

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/160946/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Conrad, Nathalie, Misra, Shivani, Verbakel, Jan Y, Verbeke, Geert, Molenberghs, Geert, Taylor, Peter N, Mason, Justin, Sattar, Naveed, McMurray, John J V, McInnes, Iain B, Khunti, Kamlesh and Cambridge, Geraldine 2023. Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. *The Lancet* 401 (10391), pp. 1878-1890. 10.1016/S0140-6736(23)00457-9

Publishers page: [https://doi.org/10.1016/S0140-6736\(23\)00457-9](https://doi.org/10.1016/S0140-6736(23)00457-9)

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1 **Incidence, prevalence and co-occurrence of autoimmune disorders, trends over time and by age, sex and**
2 **socioeconomic status. A population-based study in 22 million individuals.**

3

4 Nathalie Conrad^{1,2,3} PhD, Shivani Misra⁴ PhD, Jan Y Verbakel^{1,5} PhD, Prof. Geert Verbeke⁶ PhD, Prof. Geert
5 Molenberghs⁶ PhD, Peter N Taylor⁷ PhD, Prof. Justin Mason⁸ PhD, PhD, Prof. Naveed Sattar⁹ MD, Prof. John J.
6 V. McMurray² MD, Prof. Iain B. McInnes⁹ PhD, Prof. Kamlesh Khunti¹⁰ MD, Prof. Geraldine Cambridge¹¹ PhD

7

8 1 Department of Public Health and Primary Care, KU Leuven, Belgium

9 2 Institute of Cardiovascular and Medical Sciences, University of Glasgow, United Kingdom

10 3 Deep Medicine, Nuffield Department of Women's and Reproductive Health, University of Oxford, United
11 Kingdom

12 4 Faculty of Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College London,
13 United Kingdom

14 5 Nuffield Department of Primary Care Health Sciences, University of Oxford, United Kingdom

15 6 Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), Hasselt University and KU
16 Leuven, Belgium

17 7 Division of Infection and Immunity, School of Medicine, Cardiff University, United Kingdom

18 8 Faculty of Medicine, National Heart & Lung Institute, Imperial College London, United Kingdom

19 9 College of Medical Veterinary and Life Sciences, University of Glasgow, United Kingdom

20 10 Diabetes Research Centre, University of Leicester, United Kingdom

21 11 Division of Medicine, Dept Rheumatology, University College London, United Kingdom

22

23

24

25

26 **Correspondence to:**

27 Dr. Nathalie Conrad

28 Department of Public Health and Primary Care

29 Kapucijnenvoer 33 - 3000 Leuven – Belgium

30 Tel: +44 74 70 421 005

31 E-mail: nathalie.conrad@kuleuven.be

32

33

34 Word count abstract: 417

35 Word count manuscript: 3 650

36

37

38 **Abstract**

39 **Background:** A rise in the incidence of selected autoimmune disorders has been described, however,
40 contemporary estimates of the overall incidence of autoimmune diseases and trends over time, are scarce
41 and inconsistent.

42 **Methods:** We used linked primary and secondary electronic health records of 22 million individuals from the
43 Clinical Practice Research Datalink (CPRD), a cohort that is representative of the UK population in terms of age
44 and sex. We calculated age-sex-standardised incidence and prevalence of 19 autoimmune disorders (AID)
45 from 2000 to 2019 and used negative binomial regression models to investigate temporal trends and variation
46 by age, sex, socioeconomic status, season of onset and geographical region. To characterise co-occurrence of
47 autoimmune diseases, we calculated incidence rate ratios, comparing incidence rates of comorbid
48 autoimmune disease among patients with a first (index) autoimmune disease with incidence rates in the
49 general population, using negative binomial regression models, adjusted for age and sex.

50 **Findings:** Among the 22,009,375 individuals included in the study, we identified a total of 978 872 patients
51 with a new diagnosis of at least one autoimmune disease between 2000 and 2019 (mean (SD) age: 54.0 (21.4)
52 years, 64% women). Over the study period, age-standardised incidence rates of autoimmune diseases
53 increased by 4%, similarly for men and women. The largest increases were seen in Graves' disease, coeliac
54 disease and Sjogren's syndrome, for which incidences have doubled over the past two decades. Two
55 conditions exhibited a significant decrease in incidence (Hashimoto's thyroiditis and pernicious anaemia).
56 Taken together the 19 autoimmune disorders examined affected 10.2% of the population over the study
57 period (13.1% of women, 7.4% of men). A socioeconomic gradient was evident across several diseases,
58 including Graves' disease, pernicious anaemia, rheumatoid arthritis, and systemic lupus erythematosus.
59 Seasonal variations were observed for type 1 diabetes (more commonly diagnosed in winter) and vitiligo
60 (more commonly diagnosed in summer), and regional variations were observed for a range of conditions.
61 Autoimmune disorders were commonly associated with each other, particularly Sjogren's, systemic lupus
62 erythematosus and systemic sclerosis. Patients with type 1 diabetes also had significantly higher rates of
63 Addison's, coeliac, and thyroid diseases, and multiple sclerosis stood out as having low rates of co-occurrence
64 with other autoimmune diseases.

65 **Interpretation:** Autoimmune diseases affect about one in ten individuals. Their burden continues to increase
66 over time, albeit at varying rates across individual diseases. The socioeconomic, seasonal, and regional
67 disparities observed among several autoimmune disorders, implicate environmental factors in disease
68 pathogenesis. The interrelations between autoimmune diseases are commensurate with shared pathogenetic
69 mechanisms or predisposing factors, particularly among connective tissue diseases and among endocrine
70 diseases.

