

ORIGINAL REPORT

WILEY

Can we identify allergic rhinitis from administrative data: A validation study

Katja Biering Leth-Møller^{1,2,3}  | Tea Skaaby¹ | Flemming Madsen⁴ |
Janne Petersen^{1,5} | Allan Linneberg^{1,2}

¹Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark

²Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

³Global Marketing, ALK-Abelló A/S, Hørsholm, Denmark

⁴Department of Allergology, Allergy and Lung Clinic Helsingør, Helsingør, Denmark

⁵Section of Biostatistic, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Correspondence

Katja Biering Leth-Møller, Center for Clinical Research and Prevention, Frederiksberg and Bispebjerg Hospital, Frederiksberg, Denmark.
Email: katja.biering.leth-moeller@regionh.dk

Funding information

ALK Abelló A/S; Innovationsfonden; Lundbeckfonden

Abstract

Background: Important insights on, for example, prevalence, disease progression, and treatment of allergic rhinitis can be obtained from large-scale database studies if researchers are able to identify allergic individuals. We aimed to assess the validity of 13 different algorithms based on Danish nationwide prescription and/or hospital data to identify adults with allergic rhinitis.

Methods: Our primary gold standard of allergic rhinitis was a positive serum specific IgE (≥ 0.35) and self-reported nasal symptoms retrieved from two general health examination studies conducted in Danish adults (18–69 years) during 2006 to 2008 ($n = 3416$) and 2012 to 2015 ($n = 7237$). The secondary gold standard of allergic rhinitis was self-reported physician diagnosis. We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value, and corresponding 95% confidence intervals (95% CI) for each register-based algorithm in the two time periods.

Results: Sensitivity (≤ 0.40) was low for all algorithms irrespective of definition of allergic rhinitis (gold standard) or time period. The highest PPVs were obtained for algorithms requiring both antihistamines and intranasal corticosteroids; yielding a PPV of 0.69 (0.62–0.75) and a corresponding sensitivity of 0.10 (0.09–0.12) for the primary gold standard of allergic rhinitis in 2012 to 2015.

Conclusion: Algorithms based on both antihistamines and intranasal corticosteroids yielded the highest PPVs. However, the PPVs were still moderate and came at the expense of low sensitivity when applying the strict primary gold standard (sIgE and nasal symptom).

KEYWORDS

allergic rhinitis, pharmacoepidemiology, prescription algorithms, real world evidence, sensitivity, validation

1 | INTRODUCTION

Allergic rhinitis affects more than one in five in Europe¹ and the prevalence has been increasing.^{2–4} Large-scale register-based studies have the potential to provide important data on prevalence, time trends, disease progression, comorbidities, and treatment of allergic

rhinitis. However, to reach that potential, reliable identification of individuals with allergic rhinitis from databases is crucial.⁵ Due to nationwide high-quality health registries which can be linked at an individual-level Denmark offers a unique setting for such register-based research. Data on diseases that are diagnosed and treated by a general practitioner or a private practicing medical specialist are,

however, not available from any national register. Thus, identification of individuals with allergic rhinitis in Denmark can be based on prescription data and/or data from hospitals. Prescription-based algorithms to identify allergic rhinitis have also been commonly used in studies utilising data from other countries.⁶⁻¹¹

Allergic rhinitis is usually diagnosed by a combination of a history of allergic symptoms and a diagnostic test.¹² Skin-prick-test (SPT) and serum specific IgE (sIgE) are well-established diagnostic tests for allergic sensitisation.¹² However, non-clinical relevant allergic sensitisation occurs often, that is, sensitised individuals may not experience symptoms when exposed to the allergen.¹² Allergic rhinitis is commonly managed by oral/ophthalmic antihistamines (second generation) and intranasal corticosteroids (INC). However, these treatments are also indicated for other conditions such as urticaria, non-allergic rhinitis, and non-allergic rhinosinusitis. Thus, using these drugs to identify allergic rhinitis might result in substantial misclassification.

To our knowledge only two studies have validated register-based algorithms to identify allergic rhinitis and only among children.^{10,13} Additionally, none of the studies have included diagnostic tests in their gold standard. We aimed to assess the validity of algorithms based on Danish register data to reliably identify adults with allergic rhinitis.

2 | METHODS

In Denmark, health care services delivered at public hospitals and by private practitioners (general or specialist) are largely free of charge for all residents. Prescribed drugs are, however, subject to co-payment by the patient, where the degree of co-payment depends on the individual's yearly drug expenses.¹⁴ INCs are primarily available on prescription (94% of all sales in 2015), whereas oral and ophthalmic antihistamines are widely available over-the-counter. For the most commonly used oral non-sedative antihistamines (Anatomical Therapeutic Chemical [ATC] group R06AE) in Denmark only 36% of all sales were on prescription in 2015.¹⁵ Reasons for prescribing OTC drugs include reimbursement and that it is a reliable and easy way for prescribers to keep track of their patients' medication use.

2.1 | Registry data

Data on prescribed drugs for allergic rhinitis was collected from the Danish National Prescription Register,¹⁶ where information on every dispensed drug prescription from Danish pharmacies is available. Drugs are classified according to the ATC system. We included INCs (ATC: R01AD01, R01AD05, R01AD08-9, R01AD11-2, R01AD58) and oral/ophthalmic antihistamines (ATC: S01GA51, S01GX02, S01GX06-09, R06AX13, R06AX18, R06AX22, R06AX26-29, R06AE07, R06AE09). Data on diagnoses (classified according to International Classification of Diseases 10th version [ICD-10]) are available from the Danish National Patient Register,¹⁷ which holds information on

KEY POINTS

- Prescription drugs are often used as a proxy to identify allergic rhinitis.
- We validated 13 different algorithms using Danish prescription and/or hospital data to identify allergic rhinitis among adults.
- Sensitivity (≤ 0.40) was low for all algorithms due to low use of prescription drugs.
- The highest PPVs were obtained for algorithms requiring both antihistamines and intranasal corticosteroids; yielding a PPV of 0.69 (95% CI: 0.62-0.75).
- Researches using administrative data must consider how to reliably identify allergic rhinitis, and how a potential misclassification will impact their results.

in- and outpatient services at Danish hospitals. We included diagnoses of allergic rhinoconjunctivitis (ICD-10: J30, H101, H104A).

Based on prescription and/or diagnosis data we tested 13 algorithms to identify allergic rhinitis (Table 1). Five of the algorithms were inspired by the validation work by Mulder et al.¹⁰ We created four algorithms combining different drug types, hypothesising that it would decrease the risk of including individuals claiming prescriptions for other indications. Finally, we tested one algorithm solely based on hospital diagnosis and three previously used algorithms^{6,7,9,13,18} (Tables 1 and S1).

