Allergic diseases and long-term risk of autoimmune disorders

Krishna, Mamidipudi; Subramanian, Anuradhaa; Adderley, Nicola; Zemedikun, Dawit; Gkoutos, Georgios; Nirantharakumar, Krishnarajah

DOI:
10.1183/13993003.00476-2019

Allergic Diseases and Long Term Risk of Autoimmune Disorders: *longitudinal cohort study and cluster analysis*

Mamidipudi Thirumala Krishna*# PhD FRCP, Anuradhaa Subramanian MSc, Nicola J Adderley PhD**, Dawit Zemedikun** MSc, Georgios V. Gkoutos+$ PhD, Krishnarajah Nirantharakumar**$MD MRCP**

**Department of Allergy and Immunology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK*
#Institute of Immunology & Immunotherapy, University of Birmingham, UK
***Institute of Applied Health Research, University of Birmingham, UK*
+Institute of Cancer and Genomics, University of Birmingham, UK
$Midlands Health Data Research UK, Birmingham, UK

Short title: ***Allergic diseases and autoimmune disorders – association and clustering***

Address for correspondence:
Professor M. T. Krishna
Department of Allergy and Immunology
Birmingham Heartlands Hospital
Bordesley Green East
Birmingham B9 5SS
Tel: +44 121 4241807
Fax: +44 121 4243229
Email: mtkrishna@yahoo.com

Abstract

Background

The association between allergic diseases (ADs) and autoimmune disorders (AIDs) is not well established.

Objective

To determine incidence rates of AIDs in allergic rhinitis/conjunctivitis (ARC), atopic eczema (AE) and asthma, and investigate for co-occurring patterns.

Methods

Design: Retrospective cohort study (1990-2018) employing 'The Health Improvement Network' (UK primary care database).
Exposure group: ARC, AE and asthma - all ages.
Controls: For each exposed patient, up to 2 randomly selected age- and gender-matched controls with no documented AD.
Adjusted incidence rate ratios (aIRRs) were calculated using Poisson regression. A cross-sectional study was also conducted employing Association Rule Mining (ARM) to investigate disease clusters.

Results

782,320, 1,393,570 and 1,049,868 patients with ARC, AE and asthma, respectively, were included. aIRRs of systemic lupus erythematosus (SLE), Sjogren's syndrome (SS), vitiligo, rheumatoid arthritis (RA), psoriasis, pernicious anaemia, inflammatory bowel disease (IBD), coeliac disease (CD) and autoimmune thyroid disease were uniformly higher in the 3 ADs compared to controls. Specifically, aIRRs of SLE (1.45) and SS (1.88) were higher in ARC; SLE (1.44), SS (1.61) and myasthenia (1.56) higher in asthma; SLE (1.86), SS (1.48), vitiligo (1.54) and psoriasis (2.41) higher in AE.

There was no significant effect of the 3 ADs on multiple sclerosis and ARC and AE on myasthenia.

ARM: ADs clustered with multiple AIDs. Three age- and gender-related clusters were identified, with relatively complex pattern in females ≥55 years.

Conclusion

The long-term risk of AIDs are significantly higher in patients with ADs. ADs and AIDs show age- and gender-related clustering patterns.

Key words: allergic disease, autoimmune disorder, incidence, risk, clustering, multimorbidity, comorbidity

SUMMARY OF IMPORTANT FINDINGS OF THIS STUDY

This study has shown that the long-term risk of autoimmune disorders (AIDs) are significantly higher in patients with allergic diseases (ADs). ADs and AIDs co-occur and show an age- and gender-related clustering pattern.

List of Abbreviations:

AD: Allergic Disease

AE: Atopic Eczema

AID: Autoimmune Disorder

aIRR: adjusted incident rate ratio

ARC: Allergic rhinitis/allergic conjunctivitis

ARM: Association Rule Mining

BMI: Body Mass Index

CI: Confidence Interval

EMR: Electronic Medical Record

GP: General Practitioner

IQR: Interquartile Range

NHS: National Health Service

RA: Rheumatoid Arthritis

SD: Standard Deviation

SLE: Systemic Lupus Erythematosus

THIN: The Health Improvement Network

Introduction

Autoimmune disorders (AIDs) and allergic diseases (ADs) are characterised by an immune dysregulated state(1, 2). From an immunological viewpoint, it has been suggested that these conditions are largely polarised with disparate mechanisms, with AIDs and ADs primarily mediated by a T helper (Th)1 and Th2 cellular immune response respectively(3, 4). Both have a complex aetio-pathogenesis and occur as a consequence of a complex interplay between multiple genetic and environmental factors with other possible variables hitherto unknown(5).

Western countries were challenged with an 'allergy epidemic' during the last 3-4 decades resulting in a very high burden of allergic rhinitis, allergic conjunctivitis , atopic eczema (AE) and asthma(6, 7). In particular, the UK ranks as one of the highest with respect to prevalence of ADs worldwide(7). Interestingly, there has been a parallel increase in the incidence rates (IRs) of several AIDs including multiple sclerosis, inflammatory bowel disease, rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) during the same period(5). The biological plausibility and causality hypothesis have linked these observations to the possible involvement of common triggers such as infection, genetic foci, environmental factors, reduced biodiversity and birth by caesarean section(8). Furthermore, there is some emerging evidence regarding an enhanced risk of AIDs such as SLE, RA, Sjogren's syndrome, multiple sclerosis, coeliac disease and myasthenia gravis in patients with ADs and *vice versa(9-12).* Studies carried out in south Asian immigrants in western countries have shown higher rates of both ADs and AIDs, with a positive association with the duration of settlement away from their native country, and with an enhanced risk in their offspring(13-15). These studies have, however,

been limited by relatively small sample size, methodological issues, selection bias, focus on a single or limited number of ADs and/or AIDs and some by not accounting for important confounders such as smoking history. Nonetheless, the evidence generated from these studies have challenged the Th1/Th2 paradigm(4, 16).

The main aim of this study was to test the hypothesis that an underlying AD such as allergic rhinitis/allergic conjunctivitis (ARC), AE or asthma enhances the risk of development of an AID. In a large population-based study, we determined adjusted incidence rate ratios (aIRRs) of common AIDs in patients with and without ARC, AE and asthma and investigated for clustering patterns between ADs and AIDs.

Methods

Setting

Data was extracted from The Health Improvement Network (THIN), a database comprised of electronic medical records (EMR) from UK general practices using Vision software. It includes coded data recorded by the patient's general practitioner (GP) for clinical and management purposes, capturing every consultation. Symptoms and diagnoses are recorded using Read codes, a hierarchical coding system for structured storage information(17). Fifteen million patients at 787 GP practices across the UK have contributed data to THIN; there are approximately 3 million active patients comprising 6-7% of the population. The data extracted from THIN is generalisable to the UK population for major health conditions(18).