71 **Funding:** Research Foundation Flanders (FWO).

72 **Keywords:** autoimmune disorders, immune mediated inflammatory diseases, incidence, prevalence, co-
73 occurrence, cohort study, epidemiology, CPRD.

74

75

76 **Research in Context**

77 **Evidence before this study**

78 We searched Pubmed and Embase for reports published between 1 January 2000 to 30 July 2022 related to
79 “autoimmune disorders” (any of the 19 individual conditions investigated) and “incidence”, reviewed
80 references of clinical practice guidelines and consulted with experts for relevant studies. Most studies
81 investigated one autoimmune disorder at a time, and generally the more common autoimmune disorders,
82 such as type 1 diabetes or psoriasis.

83 Studies generally relied on a small number of cases and presented different designs, case identification and
84 diagnostic methods, rendering adequate synthesis and the calculation of pooled estimates and temporal
85 trends difficult. Evidence was particularly scarce for rarer autoimmune disorders. We found no study that
86 reported large-scale disease incidence and temporal trends of autoimmune disorders as a group of conditions

87 **Added value of this study**

88 We present standardised incidence rates derived from a large, representative, general population cohort,
89 setting a baseline for international comparison, monitoring of prevention strategies and for the design of
90 public health policies. Temporal trends across a broad range of autoimmune diseases do not support the idea
91 of an epidemic of autoimmunity at present day, and provide valuable reference rates for future studies
92 investigating population-level impact of newly-introduced risk factors, such as the covid-19 pandemic.

93 We provide robust evidence of socioeconomic, seasonal, and regional disparities for several autoimmune
94 diseases (particularly Graves’ disease, pernicious anaemia, rheumatoid arthritis, systemic lupus
95 erythematosus, and type 1 diabetes). Such variations are unlikely to be attributable to genetic differences
96 alone and suggest involvement of potentially modifiable risk factors in the pathogenesis of autoimmune
97 diseases.

98 We further demonstrate important interrelations between many autoimmune diseases, and confirm that co-
99 occurrence of autoimmune disease is common, yet orders of magnitude differ widely between diseases.
100 Associations were highest amongst connective tissue diseases, for patients with type 1 diabetes and Addison’s
101 disease, coeliac and thyroid diseases. More generally, increased risk of developing Addison’s disease was
102 observed following almost every autoimmune disease investigated. Multiple sclerosis stood out as having low
103 rates of co-occurrence with other autoimmune diseases, suggesting a distinct pathophysiology.

104

105 **Implications of all the available evidence**

106 The burden of autoimmune disorders appears higher than previous estimates, and continues to increase over
107 time, albeit at varying rates across individual diseases.

108 Socioeconomic, seasonal, and regional disparities in disease incidence point to potentially preventable factors
109 involved in the pathogenesis of autoimmune diseases. Co-occurrences of diseases point to common genetic
110 and environmental risk factors that interact and operate variably across these diseases.

111 Further research is needed to fully elucidate pathophysiological mechanism underlying the associations
112 observed in this study.

113

114

115

116

117 Introduction

118 Autoimmune diseases arise when immune dysregulation causes host tissue damage.¹ A wide range of
119 autoimmune diseases are described that present with variable age of onset, tissue distribution and clinical
120 and functional impact.¹ Most of these diseases are incurable and require lifelong treatment.

121 Adequate public health and service delivery planning requires reliable information about contemporary
122 population-level disease incidence. However, estimates of autoimmune disease incidence rates and their
123 temporal trends, even in high-income countries, are scarce and inconsistent.¹ Selected autoimmune disorders,
124 such as type 1 diabetes, are reported to have increased over the past several decades, raising the question as
125 to whether the overall incidence of autoimmune disorders is on the rise, driven perhaps by common
126 environmental factors or behavioural changes. Even for type 1 diabetes, whose incidence is among the best
127 studied within autoimmune diseases, reports rely on relatively small cohorts,^{2,3} and estimates vary by a factor
128 10 between studies in Europe alone.^{4,5} For many other autoimmune diseases, evidence concerning disease
129 incidence and prevalence is more limited. The relatively modest absolute numbers of patients affected by
130 individual autoimmune diseases presents as a major challenge to investigators, and hinders adequate
131 synthesis across studies.⁶ As a result, reliable estimates of disease incidence and how they evolve over time,
132 particularly as pertains to autoimmune diseases as a group, are not available.

133 Commonalities and differences between individual diseases also remain poorly understood and continue to
134 be subject to much research. While emerging evidence has suggested that autoimmune diseases tend to co-
135 occur within individuals, large-scale investigations across a broad spectrum of autoimmune diseases that
136 could provide clues about shared pathogenesis and risk factors are not currently available.^{7,8}

137 To address these knowledge gaps, we analysed a large longitudinal database of primary and secondary care
138 records that provides information on millions of individuals' diagnoses with several years of follow-up.^{9,10} We
139 investigated the incidence and prevalence for 19 of the most common autoimmune diseases, assessed trends
140 over time, by sex, age, socioeconomic status, season and region, and examined rates of co- occurrence among
141 autoimmune diseases.

142 Methods

144 Data source

145 We used electronic health records from the Clinical Practice Research Datalink (CPRD, GOLD and AURUM
146 datasets) from 1 January 1985 to 30 June 2019. The CPRD database contains anonymised patient data from
147 approximately 20% of the current UK population and is broadly representative in terms of age, sex and
148 ethnicity. CPRD is one of the largest databases of longitudinal medical records from primary care in the world
149 and has been validated for epidemiological research for a broad range of conditions.⁹ Primary care records
150 from CPRD were linked to secondary care records from Hospital Episodes Statistics (HES admitted patient care
151 and HES outpatient) data. Linkage was available for a subset of English practices from 1 January 1998
152 onwards, covering approximately 50% of all CPRD records. Scientific approval for this study was given by the
153 CPRD Independent Scientific Advisory Committee (ISAC).