2.2 | Health examination studies

We included two general health examination studies; Health2006 (n = 3416) and DanFunD (n = 7237). Both have been described in detail elsewhere.^{19,20} Briefly, the two studies were independently conducted in 2006 to 2008 and 2012 to 2015 at the Research Centre for Prevention and Health, Glostrup, Denmark (now Center for Clinical Research and Prevention) and each comprise a random sample of individuals aged 18 to 69 years, born in Denmark and with Danish citizenship, resident in 11 municipalities in the Western part of the Copenhagen Region. Allergic sensitisation was assessed to the four common aeroallergens birch, cat, grass, and house dust mite by sIgE (in Health2006: ADVIA Centaur, Bayer Corporation; in DanfunD: Phadia 250, Thermo Scientific ImmunoDiagnostic). In Health2006 a subset of 2450 participants were also skin-prick tested. The SPTs were performed using lancets (Soluprick SQ, ALK-Abelló, Denmark) and a standard panel of 10 allergens (birch, grass, mugwort, horse, cat, dog, two mites, and two molds) including a negative control and a positive control (10 mg/mL histamine dichloride). Information on self-reported nasal symptoms and physician-diagnosed hay fever was collected in a questionnaire completed at home by the participants prior to the health examination.

TABLE 1 Included algorithms

	Description
≥1 Antihistamine	At least one prescription of oral/nasal/ophthalmic antihistamine
≥2 Antihistamines	At least two prescriptions of oral/nasal/ophthalmic antihistamine claimed within 365 days
≥1 INC	At least one prescription of INC
≥2 INCs	At least two prescriptions of INCs claimed within 365 days
≥2 Prescriptions	At least two prescriptions of oral/nasal/ophthalmic antihistamine or INC claimed within 365 days
≥2 INC and antihistamine	At least one prescription of intranasal corticosteroids and at least one prescription of oral/nasal/ophthalmic antihistamine claimed within 365 days
≥3 INC and antihistamine	≥3 prescriptions where at least one is INC and at least one is oral/nasal/ophthalmic antihistamine claimed within 365 days
≥4 INC and antihistamine	≥4 prescriptions where at least one prescription is INC and at least one is oral/nasal/ophthalmic antihistamine claimed within 365 days
≥2 INC and antihistamine (April-July)	At least one prescription of oral/nasal/ophthalmic antihistamine and at least one prescription of INC claimed within between April and July
Hospital diagnosis	A diagnosis of allergic rhinitis or chronic allergic conjunctivitis
≥2 INC or oral antihistamine (Linneberg et al ¹⁸) ^a	At least two prescriptions of oral antihistamine or INC claimed within 365 days
Hospital diagnosis or ≥2 INCs/oral antihistamines or ≥1 ophthalmic antihistamine/AIT (Henriksen et al ^{6,13}) ^a	At least one of the following criteria need to be fulfilled: (a) a diagnosis of allergic rhinitis or chronic rhinitis, (b) at least two prescriptions of INCs, (c) at least two prescriptions of oral antihistamines, (d) at least one prescription of ophthalmic antihistamine or AIT
≥1 of INC during April to August for three consecutive years (Zielen et al ⁹) ^{a,b}	At least one prescription of INC during April-August for three consecutive years

Abbreviations: AIT, allergen-specific immunotherapy; INC, intranasal corticosteroids.

^aDescribed in detail in Table S1.

^bAlgorithm only used to identify individuals with allergic rhinitis due to grass. Based on prescription data 3 years prior to examination.

2.3 | Gold standards

We included a primary and a secondary gold standard of allergic rhinitis from the two health examination studies.

1. *slgE-defined allergic rhinitis*: a serum slgE ≥ 0.35 kU/L¹² (this threshold is widely used in clinical practice and provided as the cut-off by the manufacturers) and a confirmative answer to at least one of “Itchy or stuffy nose or sneezing when near grass, trees, or flowers within the last 12 months”; “Itchy or stuffy nose or sneezing when near furry animals within the last 12 months”; “Itchy or stuffy nose or sneezing when cleaning rooms or making beds, or when in bed within the last 12 months.”
2. *Self-reported physician-diagnosed hay fever*: a confirmative answer to “Has a physician ever told you that you have hay fever.”

We chose *slgE-defined allergic rhinitis* as the primary gold standard because it includes the objective measure of slgE and resembles the clinical diagnosis of allergic rhinitis. However, this definition could be too strict due to a limited number of allergens included and questions not covering the full symptom range. Thus, we included *self-reported physician-diagnosed hay fever* as a secondary gold standard. This has previously been used to define allergic rhinitis and might reflect another branch of allergic rhinitis. The

difference between the two gold standards will be discussed in detail in the discussion.

2.4 | Analysis

For each participant in Health2006 and DanFunD we included prescription data 2 years prior to date of the examination. Two years were chosen because we aimed to assess the validity of the algorithm at the time an individual was classified with allergic rhinitis, thus including more years prior to the health examination would increase the risk of misclassifying the gold standard. Two years were chosen over 1 year to ensure all participants had been exposed to at least two pollen seasons. Due to the rarity of hospital contacts related to allergic rhinitis, we included diagnosis data 5 years prior to date of the examination.

We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and corresponding 95% confidence intervals (95% CI) for each of the algorithms using the R package “epiR.” This was done individually for both the primary and the secondary gold standard. To assess performance of the algorithms at two time points we did separate analyses for Health2006 (2006-2008) and DanFunD (2012-2015). A sensitivity analysis was conducted to account for the fact that PPV is affected by the

TABLE 2 Characteristics of the health examination studies

	Health2006 (2006-2008)	DanFunD (2012-2015)
n	3416 ^a	7237
Age (median [IQR])	48.6 [38.8-58.6]	52.4 [42.3-61.4]
Women (%)	1892 (55.4)	3916 (54.1)
Antihistamines (%)	412 (12.1)	722 (10.0)
Prescriptions, n (median [IQR]) ^b	2.00 [1.00-4.00]	2.00 [1.00-6.00]
Intranasal corticosteroids (%)	326 (9.5)	768 (10.6)
Prescriptions, n (median [IQR]) ^b	1.00 [1.00-3.00]	2.00 [1.00-4.00]
<i>slgE-defined (%)</i>		
Allergic rhinitis (primary gold standard)	566 (17.0)	1364 (19.2)
Atopy	791 (23.3)	1887 (26.4)
<i>SPT-defined (%)</i>		
Allergic rhinitis (4 allergens)	419 (18.3)	NA
Allergic rhinitis (10 allergens)	438 (19.1)	NA
Atopy (4 allergens)	642 (27.4)	NA
Atopy (10 allergens)	716 (30.6)	NA
<i>Self-reported physician diagnoses (%)</i>		
Hay fever	604 (18.0)	1367 (19.2)
Asthma	361 (10.7)	734 (10.3)
Diabetes	129 (3.8)	313 (4.4)
<i>BMI (%)</i>		
<18.5	65 (1.9)	110 (1.5)
18.5-25	1577 (46.2)	3228 (44.6)
25-30	1223 (35.8)	2656 (36.7)
>30	549 (16.1)	1237 (17.1)
<i>Self-reported smoking (%)</i>		
Current	877 (25.9)	1202 (16.7)
Former	1089 (32.2)	2689 (37.3)
Never	1416 (41.9)	3321 (46.0)
<i>Self-rated health (%)</i>		
Very bad	26 (0.8)	48 (0.7)
Bad	265 (7.8)	623 (8.7)
Moderate	1425 (42.2)	2915 (40.5)
Good	1333 (39.5)	2955 (41.0)
Very good	328 (9.7)	659 (9.2)

Note: Allergic rhinitis is defined by a combination of nasal symptoms and slgE- or SPT-defined sensitisation. Atopy is defined by slgE- (≥ 0.35 kU/L) or SPT-defined sensitisation (mean wheal diameter ≥ 3 mm) to at least one of birch, cat, grass, and house dust mite. SPT (10 allergens) includes at least one of birch, grass, mugwort, horse, cat, dog, two mites, and two molds.