In order to ensure data quality and allowing sufficient time to document important covariates, general practices were included in this study only if they met the following criteria: (a) 12 months after the practice reported an acceptable mortality rate(19), and (b) 12 months after date the practice began using EMR, and study start date.

Study Design

Incidence of autoimmune disorders

A longitudinal retrospective cohort study was carried out (01 Jan'90 -17 Jan'18). Patients of all ages registered for at least 365 days prior to study entry were included.

The exposed cohort were defined as patients with a prevalent or incident diagnosis of any one or more of the three ADs of interest: asthma, ARC and AE. Food allergy was not included in our analysis, as the diagnostic coding may not be accurate in primary care records, unless patients had undergone a review by an allergy specialist. In a sensitivity analysis, the exposed cohort were defined as incident diagnoses only (newly diagnosed after registration with the practice). For patients with multiple ADs, the earliest diagnosis was considered as the index condition. Diagnosis of all ADs was indicated by the presence of a corresponding clinical (Read) code in the patient's medical records. Allergic rhinitis and Allergic conjunctivitis were treated as a single exposure group (i.e., ARC) due to high rates of co-occurrence and as their inflammatory responses in the mucosa and conjunctiva unfold similarly(20, 21).

For each exposed patient, up to 2 controls were randomly selected within the same GP practice from an age- (within 1 year) and gender-matched pool of patients with no AD diagnosis. The following method was used to randomly select matched controls: Patients with an AD were identified, and their order was shuffled by randomly permuting the patient list (following the Fisher Yates algorithm), using a linear congruential generator as the source of randomness. All permutations occur with equal likelihood. Following this, controls were selected. Shuffling ensures that all patients in the exposure group have an equal chance of being matched to a control in instances where that control could potentially be matched with more than one exposed patient. When the number of possible controls for a particular exposed patient exceeded the number required (i.e., 2), a linear congruential generator was used to generate a random number between 1 and the number of potential controls;

the potential control at the position of this random number was selected, and the process repeated for subsequent controls.

Outcomes were the following AIDs (considered separately): SLE, Sjogren's syndrome, vitiligo, RA, psoriasis, pernicious anaemia, myasthenia gravis, inflammatory bowel disease, coeliac disease, autoimmune thyroid disease and multiple sclerosis.

Index date for exposed patients was the date of diagnosis of the first AD recorded for newly diagnosed patients or date of study entry for patients with an existing diagnosis. Unexposed controls were assigned the same index date as their corresponding exposed patient to avoid immortal time bias(22). Exposed and control patients were followed from the index date until the earliest of the following end points: date of outcome recording (AID of interest), death date, date patient left the practice, date the practice ceased contributing to THIN, and study end date.

Read codes: Read code lists (e-Methods-1) for each condition were compiled as follows: relevant codes were extracted from a database containing all current and past version-2 Read codes; each code was reviewed and assessed for relevance; where available, code lists in previously published literature(23-26) were used to inform and validate code selection; in the case of uncertainty regarding the relevance of a code, consensus was reached in consultation with an experienced primary care physician.

Disease clusters

A cross-sectional study was performed with a census date of 01 January 2018. Patients of all ages who had been registered with the practice for at least 365 days were included.

Definitions of Covariates

Age, sex, body mass index (BMI), deprivation quintile, ethnicity and smoking status were included as covariates in the cohort analyses. For BMI and smoking status, the last value recorded prior to index date was used. BMI was categorised as <18.5 kg/m^2 (underweight), 18-25 kg/m^2 (normal weight), 25-30 kg/m^2 (overweight) and '>30 kg/m^2' (obese). Smoking status was categorised as non-smoker, previous smoker, and current smoker. Social deprivation was recorded as Townsend deprivation quintile(27). Missing BMI, deprivation quintile, ethnicity and smoking data were handled by employing a 'missing' category within the corresponding categorical variables.

Analysis

Incidence of AIDs

Baseline covariates were summarised for patients in the three separate exposure groups including asthma, AE and ARC, and their corresponding controls using appropriate descriptive statistics: mean (standard deviation, SD) and median (interquartile range, IQR) for continuous variables, and number (%) for categorical variables. Incidence rates of each AID in each of the three ADs were calculated by dividing the number of patients with a newly diagnosed AID (numerator) by the total number of person-years at risk (denominator) for the given AD. Crude incidence rate

ratios (IRR) and adjusted incidence rate ratios (aIRRs) and their corresponding 95% confidence intervals (CIs) were calculated using Poisson regression. Separate Poisson regression models were used for each AD/AID combination. All models were adjusted for the following covariates: age, sex, BMI, ethnicity, Townsend deprivation quintile and smoking status. Two-sided p-values were calculated; $p < 0.05$ was considered to indicate statistical significance. All analyses were performed in Stata IC version 14.

The study flow diagram is shown in Figure 1.

Disease clusters

Association Rule Mining (ARM) was used to assess how the three ADs associated with the 11 AIDs - as described by Zemedikun et al(28). This is a data mining technique used to investigate frequent co-occurring associations among variables in large databases. Association rules were created specifying ADs as antecedent and AIDs as consequent so that the reverse associations were not detected. Significance was measured using three parameters including 'support' (how frequently the disease combination appears in the dataset), 'confidence' (the conditional probability that a subject with the antecedent disease will also have the consequent condition) and 'lift' (the ratio of the observed support to that expected if the 2 events were independent). The minimum parameter threshold for 'support' and 'confidence' were set at 0.0005 to extract sufficient number of association rules. 'Lift', also referred to as the interestingness measure, measures the importance of a relationship. Specifically, a 'lift' of >1.0 indicates that the antecedent and consequent condition appear more often than expected, i.e., the antecedent (AD) exerts a positive effect on the occurrence of the consequent (AID). ARM analysis was performed on the full

cross-sectional dataset, and separately for age and sex restricted subgroups. ARM analysis was performed using R software, version 3.3.1.

Results

Incidence of AIDs

782,320, 1,393,570 and 1,049,868 patients with ARC, AE and asthma, respectively, together with their matched controls, were included in the analyses (Figure 1). Mean (SD) age at baseline was 36.4 (19.4), 29.8 (25.3) and 35.6 (21.3) years in patients with ARC, AE and asthma respectively; 47.1%, 45.9% and 47.9% were male, respectively. Baseline gender, smoking status, BMI, ethnicity and Townsend deprivation quintile were similar between patients with each AD and their respective controls; unexposed controls were slightly older than their corresponding exposure group (Table 1); this arose as a result of incomplete matching of some patients due to lack of availability of matched controls within the same practice. Patients with ADs had more AIDs at baseline.