154 Case identification

155 We investigated 19 of the most common autoimmune disorders (AID): Addison's disease; ankylosing
156 spondylitis; coeliac disease; childhood-onset type 1 diabetes; Graves' disease; Hashimoto's thyroiditis;
157 inflammatory bowel disease (Crohn's disease or ulcerative colitis); multiple sclerosis; myasthenia gravis;
158 pernicious anaemia; polymyalgia rheumatica; primary biliary cholangitis; psoriasis; rheumatoid arthritis
159 (including its specific subtypes such as Still's disease, Caplan syndrome, rheumatoid spondylitis, and others);
160 Sjogren's syndrome; systemic lupus erythematosus; systemic sclerosis; vasculitis; and vitiligo.¹¹⁻¹³ For selected
161 conditions, we performed sensitivity analyses using more restrictive disease definitions and present these in
162 the **appendix**. Diseases were considered individually and as a composite outcome of all AID combined. For the
163 combined analyses, we calculated primary incidence (first recorded AID, reflecting the number of patients
164 affected by AIDs) and cumulative incidence (all recorded AIDs, reflecting the cumulative number of AID
165 diagnoses).

167 Although some of these diseases remain debated in terms of their autoimmune aetiology and may be more
168 appropriately described as 'immune mediated inflammatory diseases', to assist readability we refer to this
169 group of conditions as 'autoimmune diseases'.

170 For each condition, we identified diagnoses from primary care or secondary care records based on diagnostic
171 codes from hospital/death (International Classification of Diseases, tenth revision (ICD-10)) and primary care
172 (Read¹⁴ and SNOMED codes¹⁵) coding schemes, and, for selected conditions, prescriptions of certain drugs
173 (**appendix**). Specifically, individuals with childhood-onset type 1 diabetes were identified as those with at least
174 one diagnostic code referring to type 1 or insulin-dependent diabetes, at least 1 insulin prescription, and aged
175 19 years or less at first diagnosis. Individuals with Hashimoto thyroiditis were identified as those with at least
176 one levothyroxine prescription and no history of hyperthyroidism, pituitary disease, thyroid surgery, or
177 thyroid-altering medication (amiodarone, lithium, sodium valproate, carbimazole, propylthiouracil,
178 thalidomide, sunitinib). When conflicting arthritis diagnoses (ankylosing spondylitis and rheumatoid arthritis)
179 were recorded in the same individual, the last recorded diagnosis was used. Incident diagnoses were defined
180 as the first record of that condition in primary or secondary care records from any diagnostic position.

181 **Study population**

182 Included in the study were men and women with records labelled as 'acceptable', approved for HES and ONS
183 linkage, and registered with their general practice for at least 12 months during the study period (01/01/1998
184 to 30/06/2019). For incidence calculations, we excluded all individuals who had a diagnosis of the disease of
185 interest prior to study start date (1 January 2000), or within the first 12 months of registration with their
186 general practice.

187 **Covariates**

188 Smoking status and body mass index (BMI) were abstracted from electronic health records as the most recent
189 measurement within 2 years prior to diagnosis. Socioeconomic status was defined as the Index of Multiple
190 Deprivation (IMD) 2015 quintile,¹⁶ a composite measure of relative deprivation at a small area level, covering
191 an average population of 1500 people, ranked in ascending order of deprivation score and grouped in equal
192 fifths, with quintiles 1 and 5 representing the least and most deprived areas, respectively. Data on sex are as
193 reported by the patient when they registered with their general practitioner. Ethnicity was extracted from
194 both primary and secondary care records. When ethnicity differed between primary and secondary care
195 records, secondary care data was used.

196 **Statistical analyses**

197 Baseline characteristics are presented as frequencies (%) for categorical data, medians and interquartile range
198 (IQR) for non-normally distributed continuous data, or means and standard deviation (SD) for normally
199 distributed continuous data, over the whole autoimmune disease cohort and stratified by sex, socioeconomic
200 quintile, and period of diagnosis. Number and percentage of records with missing data are displayed for
201 variables with missing entries.

202 Observed incidence rates were computed by dividing the number of incident cases by the number of patient-
203 years in the cohort. Category-specific rates were computed separately for subgroups of age, sex,
204 socioeconomic status, region, calendar year of diagnosis and season of diagnosis. Winter was defined as the
205 period from January to March, and Summer as June to August. Time at risk was calculated to start at the latest
206 of patient's registration date plus 12 months, birthyear, or study start date; and stopped at the earliest of
207 death, transfer out of practice, practice's last collection date, incidence of the disease of interest, or study end
208 date. Observed prevalence rates were computed considering all patients ever diagnosed with autoimmune
209 disease (numerator) among patients alive and registered with a general practitioner on June 30th in each year
210 (denominator). To allow comparison with other studies, we further calculated incidence and prevalence rates
211 of childhood onset type 1 diabetes restricting the denominator to those aged 19 or less (**appendix**).

212 Standardised rates were computed by applying direct age and sex standardisation¹⁷ to the 2013 European
213 Standard Population¹⁸ using 5-year age bands up to 90 years of age.