Abbreviations: IQR, interquartile range; slgE, specific IgE; SPT, skin-prick test.

^aFor the SPT population n = 2450.

^bAmong those who have claimed a prescription.

prevalence of the condition in the study population.²¹ We calculated PPVs for a range of allergic rhinitis prevalence.¹⁰ Further, to validate if birch, cat, grass, and house dust mite (included in the primary gold standard) sufficiently cover allergic rhinitis we included all 10 allergens (birch, grass, mugwort, horse, cat, dog, two mites, and two molds) from the SPT in Health2006. Finally, we stratified the primary gold

standard according to allergen (pollen, cat, and house dust mite). Based on clinical experience we hypothesised that identifying allergic rhinitis due to pollen (birch and grass) would be more reliable than allergic rhinitis due to cat or house dust mite.

For each analysis, only individuals with complete information in the primary or secondary gold standard variable were included.

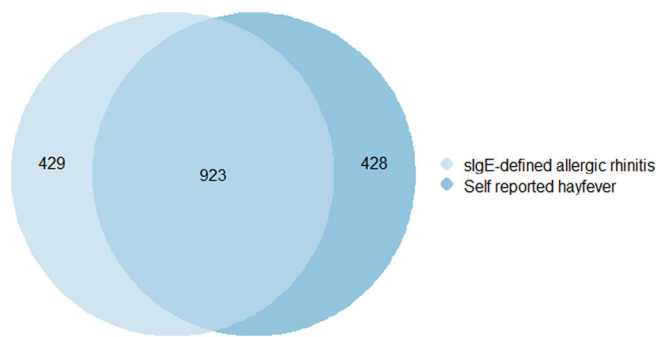


FIGURE 1 Comparison of allergic rhinitis cases classified by sIgE-defined allergic rhinitis and self-reported physician-diagnosed hay fever in 2012 to 2015. sIgE, specific IgE [Colour figure can be viewed at wileyonlinelibrary.com]

2.5 | Ethics

Health2006 and DanFunD were approved by the Ethics Committee of Copenhagen and the Danish Data Protection Agency. Recommendations of the Declaration of Helsinki were followed, and each participant gave informed written consent. This study was approved by the Danish Data Protection Agency (ref. 2012-58-0004).

3 | RESULTS

We included 3416 and 7237 participants from Health2006 and DanFunD. Characteristics of the participants are shown in Table 2. In 2012 to 2015 (DanFunD) 19% had sIgE-defined allergic rhinitis and of those 923 (69%) individuals also reported physician-diagnosed hay fever (Figure 1). Sensitisation patterns are shown in Figures S1 and S2.

A total of 2.7% were excluded from the sIgE-defined allergic rhinitis analyses and 1.9% from the self-reported physician-diagnosed hay fever analysis due to missing values in 2006 to 2008. The corresponding numbers were 1.8% and 1.5% in 2012 to 2015.

Due to low use of prescription medication among those with allergic rhinitis the sensitivity was low (≤ 0.40 ; Table 3) for all prescription-based algorithms irrespectively of definition of allergic rhinitis (gold standard). For the algorithm solely based on diagnoses from hospitals, sensitivity was even lower (< 0.02 ; Table 3).

For both the primary and secondary gold standard of allergic rhinitis, the algorithms with the highest validity in 2012 to 2015 were those requiring both antihistamines and INC (Table 3). The algorithm ≥ 2 INC and antihistamine yielded a PPV of 0.64 (95% CI 0.58–0.70) for the primary gold standard (sIgE-defined allergic rhinitis; Table 3). Restricting the period where prescription could be claimed to April to July resulted in a higher PPV of 0.69 (95% CI 0.62–0.75), but at the expense of somewhat lower sensitivity. Additionally, the restriction changed the distribution of the identified allergens slightly. Without the restriction 92%, 35%, and 25% had allergic rhinitis due to pollen (birch and grass), house dust mite,

and/or cat, respectively. With the restriction the corresponding numbers were 94%, 31%, and 25%.

The validity of the algorithms was similar in 2006 to 2008 (Table S2 and Figure 2) compared to 2012 to 2015.

3.1 | Sensitivity analyses

Figure 2 shows PPVs for the algorithm ≥ 2 INC and antihistamine as a function of the underlying prevalence for the two golden standards in 2006 to 2008 and 2012 to 2015. Overall, the PPV was lower using the primary gold standard (sIgE-defined allergic rhinitis) compared to using the secondary gold standard.

As 90% of the individuals sensitised to 1 of the 10 allergens included in the SPT sample were sensitised to at least one of birch, cat, grass, and house dust mite, including more allergens did not yield a higher PPV (Figure S3).

The stratified analyses yielded a higher sensitivity and specificity for pollen compared to cat and house dust mite. Thus, the PPVs were 0.64, 0.45, and 0.44, accordingly, assuming a fixed prevalence of 19% for all three categories.

4 | DISCUSSION

In the current study we tested the validity of 13 algorithms, based on nationwide Danish registry data, to reliably identify patients with allergic rhinitis. Algorithms based on both antihistamines and intranasal corticosteroids performed best. However, low sensitivities and moderate PPVs made reliable identification dubious, irrespectively of method to assess “true” allergic rhinitis (the gold standards).

No studies have validated algorithms to identify allergic rhinitis among adults from register data, but two studies have been conducted among children. One of these studies concludes that an algorithm based on prescriptions and/or a hospital diagnosis is valid to identify allergic rhinoconjunctivitis (sensitivity = 84.4% and PPV = 68.2%) in Denmark.¹³ Compared to our study, the sensitivity was markedly higher, and it seems unlikely to be explained alone by differences in included age group and gold standard. However, Stensballe et al¹³ only included children with algorithm-identified allergic rhinoconjunctivitis, asthma, and/or atopic dermatitis in their study. With this sampling method they keep the marginal row totals fixed and are not able to estimate sensitivity and specificity.²² Thus, the presented sensitivity does not yield a reliable picture of the algorithms' performance in their target population and is not comparable to our study. A Dutch validation study concludes that prescription-based algorithms to identification of children diagnosed (gold standard is a general practitioner diagnosis) with allergic rhinitis is questionable, since high PPVs could only be obtained in combination with low sensitivity.¹⁰ This is in contrast to identification of asthma, where they find PPVs between 0.80 and 0.90 for their prescription-based algorithms.¹⁰ For both the primary and the secondary gold standard of allergic rhinitis, the sensitivity was low due to limited use of