In patients with ARC compared to patients with no diagnosed AD, aIRRs were: SLE 1.45 (95% CI 1.27-1.67), Sjogren's syndrome 1.88 (1.63-2.17), vitiligo 1.36 (1.25-1.47), RA 1.21 (1.15-1.28), psoriasis 1.38 (1.34-1.42), pernicious anaemia 1.30 (1.20-1.40), inflammatory bowel disease 1.33 (1.26-1.41), coeliac disease 1.39 (1.28-1.51), and autoimmune thyroid disease 1.18 (1.08-1.29) (Figure 2; e-Table 1).

In patients with AE, aIRRs were: SLE 1.86 (1.66-2.09), Sjogren's syndrome 1.48 (1.30-1.69), vitiligo 1.54 (1.44-1.64), RA 1.28 (1.22-1.34), psoriasis 2.41 (2.36-2.46), pernicious anaemia 1.25 (1.18-1.33), inflammatory bowel disease 1.53 (1.45-1.61), coeliac disease 1.41 (1.32-1.50), and autoimmune thyroid disease 1.13 (1.05-1.22) (Figure 2; e-Table 1).

In patients with asthma, aIRRs were: SLE 1.44 (1.27-1.62), Sjogren's syndrome 1.61 (1.41-1.84), vitiligo 1.24 (1.15-1.34), RA 1.44 (1.37-1.50), psoriasis 1.39 (1.36-1.43), pernicious anaemia 1.30 (1.22-1.38), myasthenia gravis 1.56 (1.29-1.88), inflammatory bowel disease 1.44 (1.37-1.51), coeliac disease 1.44 (1.34-1.55), and autoimmune thyroid disease 1.23 (1.14-1.33) (Figure 2; e-Table 1).

There was no statistically significant difference in incidence of multiple sclerosis in any of the ADs, or of myasthenia gravis in ARC and AE.

e-Figure-1 summarises adjusted hazard ratios of AIDs in combination of 1, 2 and 3 ADs. A sensitivity analysis including incident diagnoses of ADs only made little difference to the results (e-Table 2).

Disease clusters

In the ARM analysis of all ages and both sexes, all three ADs clustered with psoriasis, coeliac disease and inflammatory bowel disease. ARC and AE clustered with vitiligo; ARC and asthma clustered with RA and autoimmune thyroid disease; and asthma and AE clustered with pernicious anaemia. The associations between ARC and vitiligo, and AE and psoriasis were particularly notable, with lifts of 1.583 and 1.414 respectively, indicating that the probability of vitiligo is 58% greater in the presence of ARC, and the probability of psoriasis is 41% greater in the presence of AE. The latter association was strengthened when ARC, asthma or both were present (Figure 3; Table 2).

Subgroup analysis highlighted three age- and gender-related clusters, <30, 30-54 and ≥55 years, with a relatively complex pattern of association in females ≥55 years (e-Tables 3-8; e-Figures 2 and 3). The three ADs clustered with vitiligo, coeliac disease, psoriasis, pernicious anaemia, inflammatory bowel disease, autoimmune

thyroid disease and RA. In men, the strongest associations were between ARC and vitiligo (<30 years), and AE and psoriasis (>30 years). In women, the strongest associations were between asthma/other ADs and psoriasis or inflammatory bowel disease.

Discussion

Strengths: To our knowledge, this is the largest study conducted thus far to systematically investigate associations between each AD independently with a range of common AIDs and report multimorbidity and clustering patterns. Our data covers a wide geographical area of the UK and is generalisable to the British population(18). Importantly, our study was conducted in a primary care setting, thereby reducing selection bias. Patients with missing data for some covariates were included in order to maximise the study population and ensure generalisability.

Limitations: First, it involved retrospective data extraction. Second, multiple clinicians have contributed to the THIN database and differences in clinical practice, and changes in governance systems, service framework and clinical practice standards over 3 decades are possible confounders. Third, there may be lack of standardisation with respect to diagnosis and management criteria in such a large population study involving a wide geographical area. Fourth, the majority of disorders investigated are heterogeneous in nature and it was not within the scope of this study to characterise them. Fifth, atopic status and autoimmune serology for AIDs could not be captured from the THIN database. Specifically, our data cannot differentiate between allergic asthma and non-allergic asthma in the asthma cohort. Sixth, this study did not specifically investigate the effect of parental allergy.

Interpretation of data: This study showed higher aIRRs of common AIDs in patients with ARC, asthma and AE compared to the control (unexposed) population. Furthermore, ARM analysis applied on the entire study population showed significant clustering between ARC, asthma and AE and AIDs including vitiligo, coeliac disease, psoriasis, pernicious anaemia, inflammatory bowel disease, autoimmune thyroid disease and RA. Further subgroup analysis highlighted three age and gender-related

clusters – <30 years, 30-54 years and ≥55 years – with a relatively more complex pattern of association in older patients, females in particular. The clustering patterns could at least in part be influenced by known demographic patterns involving AIDs, but it is plausible that patients with co-existing ADs and AIDs represent a distinct phenotype with a specific genetic determinant/s.

Whilst the aIRRs of both organ-specific and systemic AIDs were uniformly higher in patients with all ADs, effect sizes were greatest in the AE cohort with the effect most pronounced for psoriasis, SLE, Sjogren's syndrome, inflammatory bowel disease, and vitiligo. Similarly, ARC was positively associated with all AIDs under consideration, but aIRRs were greatest for SLE and Sjogren's syndrome. A similar pattern was also noted in the asthma cohort with aIRRs greatest for Sjogren's syndrome, RA, myasthenia gravis, coeliac disease and inflammatory bowel disease. The effects on respective AIDs was more pronounced in the presence of allergic comorbidity (i.e., ≥2 ADs). Interestingly, aIRRs for multiple sclerosis were uniformly low across all ADs and this is in keeping with a recent study by Fakih et al(29).

Comparison with previous studies: Strong associations with multiple AIDs and asthma have been reported in patients with coeliac disease(30). Studies carried out using the National Health Insurance Research database in Taiwan reported a higher incidence of asthma in patients with Sjogren's syndrome(9). Similar associations were also described between the 3 ADs and SLE and our findings are in keeping with these observations(10). Another study from Taiwan showed a strong association between allergic conjunctivitis and myasthenia gravis(12). A further study involving 155,311 patients with allergic rhinitis, AE and asthma (considered in a single group) showed positive associations with SLE, RA, Sjogren's syndrome,

dermatomyositis/polymyositis and systemic sclerosis, with a particularly strong association with Sjogren's syndrome(11).

Biological plausibility: A number of common factors have been implicated in the aetio-pathogenesis of ADs and AIDs including reduced biodiversity, particularly of the gut microbiome, urbanisation, lifestyle factors, diet (including vitamin D status) and recently identified novel metabolic and immune dysregulatory pathways involving mitochondrial stress and intracellular and extracellular RNAs(8). A meta-analysis of 2 genome wide association studies(31) involving allergen sensitisation and self-reported allergy combined with a publicly available genome wide association study data on AIDs, reported commonalities with respect to susceptibility foci, genetic pathways and genomic regulatory sites. Candidate genes have been reported between ADs and AIDs (32-34). There is some evidence for an 'auto-reactive' immunological state in AE(35) and involvement of Th17 cells in ADs and AIDs(36, 37).