214 Negative binomial regression models were used to examine overall and category-specific incidence rate ratios
215 (IRR) and corresponding 95% confidence intervals (CI). Models were adjusted for calendar year, age
216 (categorised into five years age-bands), sex, socioeconomic status and region. Negative binomial models were
217 chosen over Poisson models to account for overdispersion in the data. Sensitivity analyses comparing Poisson
218 and negative binomial models showed very similar results.

219 To characterise co-occurrence of autoimmune diseases, we have run a series of cohort studies to calculate
220 incidence rate ratios for the development of a second (comorbid) autoimmune disease according to disease
221 status of a first (index) autoimmune disease. Incidence rate ratios were calculated using negative binomial
222 regression models adjusted for age, sex and calendar year. We performed separate analyses for each pair and
223 sequence of autoimmune diseases, following methods by Somer et al.⁷ For these analyses, time at risk started
224 at the latest of patient registration plus 12 months or diagnosis of the index disease, and stopped at the
225 earliest of the incidence of the comorbid autoimmune disease, death, the patient ceasing registration with
226 the general practice, the practice's last collection date, or the end of the study.

227 Study findings are reported in accordance with the REporting of studies Conducted using Observational
228 Routinely-collected health Data (RECORD) recommendations.¹⁹ Statistical analyses were performed in R
229 (version 3.6.2) and validated with SAS (version 9.4).

230 **Role of the funding source**

231 The funders of the study had no role in study design, data collection, data analysis, data interpretation, or
232 writing of the report.

233 **Results**

234 A total of 22 009 375 individuals contributed data between 01/01/2000 and 31/12/2019 with 135 691 152
235 patient-years of follow-up. Among those, we identified 1 123 789 new diagnoses of autoimmune diseases,
236 affecting a total of 978 872 patients. The mean (SD) age at AID diagnosis was 54.0 (21.4) years, and 64% of
237 patients were women (**Table 1**).

238 **Temporal trends**

239 The number of patients newly diagnosed with one (or more) autoimmune diseases increased only modestly
240 over time (IRR comparing 2017-2019 vs 2000-2002: 1.04 [1.00, 1.09]). The number of new autoimmune
241 disease diagnoses increased by 22%, largely due to an increasing number of secondary autoimmune disease
242 diagnoses among patients already affected by a first autoimmune disease (IRR comparing 2017-2019 vs 2000-
243 2002: 1.22 [1.18, 1.28]). Coeliac, Graves', and Sjogren's syndrome showed the greatest increases, whereas
244 Hashimoto's thyroiditis and pernicious anaemia declined modestly over the study period (**Figure 1**). The
245 observed increase in coeliac disease, Graves' disease, and Sjogren's syndrome was largely driven by a higher
246 number of diagnoses in women. The increase in Graves' disease was largely driven by the very old (80 years
247 and older), whereas the increase in coeliac disease was largely driven by diagnoses in the very young (up to 40
248 years of age) (**appendix**).

249 **Age at diagnosis**

250 Autoimmune disorders developed over the whole life spectrum, from the first to the 95th year of life. Median
251 age at first autoimmune disease presentation varied greatly among individual autoimmune disease. For many
252 conditions, incidence increased with age; this was the case for Graves' disease, pernicious anaemia, and
253 rheumatoid arthritis. Only six conditions were commonly diagnosed before the age of five, these were:
254 Addison's, coeliac disease, type 1 diabetes, psoriasis, vitiligo and vasculitis (largely due to Henoch-Schonlein
255 purpura, Kawasaki disease, and glomerulonephritis). For others, such as multiple sclerosis, psoriasis, and
256 lupus, the incidence peaked in the middle age. Finally, some conditions presented a bi-modal age-distribution
257 with a peak in childhood or early adulthood and another one later in life; this was the case for coeliac disease,
258 inflammatory bowel disease and vasculitis (**Figure 2**).

259 **Incidence by sex**

260 Most autoimmune disorders were more common in women than men (IRR for women compared to men: 1.74
261 [1.72, 1.77] for all diseases combined). Thyroid disorders, Sjogren's syndrome, lupus, and systemic sclerosis
262 had the highest incidence rate ratio in women compared to men. Only three diseases were more common in
263 men than women, namely ankylosing spondylitis; type 1 diabetes; and myasthenia gravis (**Figure 3**).

264 **Incidence by socioeconomic status**

265 Overall, the most deprived socioeconomic groups had a higher incidence of autoimmune diseases (IRR for the
266 most deprived compared to least deprived quintile: 1.14 [1.11, 1.16] for all autoimmune diseases combined).
267 A marked socioeconomic gradient was visible across several individual diseases, including Graves' disease,
268 pernicious anaemia, rheumatoid arthritis, and lupus. For other diseases, such as Hashimoto's thyroiditis and
269 inflammatory bowel disease, no difference was observed among socioeconomic groups despite relatively
270 large number of cases; and for coeliac and polymyalgia rheumatica, disease incidence was highest in the least
271 deprived group (**Figure 4**).

272 **Regional differences**

273 Overall, variation by geographic region was limited. Notable exceptions were polymyalgia rheumatica, type 1
274 diabetes, and coeliac disease, which were considerably more common outside of the capital city, as well as
275 pernicious anaemia, which appeared to be more common in northern regions. Sjogren's syndrome, lupus and
276 vitiligo, on the other hand, presented lower incidence rates outside of London (**Figure S4**).

277 **Seasonal differences**

278 Most autoimmune diseases were diagnosed throughout the year with no significant differences between
279 winter and summer months. Seasonal variation was observed only for type 1 diabetes, which was more
280 commonly diagnosed in the winter months (January to March) compared to summer months (June to August),
281 and for vitiligo, which was more likely to be diagnosed during the summer (**Figure S5**).