TABLE 3 Validity for different algorithms to identify allergic rhinitis in 2012 to 2015 for the primary and secondary gold standard

Prevalence	Primary gold standard (sIgE and symptoms)					Secondary gold standard (self-reported physician-diagnosed hay fever)				
	19%					19%				
Algorithm	PPV (95% CI)	Sensitivity (95% CI)	NPV (95% CI)	Specificity (95% CI)	PPV (95% CI)	Sensitivity (95% CI)	NPV (95% CI)	Specificity (95% CI)	PPV (95% CI)	Sensitivity (95% CI)
≥1 Antihistamine	0.50 (0.46, 0.53)	0.26 (0.24, 0.28)	0.84 (0.83, 0.85)	0.94 (0.93, 0.94)	0.64 (0.60, 0.67)	0.33 (0.30, 0.35)	0.86 (0.85, 0.87)	0.96 (0.95, 0.96)		
≥2 Antihistamines	0.56 (0.51-0.61)	0.15 (0.13-0.17)	0.83 (0.82-0.84)	0.97 (0.97-0.98)	0.76 (0.71-0.80)	0.20 (0.18-0.22)	0.84 (0.83-0.85)	0.99 (0.98-0.99)		
≥1 INC	0.38 (0.34, 0.41)	0.21 (0.19, 0.23)	0.83 (0.82, 0.84)	0.92 (0.91, 0.92)	0.49 (0.45, 0.52)	0.27 (0.24, 0.29)	0.84 (0.83, 0.85)	0.93 (0.93, 0.94)		
≥2 INCs	0.41 (0.36-0.46)	0.11 (0.09-0.13)	0.82 (0.81-0.83)	0.96 (0.96-0.97)	0.54 (0.49-0.59)	0.14 (0.12-0.16)	0.83 (0.82-0.84)	0.97 (0.97-0.98)		
≥2 Prescriptions	0.44 (0.40-0.48)	0.21 (0.19-0.23)	0.83 (0.82-0.84)	0.94 (0.93-0.94)	0.61 (0.57-0.65)	0.28 (0.26-0.31)	0.85 (0.84-0.86)	0.96 (0.95-0.96)		
≥1 INC and ≥1 antihistamine	0.64 (0.58-0.70)	0.12 (0.11-0.14)	0.83 (0.82-0.83)	0.98 (0.98-0.99)	0.82 (0.77-0.87)	0.16 (0.14-0.18)	0.83 (0.82-0.84)	0.99 (0.99-0.99)		
≥3 INC and antihistamine	0.66 (0.59-0.73)	0.10 (0.08-0.12)	0.82 (0.81-0.83)	0.99 (0.98-0.99)	0.85 (0.79-0.89)	0.13 (0.08-0.12)	0.83 (0.82-0.84)	0.99 (0.99-1.00)		
≥4 INC and antihistamine	0.66 (0.58-0.73)	0.08 (0.06-0.09)	0.82 (0.81-0.83)	0.99 (0.99-0.99)	0.85 (0.79-0.90)	0.10 (0.08-0.12)	0.82 (0.81-0.83)	1.00 (0.99-1.00)		
≥1 INC and ≥ 1 antihistamine (April-July)	0.69 (0.62-0.75)	0.10 (0.09-0.12)	0.82 (0.81-0.83)	0.99 (0.99-0.99)	0.87 (0.81-0.91)	0.13 (0.11-0.15)	0.83 (0.82-0.84)	1.00 (0.99-1.00)		
Hospital diagnosis	0.65 (0.43-0.84)	0.01 (0.01-0.02)	0.81 (0.80-0.82)	1.00 (1.00-1.00)	0.75 (0.53-0.90)	0.01 (0.01-0.02)	0.81 (0.80-0.82)	1.00 (1.00-1.00)		
≥2 INC or oral antihistamine (Linneberg et al ¹⁸)	0.44 (0.40-0.48)	0.19 (0.17-0.21)	0.83 (0.82-0.84)	0.94 (0.94-0.95)	0.61 (0.57-0.65)	0.25 (0.23-0.27)	0.84 (0.84-0.85)	0.96 (0.96-0.97)		
Hospital diagnosis or ≥2 INCs/ oral antihistamines or ≥1 ophthalmic antihistamine/ AIT (Henriksen et al ^{6,13})	0.46 (0.42-0.50)	0.25 (0.23-0.27)	0.84 (0.83-0.85)	0.93 (0.92-0.94)	0.61 (0.57-0.64)	0.32 (0.30-0.35)	0.86 (0.85-0.86)	0.95 (0.94-0.96)		
≥1 of INC during April to August for three consecutive years (Zielen et al ^{9a})	0.49 (0.41-0.58)	0.05 (0.04-0.06)	0.81 (0.80-0.82)	0.99 (0.98-0.99)	0.69 (0.61-0.77)	0.07 (0.06-0.08)	0.82 (0.81-0.83)	0.99 (0.99-0.99)		

Abbreviations: INC, intranasal corticosteroids; NPV, negative predictive value; PPV, positive predictive value; sIgE, specific IgE.

^aOriginal algorithm only used to identify individuals with allergic rhinitis due to pollen. Based on prescription data 3 years prior to examination.

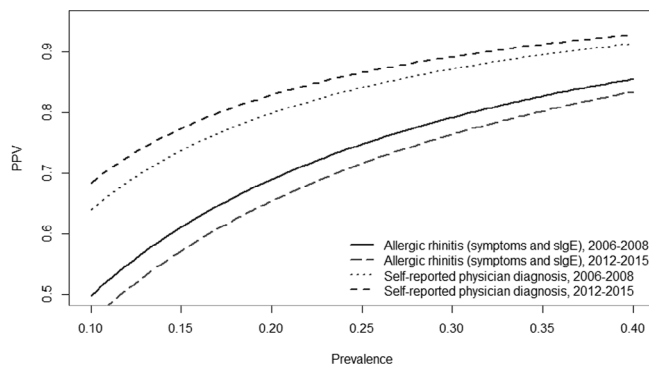


FIGURE 2 PPVs for a range of values of the underlying prevalence. PPVs are based on the algorithm ≥ 1 INC and ≥ 1 antihistamine within 365 days. INC, intranasal corticosteroids; PPVs, positive predictive values

prescription medication among those classified with allergic rhinitis. This is in line with previous studies showing a high degree of underdiagnosis and undertreatment of allergic rhinitis in the general population.^{1,23} Additionally, the low sensitivity, was not unexpected since oral/ophthalmic antihistamines are widely available over-the-counter in Denmark, meaning, a proportion of those individuals not claiming antihistamines on prescription, could have bought them over-the-counter instead.