Conclusion: The risk of AID/s are significantly higher in patients with an underlying AD. ADs and AIDs co-occur with an age and gender-related clustering pattern. Further studies in well characterised patients involving genomics are needed to confirm these observations and unravel novel mechanisms to pave the way for novel biomarkers, precision and personalised medicine, and primary and secondary prevention strategies in ADs and AIDs.

STATEMENTS

Ethics approval

The THIN data collection scheme and research carried out using THIN data were approved by the NHS South-East Multicentre Research Ethics Committee in 2003. Under the terms of the approval, studies must undergo independent scientific review. Approval for this study was obtained from the Scientific Review Committee (for the use of THIN data) in August 2018 (SRC reference 18THIN64).

Funding

This study did not receive any external funds.

Acknowledgement

The authors thank Dr A P Huissoon and Dr R Baretto, University Hospitals Birmingham NHS Foundation Trust for their comments and feedback on the draft manuscript.

References

1. Akdis M, Akdis CA. Mechanisms of allergen-specific immunotherapy: multiple suppressor factors at work in immune tolerance to allergens. J Allergy Clin Immunol. 2014;133(3):621-31.
2. Buckley CD, McGettrick HM. Leukocyte trafficking between stromal compartments: lessons from rheumatoid arthritis. Nat Rev Rheumatol. 2018.
3. Robinson DS. T-cell cytokines: what we have learned from human studies. Paediatr Respir Rev. 2004;5 Suppl A:S53-8.
4. Romagnani S. Immunologic influences on allergy and the TH1/TH2 balance. J Allergy Clin Immunol. 2004;113(3):395-400.
5. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med. 2002;347(12):911-20.
6. Platts-Mills TA. The allergy epidemics: 1870-2010. J Allergy Clin Immunol. 2015;136(1):3-13.
7. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Lancet. 1998;351(9111):1225-32.
8. Garn H, Bahn S, Baune BT, Binder EB, Bisgaard H, Chatila TA, et al. Current concepts in chronic inflammatory diseases: Interactions between microbes, cellular metabolism, and inflammation. J Allergy Clin Immunol. 2016;138(1):47-56.
9. Shen TC, Chen HJ, Wei CC, Chen CH, Tu CY, Hsia TC, et al. Risk of asthma in patients with primary Sjogren's syndrome: a retrospective cohort study. BMC Pulm Med. 2016;16(1):152.
10. Hsiao YP, Tsai JD, Muo CH, Tsai CH, Sung FC, Liao YT, et al. Atopic diseases and systemic lupus erythematosus: an epidemiological study of the risks and correlations. Int J Environ Res Public Health. 2014;11(8):8112-22.
11. Hou YC, Hu HY, Liu IL, Chang YT, Wu CY. The risk of autoimmune connective tissue diseases in patients with atopy: A nationwide population-based cohort study. Allergy Asthma Proc. 2017;38(5):383-9.
12. Yeh JH, Kuo HT, Chen HJ, Chen YK, Chiu HC, Kao CH. Higher risk of myasthenia gravis in patients with thyroid and allergic diseases: a national population-based study. Medicine (Baltimore). 2015;94(21):e835.
13. Haahtela T, Holgate S, Pawankar R, Akdis CA, Benjaponpitak S, Caraballo L, et al. The biodiversity hypothesis and allergic disease: world allergy organization position statement. World Allergy Organ J. 2013;6(1):3.
14. Cabieses B, Uphoff E, Pinart M, Anto JM, Wright J. A systematic review on the development of asthma and allergic diseases in relation to international immigration: the leading role of the environment confirmed. PLoS One. 2014;9(8):e105347.
15. Hussen HI, Moradi T, Persson M. The risk of type 1 diabetes among offspring of immigrant mothers in relation to the duration of residency in Sweden. Diabetes Care. 2015;38(5):934-6.
16. Romagnani S. The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? Immunology. 2004;112(3):352-63.
17. Booth N. What are the Read Codes? Health Libr Rev. 1994;11(3):177-82.
18. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care. 2011;19(4):251-5.
19. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiol Drug Saf. 2009;18(1):76-83.
20. Perkin MR, Bader T, Rudnicka AR, Strachan DP, Owen CG. Inter-Relationship between Rhinitis and Conjunctivitis in Allergic Rhinoconjunctivitis and Associated Risk Factors in Rural UK Children. PLoS One. 2015;10(11):e0143651.
21. Williams DC, Edney G, Maiden B, Smith PK. Recognition of allergic conjunctivitis in patients with allergic rhinitis. World Allergy Organ J. 2013;6(1):4.

22. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ. 2010;340:b5087.
23. Al Sallakh MA, Vasileiou E, Rodgers SE, Lyons RA, Sheikh A, Davies GA. Defining asthma and assessing asthma outcomes using electronic health record data: a systematic scoping review. Eur Respir J. 2017;49(6).
24. Nissen F, Quint JK, Wilkinson S, Mullerova H, Smeeth L, Douglas IJ. Validation of asthma recording in electronic health records: a systematic review. Clin Epidemiol. 2017;9:643-56.
25. Dizon MP, Yu AM, Singh RK, Wan J, Chren MM, Flohr C, et al. Systematic review of atopic dermatitis disease definition in studies using routinely collected health data. Br J Dermatol. 2018;178(6):1280-7.
26. Ban L, Langan SM, Abuabara K, Thomas KS, Abdul Sultan A, Sach T, et al. Incidence and sociodemographic characteristics of eczema diagnosis in children: A cohort study. J Allergy Clin Immunol. 2018;141(5):1927-9 e8.
27. Adams J, Ryan V, White M. How accurate are Townsend Deprivation Scores as predictors of self-reported health? A comparison with individual level data. J Public Health (Oxf). 2005;27(1):101-6.
28. Zemedikun DT, Gray LJ, Khunti K, Davies MJ, Dhalwani NN. Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data. Mayo Clin Proc. 2018;93(7):857-66.
29. Fakih R, Diaz-Cruz C, Chua AS, Gonzalez C, Healy BC, Sattarnezhad N, et al. Food allergies are associated with increased disease activity in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2018.
30. Assa A, Frenkel-Nir Y, Tzur D, Katz LH, Shamir R. Large population study shows that adolescents with celiac disease have an increased risk of multiple autoimmune and nonautoimmune comorbidities. Acta Paediatr. 2017;106(6):967-72.
31. Kreiner E, Waage J, Standl M, Brix S, Pers TH, Couto Alves A, et al. Shared genetic variants suggest common pathways in allergy and autoimmune diseases. J Allergy Clin Immunol. 2017;140(3):771-81.
32. Jimenez-Morales S, Velazquez-Cruz R, Ramirez-Bello J, Bonilla-Gonzalez E, Romero-Hidalgo S, Escamilla-Guerrero G, et al. Tumor necrosis factor-alpha is a common genetic risk factor for asthma, juvenile rheumatoid arthritis, and systemic lupus erythematosus in a Mexican pediatric population. Hum Immunol. 2009;70(4):251-6.
33. Becker KG, Barnes KC. Underlying disease specificity of genetic loci in atopic dermatitis. J Invest Dermatol. 2001;117(5):1325-7.
34. Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, et al. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. Nat Genet. 2018;50(1):42-53.
35. Tang TS, Bieber T, Williams HC. Does "autoreactivity" play a role in atopic dermatitis? J Allergy Clin Immunol. 2012;129(5):1209-15 e2.
36. Buehler U, Schulenburg K, Yurugi H, Solman M, Abankwa D, Ulges A, et al. Targeting prohibitins at the cell surface prevents Th17-mediated autoimmunity. EMBO J. 2018.
37. Sorbello V, Ciprandi G, Di Stefano A, Massaglia GM, Favata G, Conticello S, et al. Nasal IL-17F is related to bronchial IL-17F/neutrophilia and exacerbations in stable atopic severe asthma. Allergy. 2015;70(2):236-40.