282 **Co-occurrence of autoimmune diseases**

283 Autoimmune disorders were commonly associated with each other. Increased risk of developing a second
284 autoimmune disorder was seen across many autoimmune diseases, but orders of magnitude differed widely
285 between diseases. Associations were generally highest among connective tissue diseases, particularly
286 between Sjogren's, systemic lupus erythematosus and systemic sclerosis. Patients with type 1 diabetes also
287 had significantly higher rates of Addison's, coeliac and thyroid diseases. More generally, Addison's disease
288 occurred at a considerably higher incidence among patients with pre-existing autoimmune disease than in the
289 general population. Multiple sclerosis stood out as having low rates of co-occurrence with other autoimmune
290 diseases and even showed an inverse association with some autoimmune disorders, but this was only
291 significant for vitiligo (**Figure 5, Table S1**).

292 **Prevalence trends over time**

293 Together the 19 autoimmune disorders investigated in this study affected 10.2% of the population over the
294 study period (13.1% of women, 7.4% of men). Age- and sex- standardised prevalence increased over time
295 from 7.7% in 2000-2002 to 11.0% in 2017-2019 (RR comparing 2017-2019 vs 2000-2002: 1.41 [1.37, 1.44]).

296 **Discussion**

297 Our large-scale population-based study provides several novel insights into the burden of autoimmune
298 disorders, its variation over time, by individual diseases, and patient subgroups of age, sex and socioeconomic
299 status. Our findings confirm and extend evidence from previous studies demonstrating an increasing
300 incidence of several autoimmune disorders²⁰, and shows that the increase was particularly pronounced for
301 Graves' disease, coeliac disease, and rheumatic disorders. Our results reflect disease incidence based on
302 diagnostic criteria, screening practices, availability and accuracy of diagnostic tests in place at that precise
303 time and hence must be interpreted within this context. As such and in consideration of the increasing
304 awareness for some of these conditions, improved coding practices, and earlier recognition of these
305 conditions over the past two decades, the observed increase remains modest.

306 The epidemiology of type 1 diabetes is perhaps the best studied of all. Previous studies had reported varying
307 estimates and trends over time,^{4,5} and several surveys from Scandinavian countries have reported steep
308 increases since the 1950s with a plateau since mid 2000.^{21,22} In our study, we observed only a modest increase
309 in disease incidence over the past two decades. Overall estimates were comparable with estimates from a
310 recent simulation study on the incidence of type 1 diabetes²³ in Western Europe and similar to those reported
311 by the International Diabetes Federation Atlas for the UK,^{24,25} but lower than those recently reported in
312 northern European countries.²⁶ Another key finding was the reduction in Hashimoto's thyroiditis over time,
313 which could be due to more careful initiation of levothyroxine in older persons with subclinical
314 hypothyroidism following trials demonstrating limited benefits in this population.^{27,28} Pernicious anaemia also
315 showed an apparent decline in incidence over time, a decline that appears to coincide with increased use of
316 dietary supplements over the same period, and possibly a more widespread recognition of other causes of
317 vitamin B12 deficiency such as helicobacter pylori, although no causal inference can be made from our data.²⁹

318 The observed increase in incidence of rheumatic diseases will likely have important implications for health
319 services and already substantial medication expenditures linked to biologics. Improved coding practices
320 during the study period, the introduction, in 2013, of a quality audit in primary care (so-called Quality and
321 Outcomes Framework, QOF) rewarding general practitioners for maintaining a register and evaluating
322 cardiovascular/fracture risks in patients with rheumatoid arthritis³⁰, and novel classification criteria for
323 ankylosing spondylitis, are all likely to have contributed to the observed trend. The increase in incidence of
324 axial spondyloarthritis in women starting around the time of publication of ASAS (Assessment of
325 SpondyloArthritis international Society)³¹ classification criteria - introducing the concept of non-radiographic
326 axial spondyloarthritis - is consistent with other studies indicating that whilst ankylosing spondylitis is male-
327 predominant, non-radiographic axial spondyloarthritis is similarly common in women and men.^{32,33}

328 Our stratified analyses by socioeconomic status, region, and seasonal variations in disease incidence provide
329 further insights into the possible role of environmental factors in the development of autoimmune diseases.
330 Four diseases - Graves' disease, pernicious anaemia, rheumatoid arthritis, and systemic lupus erythematosus -
331 presented a clear socioeconomic gradient with those in the most deprived group up to 50 % more likely to
332 develop the disease than their affluent counterpart. Such socioeconomic disparities could indicate that diet,
333 smoking, obesity, air pollution, or other currently unrecognised environmental exposures might play a role in
334 the development of these diseases. Two conditions – coeliac disease and polymyalgia rheumatica – presented
335 an inverse socioeconomic gradient, a phenomenon rarely observed in public health research, and which could
336 be linked to increased awareness and testing for these diseases in more affluent populations or indeed
337 lifestyle differences.

338 Seasonal variations on the other hand were limited. Vitiligo was more commonly diagnosed in the summer,
339 perhaps due to increased visibility of depigmented skin areas in summer months, and type 1 diabetes was
340 more commonly diagnosed in the winter and outside of the capital region, a finding compatible with
341 hypotheses of viral triggers, diet, higher weight, or ethnicity playing a role in the disease's pathogenesis.³⁴⁻³⁶
342 Other regional variations, such as coeliac disease more common outside of the capital, remain unexplained
343 for now. While numerous reports have linked smoking to the incidence of certain autoimmune diseases such
344 as rheumatoid arthritis, multiple sclerosis, or psoriasis³⁷, we did not observe a decline in these diseases'
345 incidences despite considerable reduction in smoking prevalence over the same period³⁸. This could be due to
346 a parallel increase in other risk factors, such as obesity, over the same period, confounding via socioeconomic
347 status, or currently unknown reasons.