Strengths of this validation study includes a large sample size, sampling from the general population, inclusion of an objective measure of allergic sensitisation (sIgE) in the primary gold standard, and gold standards obtained at two different time points (2006-2008 and 2012-2015). We included two different gold standards for allergic rhinitis, and especially for PPV the algorithms' performance depended on the chosen gold standard. The validity of our algorithms was markedly higher when the gold standard was self-reported physician-diagnosed hay fever compared to the primary gold standard, especially in 2012 to 2015. This indicates a higher correlation between drug use and perceived hay fever diagnosis than between drug use and allergic sensitisation and nasal symptoms (sIgE-defined allergic rhinitis). Several factors could explain this difference. Firstly, if individuals claiming antihistamine and INC were diagnosed with allergic rhinitis to other allergens than the four included (eg, horse or alder), they would not have allergic rhinitis according to our primary gold standard but should still report physician-diagnosed hay fever. However, the sensitivity analysis showed that birch, cat, grass, and house dust mite is the most clinically significant aeroallergens in Denmark, and misclassification in the primary gold standard seems implausible to explain the difference between the two gold standards alone. Secondly, individuals could be misclassified in the primary gold standard if they have symptoms when exposed to an allergen, but their symptom experience is different than those described in the questionnaire or they do not recall their symptoms correctly, for example, if they are reporting outside the pollen season. Still, using the algorithm with the highest PPV (1 INC and 1 antihistamine in April-July), less than 75% of those identified have allergic sensitisation. Thus, misclassification of symptoms

does not drive the differences between the primary and the secondary gold standard. Thirdly, of those reporting physician-diagnosed hay fever more than 25% have no allergic sensitisation indicating they might recall wrongly or have been diagnosed by their physician without diagnostic testing (sIgE or SPT). Finally, the discrepancy could potentially be due our chosen threshold of sIgE ≥ 0.35 kU/L if it is too strict. However, this seems less likely since this threshold is widely used in clinical practice in Denmark and the overlap with SPT is high (Figure S3). Thus, the higher PPV (≈ 0.80) for physician-diagnosed hay fever indicates treatment of symptoms related to those of allergic rhinitis, but might also include differential diagnoses, such as non-allergic rhinitis, local allergic rhinitis (nasal sIgE response without a positive SPT or serum sIgE^{12,24}), and polyps. Which of the two gold standards and their correspondingly validity measures that are most relevant depends on the study of interest and the importance of sIgE sensitisation in a given research question.

Even though this validation study applies to Danish data, the clinical setting with availability of over-the-counter antihistamines and intranasal corticosteroids as well as underdiagnosis and -treatment of allergic rhinitis is not a Danish phenomenon.^{1,9,12,23} Thus, the results should be relevant for researchers using other data sources. Additionally, Mulder et al¹⁰ used a different gold standard and focused on a paediatric Dutch population and still reached a similar conclusion as we do: identification of allergic rhinitis based on prescription data should be done with caution. Researchers must, therefore, consider how the potential high risk of misclassification will impact their results when using prescription data to identify allergic rhinitis. Which accuracy measure to prioritise depends on the intended use of the algorithm,²¹ for example, a high PPV is important if the researcher needs to identify a cohort who truly have allergic rhinitis. This can be the case in a pharmacoepidemiologic study where such a cohort is used as the control group.^{7-9,18,25-27} Some studies have used diagnosis codes only (including general practitioners and private practicing medical specialists) to identify allergic rhinitis,^{10,25-28} which might be more reliable. However, that identification strategy has, to our knowledge, not been validated. We found moderate PPVs when using hospital diagnosis (Tables 3 and S2), although we only had few cases (13 in 2006-2008 and 23 in 2012-2015). The Danish National Prescription Registry includes information regarding the indication of the prescription which potentially could increase the reliability of using prescriptions to identify allergic rhinitis in Denmark. However, this variable has not been validated and have a high degree of missing data. Thus, it is not yet recommended for research purposes.²⁹

To conclude, due to low use of prescription medication among those with allergic rhinitis the sensitivity was low (≤ 0.40) for all algorithms irrespectively of the definition of allergic rhinitis. Algorithms based on both antihistamines and intranasal corticosteroids yielded the highest PPVs. However, the PPVs were still moderate, and came at the expense of low sensitivity, when applying a strict gold standard (sIgE and nasal symptoms). Researchers using administrative data must consider how to reliably identify allergic rhinitis, for example, using different data sources, and how a potential misclassification will impact their results.

ETHICS STATEMENT

Health2006 and DanFunD were approved by the Ethics Committee of Copenhagen and the Danish Data Protection Agency. Recommendations of the Declaration of Helsinki were followed, and each participant gave informed written consent. This study was approved by the Danish Data Protection Agency (ref. 2012-58-0004).

ACKNOWLEDGEMENTS

Katja Biering Leth-Møller was financially supported by Innovation Fund Denmark. Tea Skaaby was funded by the Lundbeckfonden. The authors thank the project team behind Health2006 and DanFunD at the Research Center for Prevention and Health (now: Center for Clinical Research and Prevention) for their invaluable commitment in collecting and assuring data. The DanFunD Steering Committee consisting of Prof. Torben Jørgensen, MD, DMSc (principal investigator), Prof. Per Fink, MD, DMSc, Lene Falgaard Eplow MD, PhD (senior consultant), Betina Heinsbæk Thuesen, MSc, PhD, Susanne Brix Pedersen, MSc, PhD, Michael Eriksen Benros, MD, PhD, and Thomas M. Dantoft, MSc, PhD (DanFunD scientific officer). The DanFunD cohort was funded by the Danish Foundations, The Lundbeckfonden (grant number R155-2013-14070) and TrygFonden (grant number 7-11-0213). The funding sources had no role in the design, conduct, or reporting of the study.

CONFLICT OF INTEREST

Katja Biering Leth-Møller is an industrial PhD student at ALK-Abelló A/S and University of Copenhagen. Allan Linneberg has received lecture fees from ALK-Abelló A/S. Tea Skaaby is employed by Leo Pharma. Flemming Madsen and Janne Petersen declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

All authors contributed to the design of the study, interpretation of the study results, critically revised the article, and approved the final version. Allan Linneberg was involved in the data collection process. Katja Biering Leth-Møller initiated the study, analysed the data, and drafted the article.