Figure titles and legends:

Figure 1: Study flow diagram
Legend: Flow diagram summarising study design

Figure 2: Adjusted Incident Rate Ratios of AIDs in ADs
Legend: Forest plots summarising aIRRs of AIDs in each AD (incident and prevalent cases).

Figure 3: ARM analysis of the entire study population.
Legend: This figure should be reviewed in conjunction with Table-2. The figure illustrates clustering and association between ADs and AIDs. Each arrow connects an AD (antecedent) with an AID (consequent) *via* circles (or nodes). The size and the intensity of colour within each circle represent the level of 'support' and level of 'lift' respectively, i.e., larger the circle and darker the colour, greater is the 'support' and 'lift' respectively. For example, there is clustering between asthma, AE and ARC and psoriasis; asthma and ARC (but not AE) with RA; AE and ARC (but not asthma) with vitiligo. This suggests that these ADs have a positive effect on the occurrence of the respective AIDs.

e-Figure-1: Adjusted hazard ratios of AIDs in combinations of ADs.
Legend: Forest plot summarising adjusted hazard ratios (aHR) of AIDs in combinations of 1, 2 or 3 ADs. The figure illustrates an overall trend towards higher aHRs for respective AIDs in the presence of ≥2 ADs. aHRs calculated using Cox proportional hazards regression.

e-Figure 2: Sub-group ARM analysis by age in males.
Legend: To be reviewed in conjunction with e-Tables 3-5. Each arrow connects an AD (antecedent) with an AID (consequent) *via* circles (or nodes). The size and the intensity of colour within each circle represent the level of 'support' and level of 'lift' respectively, i.e., larger the circle and darker the colour, greater is the 'support' and 'lift' respectively.

e-Figure 3: Sub-group ARM analysis by age in females.
Legend: To be reviewed in conjunction with e-Tables 6-8. Each arrow connects an AD (antecedent) with an AID (consequent) *via* circles (or nodes). The size and the intensity of colour within each circle represent the level of 'support' and level of 'lift' respectively, i.e., larger the circle and darker the colour, greater is the 'support' and 'lift' respectively.

Table 1: Baseline patient characteristics

	Allergic rhinitis/conjunctivitis		Atopic eczema		Asthma	
Characteristics	**Exposed**	**Unexposed**	**Exposed**	**Unexposed**	**Exposed**	**Unexposed**
All	**(n=782,320)**	**(n=1,291,550)**	**(n=1,393,570)**	**(n=2,170,618)**	**(n=1,049,868)**	**(n=1,732,480)**
Age at allergy diagnosis, mean (SD)	29.89 (19.74)		24.25 (25.29)		25.09 (21.95)	
Age, mean (SD)	36.44 (19.44)	38.14 (19.11)	29.77 (25.30)	32.74 (25.45)	35.61 (21.26)	37.98 (20.91)
Age Categories, n (%)						
0 - 9	56995 (7.29)	78503 (6.08)	443272 (31.81)	602343 (27.75)	110635 (10.54)	142078 (8.20)
10 - 19	97289 (12.44)	125135 (9.69)	152580 (10.95)	191030 (8.80)	146137 (13.92)	184127 (10.63)
20 - 29	156848 (20.05)	247795 (19.19)	176608 (12.67)	277761 (12.80)	218329 (20.80)	356637 (20.59)
30 - 39	165941 (21.21)	292551 (22.65)	157265 (11.29)	272643 (12.56)	178980 (17.05)	323138 (18.65)
40 - 49	115694 (14.79)	209110 (16.19)	126174 (9.05)	226457 (10.43)	126552 (12.05)	234123 (13.51)
50 - 59	78520 (10.04)	141956 (10.99)	109781 (7.88)	198495 (9.14)	98207 (9.35)	182006 (10.51)
60 - 69	58414 (7.47)	103498 (8.01)	96955 (6.96)	171496 (7.90)	81459 (7.76)	147338 (8.50)
70 - max	52619 (6.73)	93002 (7.20)	130935 (9.40)	230393 (10.61)	89569 (8.53)	163033 (9.41)
Sex						
Male	368819 (47.14)	632521 (48.97)	639277 (45.87)	1002759 (46.20)	502841 (47.90)	844310 (48.73)
Female	413501 (52.86)	659029 (51.03)	754293 (54.13)	1167859 (53.80)	547027 (52.10)	888170 (51.27)
BMI, mean (SD)	25.94 (5.26)	25.82 (5.31)	26.14 (5.43)	25.98 (5.33)	26.74 (6.01)	25.83 (5.35)
BMI Categories, n (%)						
Underweight (<18.5)	13778 (1.76)	24658 (1.91)	18525 (1.33)	32012 (1.47)	18722 (1.78)	32606 (1.88)
Normal weight (18.5-25)	240550 (30.75)	391963 (30.35)	296729 (21.29)	485257 (22.36)	273231 (26.03)	488611 (28.20)
Overweight (25-30)	167639 (21.43)	260249 (20.15)	215020 (15.43)	345277 (15.91)	205041 (19.53)	323882 (18.69)
Obese (>30)	95186 (12.17)	148894 (11.53)	131161 (9.41)	200892 (9.26)	156489 (14.91)	189050 (10.91)
Missing	265167 (33.89)	465786 (36.06)	732135 (52.54)	1107180 (51.01)	396385 (37.76)	698331 (40.31)
Smoker categories, n (%)						
Non-smoker	393523 (50.30)	581795 (45.05)	472948 (33.94)	762193 (35.11)	454001 (43.24)	735619 (42.46)
Previous smoker	103476 (13.23)	158144 (12.24)	151830 (10.90)	228211 (10.51)	153849 (14.65)	198447 (11.45)
Current smoker	117214 (14.98)	255481 (19.78)	181665 (13.04)	304165 (14.01)	202006 (19.24)	329486 (19.02)
Missing	168107 (21.49)	296130 (22.93)	587127 (42.13)	876049 (40.36)	240012 (22.86)	468928 (27.07)
Ethnicity, n (%)						
Caucasian	276634 (35.36)	419125 (32.45)	501207 (35.97)	737165 (33.96)	398059 (37.92)	543068 (31.35)
Black Afro-Caribbean	16551 (2.12)	18529 (1.43)	19286 (1.38)	26148 (1.20)	9608 (0.92)	18120 (1.05)