348 Finally, co-occurrences of diseases provide valuable insights into a possibly common aetiology across some of
349 these conditions. For example, we found high rates of co-occurrence among a range of connective tissue
350 diseases, particularly Sjogren's syndrome, systemic lupus erythematosus, and systemic sclerosis, regardless of
351 diagnostic sequence and after accounting for age and sex. Interestingly polymyalgia rheumatica did not have
352 much association with other connective tissue disorders, except for vasculitis. Myasthenia gravis and multiple
353 sclerosis also tended to co-occur. However, the association was weak and given the nature of our data, which
354 is based on routine clinical practice data, we cannot rule out the possibility of inaccurate diagnoses or coding.

355 Our study further provides robust evidence for an increased incidence of a range of autoimmune diseases,
356 particularly coeliac disease and thyroid disorders, among patients with type 1 diabetes. Such observations
357 have been reported before and may point to overlapping genetic or environmental risk factors as well as

358 people with a first autoimmune disease being more likely to undergo screening for these conditions.³⁹ While
359 we did not observe associations between childhood-onset diabetes and diseases typically occurring at older
360 age groups, our data was limited by follow-up duration, so that we cannot exclude that such associations
361 might exist. Similarly, in our study, psoriasis presented limited co-occurrence with autoimmune diseases,
362 possibly because of the large proportion of mild cases typically observed among cohorts based on primary
363 care data.⁴⁰

364 Another consideration is whether treatment of an index autoimmune disease affects risk of developing a
365 comorbid disease. Although nonspecific immunosuppressives could in theory decrease risk of a comorbid
366 autoimmune disease, targeted therapies could have differential effects. For example, an intriguing finding was
367 the high incidence of Addison's disease following almost every other autoimmune disorder, which could
368 perhaps be related to glucocorticoid induced adrenal insufficiency (which is typically recorded as Addison's
369 disease in the UK).⁴¹ Such a mechanism might also explain the higher rates of cardiovascular diseases
370 observed in patients with Addison's disease in this same cohort.⁴² Nevertheless, given the relatively small
371 number of patients with Addison's disease, these results must be interpreted with caution. Overall, this
372 research confirms that some autoimmune diseases co-occur with one another at a rate greater than expected
373 by chance or surveillance bias alone, but it reveals that this phenomenon is not generalized across all
374 autoimmune diseases.

375 A major strength of this study is the selection of a statistically powerful data source with over 130 million
376 person-years of data to investigate the incidence, prevalence, and co-occurrence of autoimmune disorders.
377 The very large size of our cohort allowed us to perform stratified analyses of unprecedented granularity, over
378 a broad spectrum of conditions, as well as allowing examination of the influence of age, sex, and
379 socioeconomic status, as well as trends over 20 years. The use of routinely reported diagnoses also captures
380 the burden of disease as experienced by physicians and health services, and likely increases the
381 generalizability of findings. One of the key limitations of our study was the limited diversity in ethnic
382 backgrounds in our cohort and the unavailability or significant missingness of additional variables potentially
383 relevant to autoimmune disease pathogenesis, such as smoking, body mass index or blood biomarkers such as
384 vitamin D deficiency. The non-randomised design of our study further meant that we were unable to account
385 for the effect of drug therapy, such as antirheumatic drugs, steroids, or biologics, on the incidence of a second
386 autoimmune disease. Research using electronic health records is also reliant on the accuracy of clinical coding
387 carried out during consultations and hospital admissions. The validity of diagnoses underlying our study has
388 been carefully assessed and was considered appropriate in light of the over two hundred independent studies
389 that have investigated the validity of diagnoses recorded in CPRD which reported an average positive
390 predictive value of about 90% for a broad range of conditions.⁴³ These include recent studies on several of the
391 diseases studied here, which have demonstrated that algorithms based on diagnostic codes perform well at
392 identifying patients with these conditions in primary care records.^{33,40} Generally, observed age distributions
393 were consistent with prior studies and add to the validity of our approach. Nevertheless, our results,
394 particularly trends over time and co-occurrence of diseases must be interpreted within the context and
395 limitations of routinely collected health records data and the possibility that some level of miscoding is
396 present. Finally, a large number of tests and subgroup analyses within one study must also be interpreted
397 with adequate caution.

398 Our findings present an important new piece in the puzzle of autoimmune disease aetiology, a group of
399 conditions that are apparent in near 10% of the population and which consume considerable health
400 resources. The socioeconomic, seasonal, and regional disparities observed among several autoimmune
401 disorders, point towards involvement of environmental factors in the pathogenesis of selected autoimmune
402 diseases. The interrelations between many autoimmune diseases further suggest a shared pathogenesis,
403 particularly among connective tissue diseases as well as between diabetes, coeliac and thyroid disorders. To
404 this day, the exact causes of many of the autoimmune diseases studied here remain unknown and require
405 further research.

406 **Contributions**

407 NC, GC and JCM conceived and designed the study. NC, GM, and GV designed the statistical analysis plan and
408 NC performed the statistical analysis. All authors contributed to interpreting the results, drafting the
409 manuscript, and the revisions. NC and GC had full access to the data in the study and had final responsibility

410 for the decision to submit for publication. NC, GM, GV, and JYV had permission to access the data and NC and
411 GM verified the data (CRPD requests that access to data is given only as absolutely necessary). All authors
412 gave final approval of the version to be published and accept responsibility to submit the manuscript for
413 publication.