ORCID

Katja Biering Leth-Møller  <https://orcid.org/0000-0002-1624-9789>

REFERENCES

- Bauchau V, Durham SR. Prevalence and rate of diagnosis of allergic rhinitis in Europe. *Eur Respir J*. 2004;24(5):758-764.
- Bjerg A, Ekerljung L, Eriksson J, et al. Increase in pollen sensitization in Swedish adults and protective effect of keeping animals in childhood. *Clin Exp Allergy*. 2016;46(10):1328-1336.
- Linneberg A, Jørgensen T, Nielsen NH, Madsen F, Frolund L, Dirksen A. The prevalence of skin-test-positive allergic rhinitis in Danish adults: two cross-sectional surveys 8 years apart. The Copenhagen allergy study. *Allergy*. 2000;55(8):767-772.
- Leth-Møller KB, Skaaby T, Linneberg A. Allergic rhinitis and allergic sensitisation are still increasing among Danish adults. *Allergy*. 2019;75:660-668.
- Almqvist C, Lundholm C. Population-based data on asthma and allergic disease call for advanced epidemiologic methods. *J Allergy Clin Immunol*. 2015;136(3):656-657.
- Henriksen L, Simonsen J, Haerskjold A, et al. Incidence rates of atopic dermatitis, asthma, and allergic rhinoconjunctivitis in Danish and Swedish children. *J Allergy Clin Immunol*. 2015;136(2):360-366.e362.
- Devillier P, Wahn U, Zielen S, Heinrich J. Grass pollen sublingual immunotherapy tablets provide long-term relief of grass pollen-associated allergic rhinitis and reduce the risk of asthma: findings from a retrospective, real-world database subanalysis. *Expert Rev Clin Immunol*. 2017;13(12):1199-1206.
- Wahn U, Bachert C, Heinrich J, Richter H, Zielen S. Real-world benefits of allergen immunotherapy for birch pollen-associated allergic rhinitis and asthma. *Allergy*. 2018;74:594-604.
- Zielen S, Devillier P, Heinrich J, Richter H, Wahn U. Sublingual immunotherapy provides long-term relief in allergic rhinitis and reduces the risk of asthma: a retrospective, real-world database analysis. *Allergy*. 2017;73:165-177.
- Mulder B, Groenhouf F, Kocabas LI, et al. Identification of Dutch children diagnosed with atopic diseases using prescription data: a validation study. *Eur J Clin Pharmacol*. 2016;72(1):73-82.
- Mulder B, Schuiling-Veninga CC, Bos HJ, De Vries TW, Jick SS, Hak E. Prenatal exposure to acid-suppressive drugs and the risk of allergic diseases in the offspring: a cohort study. *Clin Exp Allergy*. 2014;44(2):261-269.
- Bousquet J, Khaltaev N, Cruz AA, et al. Allergic rhinitis and its impact on asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy*. 2008;63(Suppl 86):8-160.
- Stensballe LG, Klanso L, Jensen A, Haerskjold A, Thomsen SF, Simonsen J. The validity of register data to identify children with atopic dermatitis, asthma or allergic rhinoconjunctivitis. *Pediatr Allergy Immunol*. 2017;28(6):535-542.
- Olejz M, Juul Nielsen A, Rudkjøbing A, Okkels Birk H, Krasnik A, Hernández-Quevedo C. Denmark health system review. *Health Syst Transit*. 2012;14(2) i-xxii:1-192.
- Register of Medical Product Statistics. Medstat.dk. <http://www.medstat.dk/>
- Pottegård A, Schmidt SAJ, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data resource profile: the Danish National Prescription Registry. *Int J Epidemiol*. 2017;46(3):798-798f.
- Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449-490.
- Linneberg A, Jacobsen RK, Jespersen L, Abildstrom SZ. Association of subcutaneous allergen-specific immunotherapy with incidence of autoimmune disease, ischemic heart disease, and mortality. *J Allergy Clin Immunol*. 2012;129(2):413-419.
- Thuesen BH, Cerqueira C, Aadahl M, et al. Cohort profile: the Health2006 cohort, research centre for prevention and health. *Int J Epidemiol*. 2014;43(2):568-575.
- Dantoft TM, Ebstrup JF, Linneberg A, et al. Cohort description: the Danish study of functional disorders. *Clin Epidemiol*. 2017;9:127-139.
- Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. *J Clin Epidemiol*. 2012;65(3):343-349. e342.
- Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer: New York, NY; 2009.
- Gronhoj Larsen C, Gyldenlove M, Linneberg A. Allergic rhinitis is often undiagnosed and untreated: results from a general population study of Danish adults. *Clin Respir J*. 2013;7(4):354-358.

24. Rondon C, Campo P, Togias A, et al. Local allergic rhinitis: concept, pathophysiology, and management. *J Allergy Clin Immunol*. 2012;129(6):1460-1467.
25. Schmitt J, Schwarz K, Stadler E, Wustenberg EG. Allergy immunotherapy for allergic rhinitis effectively prevents asthma: results from a large retrospective cohort study. *J Allergy Clin Immunol*. 2015;136(6):1511-1516.
26. Schmitt J, Wustenberg E, Kuster D, Mucke V, Serup-Hansen N, Tesch F. The moderating role of allergy immunotherapy in asthma progression: results of a population-based cohort study. *Allergy*. 2019;75:596-602.
27. Hankin CS, Cox L, Lang D, et al. Allergen immunotherapy and health care cost benefits for children with allergic rhinitis: a large-scale, retrospective, matched cohort study. *Ann Allergy Asthma Immunol*. 2010;104(1):79-85.
28. Kim BK, Kim JY, Kang MK, et al. Allergies are still on the rise? A 6-year nationwide population-based study in Korea. *Allergol Int*. 2016;65(2):186-191.
29. Kildemoes HW, Sorensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health*. 2011;39(7 Suppl):38-41.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Leth-Møller KB, Skaaby T, Madsen F, Petersen J, Linneberg A. Can we identify allergic rhinitis from administrative data: A validation study. *Pharmacoepidemiol Drug Saf*. 2020;29:1423-1431. <https://doi.org/10.1002/pds.5120>

Table S1: Detailed description of previously used algorithms

Study	Description
Linneberg et al. ¹⁸	Country: Denmark Study period: 1997-2006 Age: >18 years Algorithm: at least two prescriptions of oral antihistamine (ATC codes: R06AX13, R06AX18, R06AX22, R06AX26, R06AX27, R06AE07) or INC (R01AD01, R01AD05, R01AD08, R01AD09, R01AD11, R01AD12) claimed within 365 days
Henriksen et al. ^{6,13}	Country: Denmark Study period: 1997-2011 Age: <16 years Algorithm: CRITERIA 1 (based on ICD-10): ≥1 hospital contact for: J30 "hay fever and allergic rhinitis" J30.0 "vasomotor rhinitis" J30.1 "allergic rhinitis due to pollen" J30.2 "other seasonal allergic rhinitis" J30.3 "other allergic rhinitis" J30.4 "allergic rhinitis, unspecified" J31.0 "chronic rhinitis" CRITERIA 2 (based on ATC and ICD-10): ≥ 2 filled prescriptions of: R01AD01 –R01AD60 "inhaled corticosteroids for rhinitis" and no hospital contact for (exclusions criteria): J33 "nasal polyps" J330 "polyps in nasal cavity" J331 "polyp related sinus degeneration" J331A "woakes' ethmoiditis" J338 "nasal polyps, other" J338A "polypus sinus sphenoidalis" J339 "nasal polyps, unspecified" J010-J019 "acute sinusitis" J320 –J329 "chronic sinusitis" CRITERIA 3 (based on ICD-10 and ATC) ≥2 filled prescriptions of: R06A "antihistamines for systemic use" and no hospital contact for: L29 "pruritus" or DL50 "allergic urticaria" CRITERIA 4 (based on ATC) ≥ 1 filled prescriptions of: V01A "specific immune therapy, allergen substract therapy" or/andS01GX "medication for allergic conjunctivitis"
Zielen et al. ⁹	Country: Germany Study period: 2009-2016

