Mar 2008;17(3):215-223. https://doi.org/10.1002/pds.1545
13. Cavelaars M, Rousseau J, Parlayan C, et al. OpenClinica. *J Clin Bioinforma*. 2015;5(1):1-2.
14. SAS Statistical Analysis System. *SAS Release 9.4 for Windows*. Cary, NC: SAS Institute Inc; 2016.
15. Tanno LK, Ganem F, Demoly P, Toscano CM, Bierrenbach AL. Undernotification of anaphylaxis deaths in Brazil due to difficult coding under the ICD-10. *Allergy*. Jun 2012;67(6):783-789. https://doi.org/10.1111/j.1398-9995.2012.02829.x
16. Grabenhenrich LB, Dölle S, Ruëff F, et al. Epinephrine in severe allergic reactions: the European anaphylaxis register. *J Allergy Clin Immunol Pract*. Nov-Dec 2018;6(6):1898-1906.e1. https://doi.org/10.1016/j.jaip.2018.02.026
17. Langelaan M, Baines RJ, de Bruijne MC, Wagner C. Association of admission and patient characteristics with quality of discharge letters: posthoc analysis of a retrospective study. *BMC Health Serv Res*. Mar 21 2017;17(1):225. https://doi.org/10.1186/s12913-017-2149-8
18. Unnewehr M, Schaaf B, Marev R, Fitch J, Friederichs H. Optimizing the quality of hospital discharge summaries–a systematic review and practical tools. *Postgrad Med*. Aug 2015;127(6):630-639. https://doi.org/10.1080/00325481.2015.1054256
19. Wimsett J, Harper A, Jones P. Review article: components of a good quality discharge summary: a systematic review. *Emerg Med Australas*. Oct 2014;26(5):430-438. https://doi.org/10.1111/1742-6723.12285
20. Clark S, Wei W, Rudders SA, Camargo CA Jr. Risk factors for severe anaphylaxis in patients receiving anaphylaxis treatment in US emergency departments and hospitals. *J Allergy Clin Immunol*. Nov 2014;134(5):1125-1130. https://doi.org/10.1016/j.jaci.2014.05.018
21. Worm M, Francuzik W, Renaudin JM, et al. Factors increasing the risk for a severe reaction in anaphylaxis: an analysis of data from the European anaphylaxis registry. *Allergy*. Jun 2018;73(6):1322-1330. https://doi.org/10.1111/all.13380
22. Yu W, Zheng C, Xie F, et al. The use of natural language processing to identify vaccine-related anaphylaxis at five health care systems in the vaccine safety Datalink. *Pharmacoepidemiol Drug Saf*. Feb 2020;29(2):182-188. https://doi.org/10.1002/pds.4919
23. Banerji A, Lai KH, Li Y, et al. Natural language processing combined with ICD-9-CM codes as a novel method to study the epidemiology of allergic drug reactions. *J Allergy Clin Immunol Pract*. Dec 16 2020;8(3):1032-1038. https://doi.org/10.1016/j.jaip.2019.12.007
24. Ball R, Toh S, Nolan J, Haynes K, Forshee R, Botsis T. Evaluating automated approaches to anaphylaxis case classification using unstructured data from the FDA sentinel system. *Pharmacoepidemiol Drug Saf*. Oct 2018;27(10):1077-1084. https://doi.org/10.1002/pds.4645
25. Tanno LK, Chalmers R, Jacob R, et al. Global implementation of the world health organization's international classification of diseases (ICD)-11: the allergic and hypersensitivity conditions model. *Allergy*. Sep 2020;75(9):2206-2218. https://doi.org/10.1111/all.14468

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: de Sordi D, Kappen S, Otto-Sobotka F, et al. Validity of hospital ICD-10-GM codes to identify anaphylaxis. *Pharmacoepidemiol Drug Saf*. 2021;30(12):1643-1652. doi:10.1002/pds.5348