414 **Declarations of interest**

415 NC is funded by a personal fellowship from the Research Foundation Flanders (grant number 12ZU922N) and
416 declares royalties from Oxford University Innovation. SM is personally funded by a Wellcome Trust career
417 development fellowship and is supported by an institutional Biomedical Research Council award. IBM declares
418 honoraria from AbbVie; grant support paid to his university from AstraZeneca and Eli Lilly; participation on
419 data safety monitoring boards or advisory boards of AstraZeneca, Bristol Myers Squibb, Eli Lilly, Novartis,
420 Janssen, GlaxoSmithKline, AbbVie, Cabaletta, Compugen, Causeway, Gilead, Moonlake, Reflexion, UCB, and
421 XinThera; patents from Novartis; leadership roles with Evelo, Versus Arthritis, and Greater Glasgow and Clyde
422 Health Board; and stock or stock options with Evelo, Compugen, and Cabaletta. JJVM has received funding to
423 his institution from Amgen and Cytokinetics for his participation in the Steering Committee for the ATOMIC-
424 HF, COSMIC-HF, and GALACTIC-HF trials and meetings and other activities related to these trials; has received
425 payments through Glasgow University from work on clinical trials, consulting and other activities from
426 Alnylam, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Cardurion, Dal-Cor,
427 GlaxoSmithKline, Ionis, KBP Biosciences, Novartis, Pfizer, and Theracos; and has received personal lecture fees
428 from the Corpus, Abbott, Hikma, Sun Pharmaceuticals, Medscape/Heart.Org, Radcliffe Cardiology, Alkem
429 Metabolics, Eris Lifesciences, Lupin, ProAdWise Communications, Servier Director, and Global Clinical Trial
430 Partners. NS declares consulting fees or speaker honoraria, or both, from Abbott Laboratories, Afimmune,
431 Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Janssen, Merck Sharp & Dohme,
432 Novartis, Novo Nordisk, Pfizer, Roche Diagnostics, and Sanofi; and grant support paid to his university from
433 AstraZeneca, Boehringer Ingelheim, Novartis, and Roche Diagnostics. KK is supported by the National Institute
434 for Health Research (NIHR) Applied Research Collaboration East Midlands and the NIHR Leicester Biomedical
435 Research Centre. KK has also acted as a consultant, speaker or received grants for investigator-initiated
436 studies for AstraZeneca, Abbott, Amgen, Napp, Bayer, Novartis, Novo Nordisk, Sanofi-Aventis, Lilly and Merck
437 Sharp & Dohme, Boehringer Ingelheim, Oramed Pharmaceuticals, and Applied Therapeutics. PNT declares
438 personal consulting fees from Immunovant and leadership roles in the Society for Endocrinology and British
439 Thyroid Association. All other authors declare no competing interests. The views expressed are those of the
440 authors and not necessarily those of the funder.

441 **Acknowledgements**

442 We would like to thank Ms. Hilary Shepherd, Dr. Sonia Coton, and Dr. Eleanor L Axson from the Clinical
443 Practice Research Datalink for their support and expertise in preparing the dataset underlying these analyses.

444 Prof Justin Mason died before publication of this work. He was widely recognised as an outstanding clinician
445 and academic, a gifted leader, and a friend to many. Justin was an international authority on large vessel
446 vasculitis and his care for his patients and contributions to the vasculitis community will continue to make a
447 significant difference for many years to come.

448 **Data sharing**

449 Access to CPRD data is subject to a license agreement and protocol approval process that is overseen by
450 CPRD's Independent Scientific Advisory Committee (ISAC). A guide to access is provided on the [CPRD website](#).

451

452 **References:**

- 453 1 Gutierrez-Arcelus M, Rich SS, Raychaudhuri S. Autoimmune diseases — connecting risk alleles with
454 molecular traits of the immune system. *Nat Rev Genet* 2016; **17**: 160–74.
- 455 2 Ehehalt S, Blumenstock G, Willasch AM, Hub R, Ranke MB, Neu A. Continuous rise in incidence of childhood
456 Type 1 diabetes in Germany. *Diabetic Medicine* 2008; **25**: 755–7.