South Asian	23375 (2.99)	27075 (2.10)	33686 (2.42)	38714 (1.78)	15577 (1.48)	25656 (1.48)
Mixed Race/Chinese/Other	15467 (1.98)	20857 (1.61)	22697 (1.63)	33153 (1.53)	11179 (1.06)	22258 (1.28)
Missing	450293 (57.56)	805964 (62.40)	816694 (58.60)	1335438 (61.52)	615445 (58.62)	1123378 (64.84)
Townsend, n (%)						
1	173514 (22.18)	267708 (20.73)	298714 (21.44)	446876 (20.59)	189932 (18.09)	332379 (19.19)
2	143591 (18.35)	230699 (17.86)	253849 (18.22)	391908 (18.06)	178000 (16.95)	307500 (17.75)
3	140876 (18.01)	238640 (18.48)	252724 (18.14)	398279 (18.35)	193512 (18.43)	324160 (18.71)
4	123684 (15.81)	215294 (16.67)	227134 (16.30)	362669 (16.71)	189185 (18.02)	298135 (17.21)
5	87912 (11.24)	154989 (12.00)	160308 (11.50)	256742 (11.83)	189185 (13.33)	212347 (12.26)
Missing	112743 (14.41)	184220 (14.26)	200841 (14.41)	314144 (14.47)	159252 (15.17)	257959 (14.89)
Autoimmune disorders at baseline, n (%)						
SLE	791 (0.10)	1195 (0.09)	1445 (0.10)	1610 (0.07)	1064 (0.10)	1514 (0.09)
Sjogrens	516 (0.07)	511 (0.04)	693 (0.05)	813 (0.04)	572 (0.05)	735 (0.04)
Vitiligo	2191 (0.28)	2516 (0.19)	3532 (0.25)	3519 (0.16)	1996 (0.19)	2895 (0.17)
Rheumatoid arthritis	3819 (0.49)	6122 (0.47)	6933 (0.50)	10321 (0.48)	6310 (0.60)	8576 (0.50)
Psoriasis	20077 (2.57)	28648 (2.22)	40319 (2.89)	39147 (1.80)	24362 (2.32)	35489 (2.05)
Pernicious anaemia	1458 (0.19)	2095 (0.16)	2879 (0.21)	4101 (0.19)	2464 (0.23)	3166 (0.18)
Myasthenia gravis	180 (0.02)	256 (0.02)	251 (0.02)	460 (0.02)	273 (0.03)	349 (0.02)
Inflammatory bowel disease	5003 (0.64)	7030 (0.54)	7848 (0.56)	9537 (0.44)	6428 (0.61)	8843 (0.51)
Coeliac disease	1824 (0.23)	2141 (0.17)	2876 (0.21)	3221 (0.15)	2398 (0.23)	2770 (0.16)
Autoimmune thyroid disease	2069 (0.26)	2586 (0.20)	2528 (0.18)	3456 (0.16)	2291 (0.22)	3215 (0.19)
Multiple sclerosis	1479 (0.19)	2474 (0.19)	2204 (0.16)	3439 (0.16)	1519 (0.14)	3147 (0.18)

Table 2: Cluster analysis: association rules of ADs and AIDs in the full cross-sectional dataset (in order of lift).

Rule	Antecedent		Consequent	Support	Confidence	Lift
1	{ConjunctivitisRhinitis}	=>	{Vitiligo}	0.0006	0.0050	1.5835
2	{Eczema,ConjunctivitisRhinitis}	=>	{Psoriasis}	0.0020	0.0555	1.5547
3	{Asthma,Eczema,ConjunctivitisRhinitis}	=>	{Psoriasis}	0.0007	0.0524	1.4654
4	{Asthma,Eczema}	=>	{Psoriasis}	0.0018	0.0512	1.4340
5	{Eczema}	=>	{Psoriasis}	0.0090	0.0505	1.4141
6	{Eczema}	=>	{Vitiligo}	0.0008	0.0044	1.4019
7	{ConjunctivitisRhinitis}	=>	{Coeliac}	0.0006	0.0054	1.3910
8	{ConjunctivitisRhinitis}	=>	{AutoimmuneThyroid}	0.0006	0.0048	1.3427
9	{Asthma}	=>	{Coeliac}	0.0007	0.0052	1.3386
10	{Asthma}	=>	{IBD}	0.0014	0.0110	1.3114
11	{ConjunctivitisRhinitis}	=>	{IBD}	0.0013	0.0109	1.2968
12	{Asthma}	=>	{PerniciousAnaemia}	0.0006	0.0044	1.2939
13	{Asthma,ConjunctivitisRhinitis}	=>	{Psoriasis}	0.0015	0.0459	1.2857
14	{Eczema}	=>	{Coeliac}	0.0009	0.0049	1.2785
15	{Asthma}	=>	{RheumatoidArthritis}	0.0012	0.0091	1.2768
16	{Asthma}	=>	{AutoimmuneThyroid}	0.0006	0.0045	1.2527
17	{ConjunctivitisRhinitis}	=>	{Psoriasis}	0.0052	0.0443	1.2403
18	{Asthma}	=>	{Psoriasis}	0.0054	0.0423	1.1834
19	{Eczema}	=>	{IBD}	0.0017	0.0095	1.1381
20	{ConjunctivitisRhinitis}	=>	{RheumatoidArthritis}	0.0009	0.0078	1.0945
21	{Eczema}	=>	{PerniciousAnaemia}	0.0006	0.0035	1.0136