	Age: >5 years Algorithm: at least one prescription of INC (ATC code: R01AD*) during May-August for three consecutive years. *Assumed code. Zielen et al. state they used ATC R01A1, however we were not able to identify such an ATC code.
--	--

Table S2: Validity for different algorithms to identify allergic rhinitis in 2006–2008 for the primary and secondary gold standard

	Primary gold standard (sIgE and symptoms)				Secondary gold standard (self-reported physician-diagnosed hay fever)			
Prevalence	17%				18%			
Algorithm	PPV (95% CI)	Sensitivity (95% CI)	NPV (95% CI)	Specificity (95% CI)	PPV (95% CI)	Sensitivity (95% CI)	NPV (95% CI)	Specificity (95% CI)
≥1 antihistamine	0.47 (0.42–0.53)	0.33 (0.30–0.37)	0.87 (0.86–0.88)	0.92 (0.91–0.93)	0.60 (0.55–0.65)	0.40 (0.36–0.44)	0.88 (0.86–0.89)	0.94 (0.93–0.95)
≥2 antihistamines	0.51 (0.44–0.58)	0.17 (0.14–0.21)	0.85 (0.84–0.86)	0.97 (0.96–0.97)	0.69 (0.62–0.75)	0.23 (0.20–0.26)	0.85 (0.84–0.86)	0.98 (0.97–0.98)
≥1 INC	0.36 (0.31–0.42)	0.20 (0.17–0.24)	0.85 (0.84–0.86)	0.93 (0.92–0.94)	0.50 (0.44–0.56)	0.26 (0.23–0.30)	0.85 (0.84–0.87)	0.94 (0.93–0.95)
≥2 INCs	0.36 (0.28–0.45)	0.09 (0.07–0.12)	0.84 (0.83–0.85)	0.97 (0.96–0.97)	0.58 (0.49–0.66)	0.14 (0.11–0.17)	0.84 (0.82–0.85)	0.98 (0.97–0.98)
≥2 prescriptions	0.42 (0.36–0.48)	0.22 (0.19–0.26)	0.85 (0.84–0.87)	0.94 (0.93–0.95)	0.60 (0.54–0.66)	0.30 (0.26–0.34)	0.86 (0.85–0.87)	0.96 (0.95–0.96)
≥1 INC and ≥ 1 antihistamine	0.65 (0.56–0.73)	0.15 (0.12–0.18)	0.85 (0.84–0.86)	0.98 (0.98–0.99)	0.78 (0.70–0.84)	0.17 (0.14–0.21)	0.84 (0.83–0.86)	0.99 (0.98–0.99)
≥3 INC and antihistamine	0.65 (0.55–0.74)	0.11 (0.09–0.14)	0.84 (0.83–0.86)	0.99 (0.98–0.99)	0.80 (0.71–0.87)	0.14 (0.11–0.17)	0.84 (0.83–0.85)	0.99 (0.99–1.00)
≥4 INC and antihistamine	0.62 (0.50–0.73)	0.08 (0.06–0.10)	0.84 (0.83–0.85)	0.99 (0.99–0.99)	0.82 (0.72–0.90)	0.11 (0.08–0.13)	0.84 (0.82–0.85)	0.99 (0.99–1.00)
≥1 INC and ≥ 1 antihistamine (April–July)	0.67 (0.57–0.76)	0.12 (0.09–0.15)	0.85 (0.83–0.86)	0.99 (0.98–0.99)	0.82 (0.74–0.89)	0.14 (0.11–0.17)	0.84 (0.83–0.85)	0.99 (0.99–1.00)
Hospital diagnosis	0.69 (0.39–0.91)	0.02 (0.01–0.03)	0.83 (0.82–0.84)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.02 (0.01–0.04)	0.82 (0.81–0.84)	1.00 (0.77–1.00)
≥2 INC or oral antihistamine (Linneberg et al. ¹⁸)	0.42 (0.36–0.48)	0.19 (0.16–0.23)	0.85 (0.84–0.86)	0.95 (0.94–0.95)	0.62 (0.56–0.68)	0.27 (0.24–0.31)	0.86 (0.84–0.87)	0.96 (0.96–0.97)
Hospital diagnosis or ≥2 INCs/oral antihistamines or ≥1 ophthalmic antihistamine/AIT (Henriksen et al. ^{6,13})	0.46 (0.40–0.51)	0.27 (0.23–0.31)	0.86 (0.85–0.87)	0.93 (0.92–0.94)	0.62 (0.56–0.67)	0.35 (0.31–0.39)	0.87 (0.86–0.88)	0.95 (0.94–0.96)
≥1 of INC during April–August for 3 consecutive years (Zielen et al. ⁹)	0.43 (0.29–0.57)	0.04 (0.03–0.06)	0.83 (0.82–0.85)	0.99 (0.98–0.99)	0.73 (0.59–0.84)	0.07 (0.05–0.09)	0.83 (0.82–0.84)	0.99 (0.99–1.00)

Table note: INC = intranasal corticosteroids. PPV=Positive predictive value. NPV= Negative predictive value. †Original algorithm only used to identify individuals with allergic rhinitis due to pollen. Based on prescription data three years prior to examination.

Figure S1: Sensitisation patterns in Health2006 (2006–2008) measured by sIgE and SPT.