Figure-1

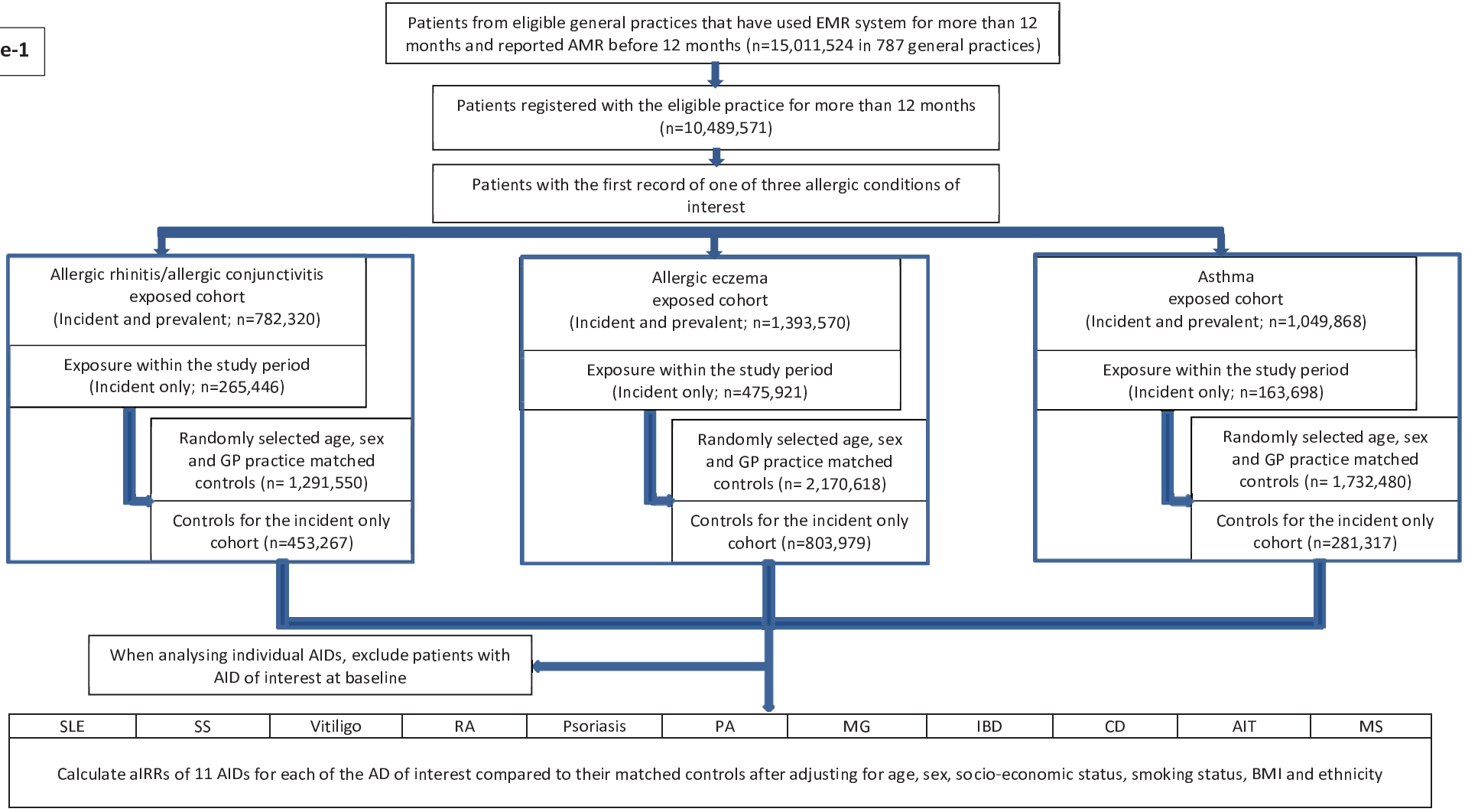

Patients from eligible general practices that have used EMR system for more than 12 months and reported AMR before 12 months (n=15,011,524 in 787 general practices)
Patients registered with the eligible practice for more than 12 months (n=10,489,571)
Patients with the first record of one of three allergic conditions of interest
Allergic rhinitis/allergic conjunctivitis exposed cohort (Incident and prevalent; n=782,320)
Exposure within the study period (Incident only; n=265,446)
Randomly selected age, sex and GP practice matched controls (n= 1,291,550)
Controls for the incident only cohort (n=453,267)
Allergic eczema exposed cohort (Incident and prevalent; n=1,393,570)
Exposure within the study period (Incident only; n=475,921)
Randomly selected age, sex and GP practice matched controls (n= 2,170,618)
Controls for the incident only cohort (n=803,979)
Asthma exposed cohort (Incident and prevalent; n=1,049,868)
Exposure within the study period (Incident only; n=163,698)
Randomly selected age, sex and GP practice matched controls (n= 1,732,480)
Controls for the incident only cohort (n=281,317)
When analysing individual AIDs, exclude patients with AID of interest at baseline
SLE
SS
Vitiligo
RA
Psoriasis
PA
MG
IBD
CD
AIT
MS
Calculate aIRRs of 11 AIDs for each of the AD of interest compared to their matched controls after adjusting for age, sex, socio-economic status, smoking status, BMI and ethnicity