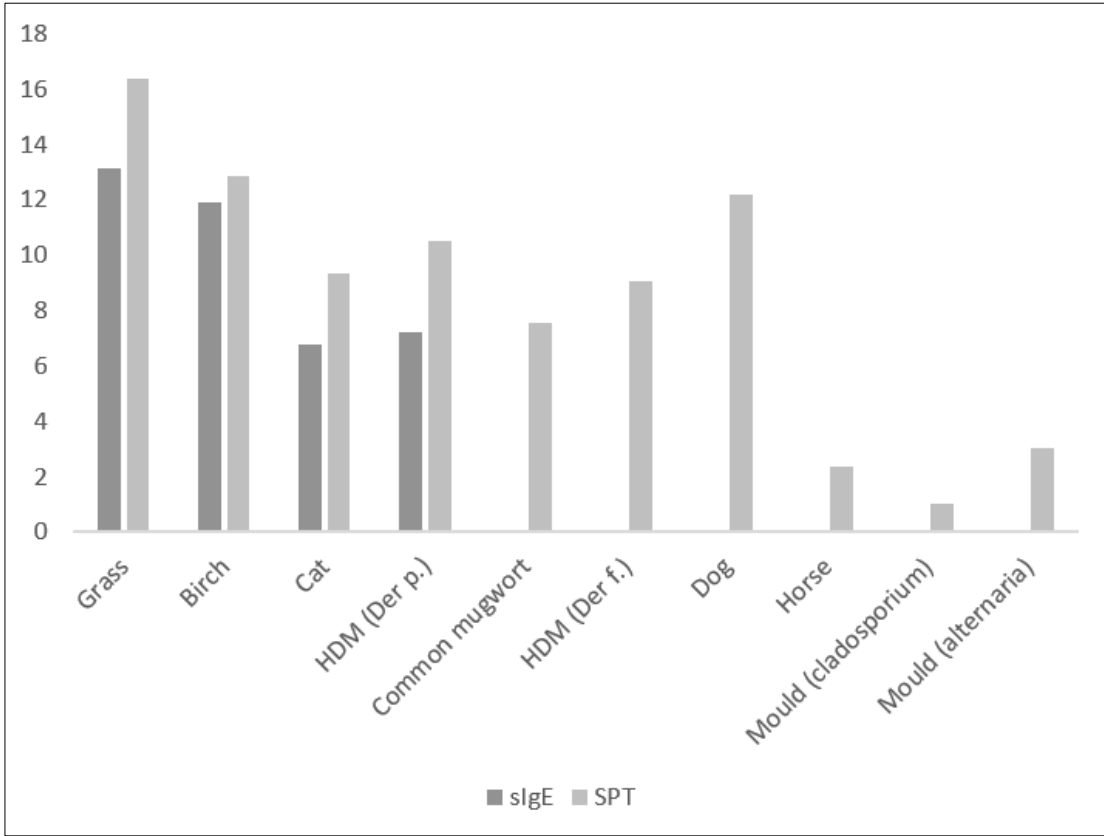
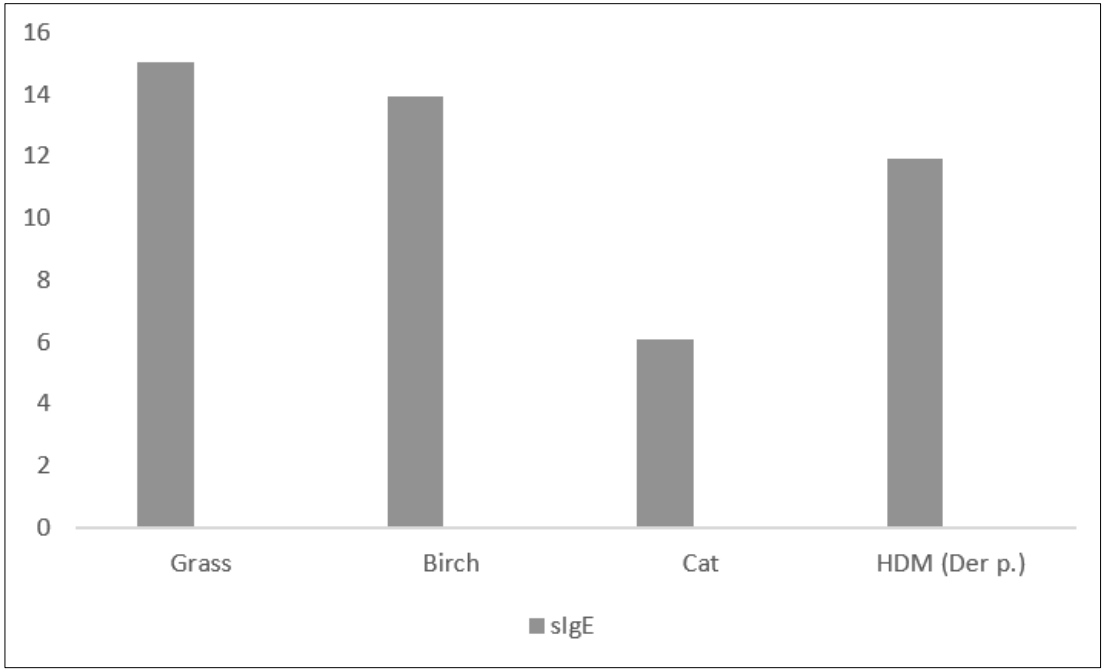


Table note: a positive SPT is defined as a mean wheal diameter ≥ 3 mm. Abbreviations: HDM = house dust mite. Der p. = Dermatophagoides pteronyssinus. Der f. = Dermatophagoides farinae.

Figure S2: Sensitisation patterns in DanFunD (2012–2015) measured by sIgE



Abbreviations: HDM = house dust mite. Der p. = Dermatophagoides pteronyssinus.

Figure S3: Positive predictive values (PPVs) for a range of values of the underlying prevalence. PPVs are based on the algorithm ≥ 1 INC and ≥ 1 antihistamine within 365 days in 2006–2008.

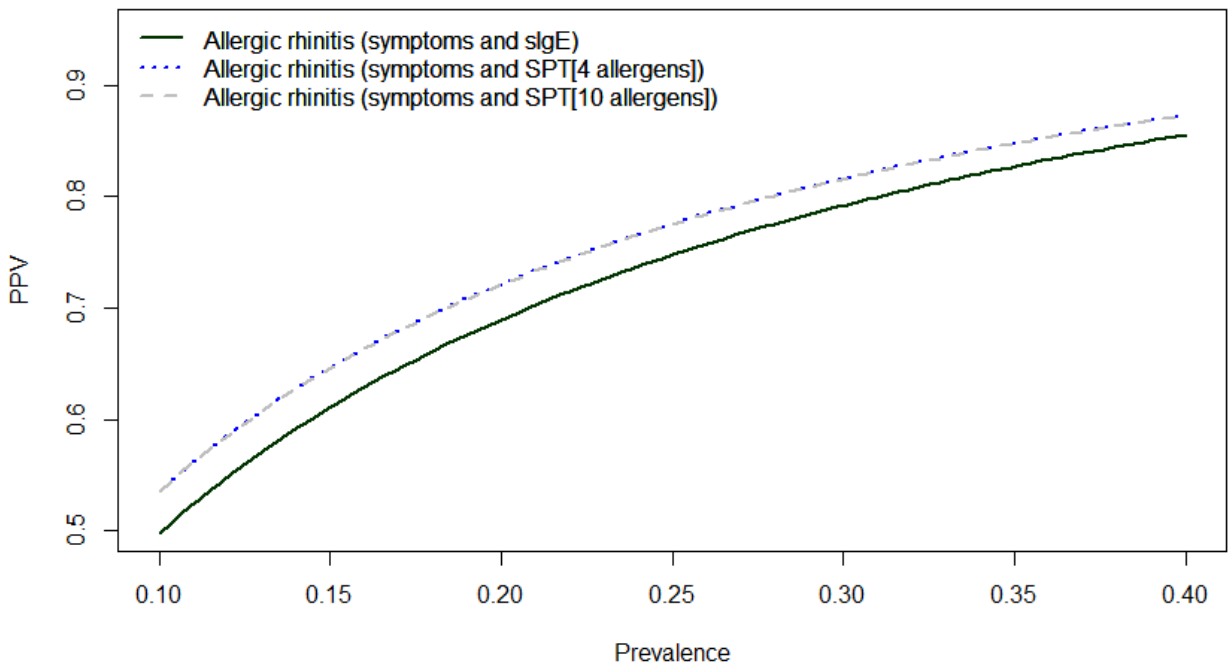


Figure note: a positive SPT is defined as a mean wheal diameter ≥ 3 mm. Abbreviations: INC = intranasal corticosteroids.