Figure-2

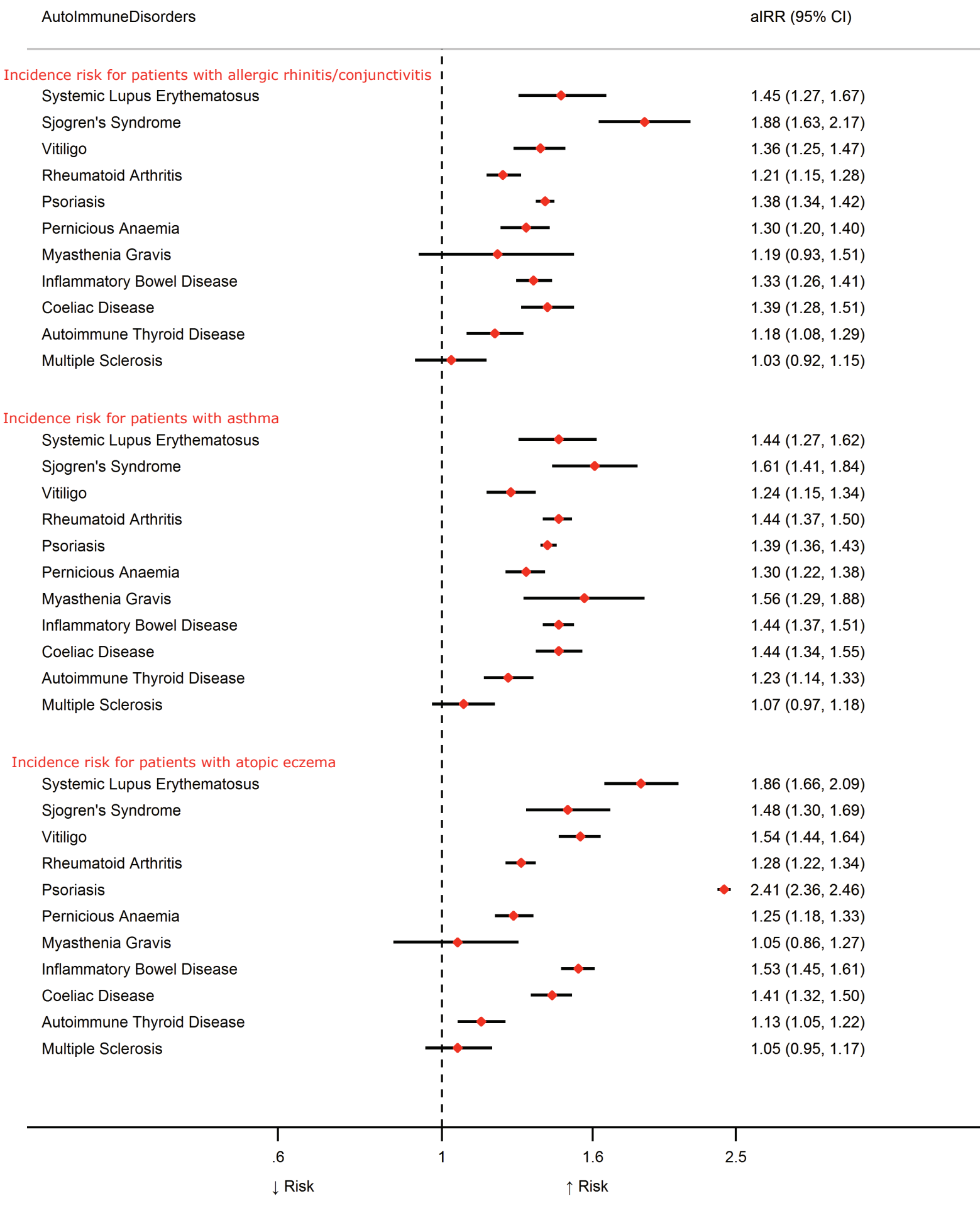
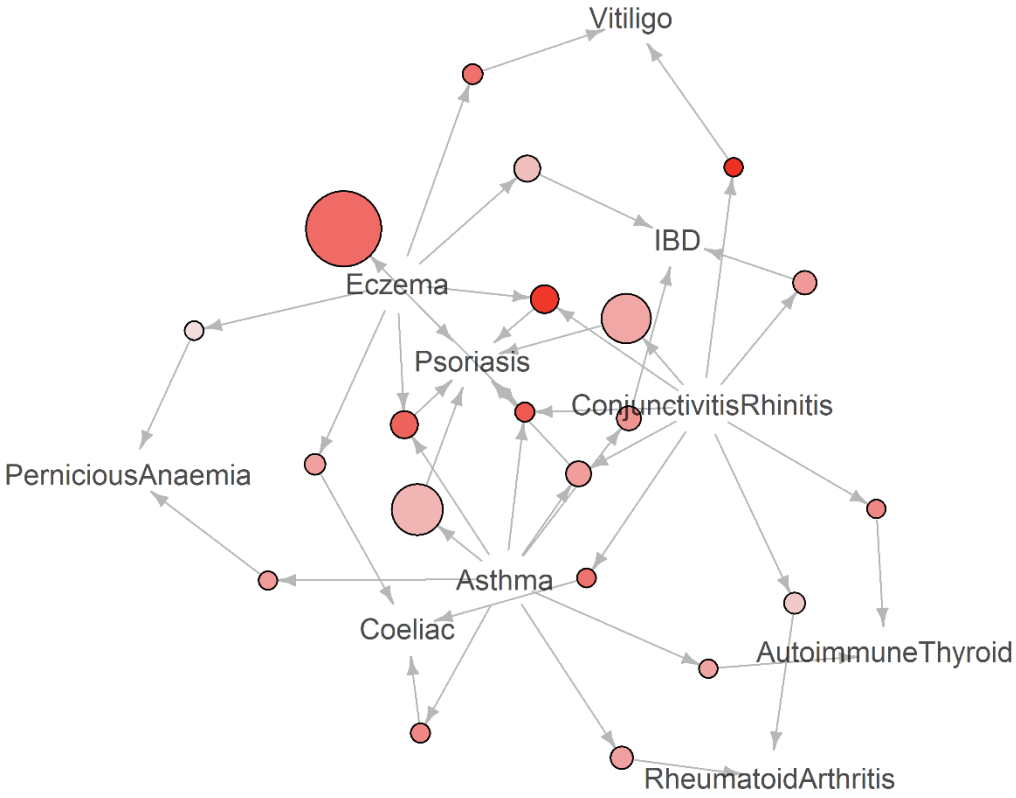

AutoImmuneDisorders
aIRR (95% CI)
Incidence risk for patients with allergic rhinitis/conjunctivitis
Systemic Lupus Erythematosus
1.45 (1.27, 1.67)
Sjogren's Syndrome
1.88 (1.63, 2.17)
Vitiligo
1.36 (1.25, 1.47)
Rheumatoid Arthritis
1.21 (1.15, 1.28)
Psoriasis
1.38 (1.34, 1.42)
Pernicious Anaemia
1.30 (1.20, 1.40)
Myasthenia Gravis
1.19 (0.93, 1.51)
Inflammatory Bowel Disease
1.33 (1.26, 1.41)
Coeliac Disease
1.39 (1.28, 1.51)
Autoimmune Thyroid Disease
1.18 (1.08, 1.29)
Multiple Sclerosis
1.03 (0.92, 1.15)
Incidence risk for patients with asthma
Systemic Lupus Erythematosus
1.44 (1.27, 1.62)
Sjogren's Syndrome
1.61 (1.41, 1.84)
Vitiligo
1.24 (1.15, 1.34)
Rheumatoid Arthritis
1.44 (1.37, 1.50)
Psoriasis
1.39 (1.36, 1.43)
Pernicious Anaemia
1.30 (1.22, 1.38)
Myasthenia Gravis
1.56 (1.29, 1.88)
Inflammatory Bowel Disease
1.44 (1.37, 1.51)
Coeliac Disease
1.44 (1.34, 1.55)
Autoimmune Thyroid Disease
1.23 (1.14, 1.33)
Multiple Sclerosis
1.07 (0.97, 1.18)
Incidence risk for patients with atopic eczema
Systemic Lupus Erythematosus
1.86 (1.66, 2.09)
Sjogren's Syndrome
1.48 (1.30, 1.69)
Vitiligo
1.54 (1.44, 1.64)
Rheumatoid Arthritis
1.28 (1.22, 1.34)
Psoriasis
2.41 (2.36, 2.46)
Pernicious Anaemia
1.25 (1.18, 1.33)
Myasthenia Gravis
1.05 (0.86, 1.27)
Inflammatory Bowel Disease
1.53 (1.45, 1.61)
Coeliac Disease
1.41 (1.32, 1.50)
Autoimmune Thyroid Disease
1.13 (1.05, 1.22)
Multiple Sclerosis
1.05 (0.95, 1.17)
.6
1
1.6
2.5
↓ Risk
↑ Risk

Figure 3


Vitiligo
IBD
Eczema
Psoriasis
ConjunctivitisRhinitis
PerniciousAnaemia
Asthma
Coeliac
AutoimmuneThyroid
RheumatoidArthritis