- 457 3 Newhook LA, Penney S, Fiander J, Dowden J. Recent incidence of type 1 diabetes mellitus in children 0-14
458 years in Newfoundland and Labrador, Canada climbs to over 45/100,000: a retrospective time trend study.
459 *BMC research notes* 2012; **5**: 628.
- 460 4 Mobasser M, Shirmohammadi M, Amiri T, Vahed N, Hosseini Fard H, Ghojzadeh M. Prevalence and
461 incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health promotion*
462 *perspectives* 2020; **10**: 98–115.
- 463 5 EURODIAB ACE Study Group. Variation and trends in incidence of childhood diabetes in Europe. EURODIAB
464 ACE Study Group. *Lancet (London, England)* 2000; **355**: 873–6.
- 465 6 McLeod DSA, Cooper DS. The incidence and prevalence of thyroid autoimmunity. *Endocrine* 2012; **42**: 252–
466 65.
- 467 7 Somers EC, Thomas SL, Smeeth L, Hall AJ. Autoimmune Diseases Co-occurring Within Individuals and
468 Within Families A Systematic Review. 2006. DOI:10.1097/01.ede.0000193605.93416.df.
- 469 8 Somers EC, Thomas SL, Smeeth L, Hall AJ. Are Individuals With an Autoimmune Disease at Higher Risk of a
470 Second Autoimmune Disorder? *American Journal of Epidemiology* 2009; **169**: 749–55.
- 471 9 Herrett E, Gallagher AM, Bhaskaran K, *et al.* Data Resource Profile: Clinical Practice Research Datalink
472 (CPRD). *International journal of epidemiology* 2015; **44**: 827–36.
- 473 10 Wolf A, Dedman D, Campbell J, *et al.* Data resource profile: Clinical Practice Research Datalink (CPRD)
474 Aurum. *International Journal of Epidemiology* 2019; published online March 11. DOI:10.1093/ije/dyz034.
- 475 11 Cooper GS, Bynum MLK, Somers EC. Recent insights in the epidemiology of autoimmune diseases:
476 Improved prevalence estimates and understanding of clustering of diseases. *Journal of Autoimmunity*
477 2009; **33**: 197–207.
- 478 12 Osborne D, Sobczyńska-Malefora A. Autoimmune mechanisms in pernicious anaemia and thyroid disease.
479 *Autoimmunity Reviews* 2015; **14**: 763–8.
- 480 13 Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of
481 ankylosing spondylitis. *Rheumatology* 2014; **53**: 650–7.
- 482 14 Chisholm J. The Read Clinical Classification: The NHS Has Acquired A Coding System Designed For The
483 Computer Age on JSTOR. *British Medical Journal* 1990; **300**: 1092.
- 484 15 SNOMED CT. NHS Digital. <https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct>
485 (accessed Feb 3, 2023).
- 486 16 Department for Communities and Local Government (DCLG). The English Index of Multiple Deprivation
487 2015: Guidance. 2015.
- 488 17 Kirkwood B, Sterne J. Essential medical statistics. 2010.
- 489 18 Eurostat’s task force. Revision of the European Standard Population Report. 2013.
- 490 19 Benchimol EI, Smeeth L, Guttman A, *et al.* The REporting of studies Conducted using Observational
491 Routinely-collected health Data (RECORD) Statement. *PLOS Medicine* 2015; **12**: e1001885.
- 492 20 Lerner A, Jeremias P, Matthias T. The World Incidence and Prevalence of Autoimmune Diseases is
493 Increasing. *International Journal of Celiac Disease* 2016; **3**: 151–5.

- 494 21 Parviainen A, But A, Siljander H, *et al.* Decreased Incidence of Type 1 Diabetes in Young Finnish Children.
495 *Diabetes Care* 2020; **43**: 2953–8.
- 496 22 Berhan Y, Waernbaum I, Lind T, Möllsten A, Dahlquist G, Group* for the SCDS. Thirty Years of Prospective
497 Nationwide Incidence of Childhood Type 1 Diabetes. *Diabetes* 2011; **60**: 577–81.
- 498 23 Ward ZJ, Yeh JM, Reddy CL, *et al.* Estimating the total incidence of type 1 diabetes in children and
499 adolescents aged 0–19 years from 1990 to 2050: a global simulation-based analysis. *The Lancet Diabetes &*
500 *Endocrinology* 2022; published online Nov 10. DOI:10.1016/S2213-8587(22)00276-5.
- 501 24 Ogle GD, James S, Dabelea D, *et al.* Global estimates of incidence of type 1 diabetes in children and
502 adolescents: Results from the International Diabetes Federation Atlas, 10th edition. *Diabetes Research and*
503 *Clinical Practice* 2022; **183**: 109083.
- 504 25 Harvey JN, Hibbs R, Maguire MJ, O’Connell H, Gregory JW, Brecon Group (The Wales Paediatric Diabetes
505 Interest Group). The changing incidence of childhood-onset type 1 diabetes in Wales: Effect of gender and
506 season at diagnosis and birth. *Diabetes Res Clin Pract* 2021; **175**: 108739.
- 507 26 Waernbaum I, Lind T, Möllsten A, Dahlquist G. The incidence of childhood-onset type 1 diabetes, time
508 trends and association with the population composition in Sweden: a 40 year follow-up. *Diabetologia*
509 2022; published online Oct 20. DOI:10.1007/s00125-022-05816-0.
- 510 27 Jonklaas J, DeSale S. Levothyroxine prescriptions trends may indicate a downtrend in prescribing.
511 *Therapeutic Advances in Endocrinology and Metaboli* 2020; **11**: 1–14.
- 512 28 Stott DJ, Rodondi N, Kearney PM, *et al.* Thyroid Hormone Therapy for Older Adults with Subclinical
513 Hypothyroidism. *New England Journal of Medicine* 2017; **376**: 2534–44.
- 514 29 Mikulic M, . OTC vitamins & minerals sales value, Great Britain, 2009-2021.
515 <https://www.statista.com/statistics/521713/otc-vitamins-and-minerals-sales-value-great-britain/>
516 (accessed Dec 17, 2022).
- 517 30 Stokes T. Rheumatoid arthritis is a new clinical domain for QOF 2013/14. Guidelines in Practice.
518 [https://www.guidelinesinpractice.co.uk/musculoskeletal-and-joints-/rheumatoid-arthritis-is-a-new-clinical-](https://www.guidelinesinpractice.co.uk/musculoskeletal-and-joints-/rheumatoid-arthritis-is-a-new-clinical-domain-for-qof-2013/14/337862.article)
519 [domain-for-qof-2013/14/337862.article](https://www.guidelinesinpractice.co.uk/musculoskeletal-and-joints-/rheumatoid-arthritis-is-a-new-clinical-domain-for-qof-2013/14/337862.article) (accessed Nov 16, 2022).
- 520 31 Rudwaleit M, Heijde D van der, Landewé R, *et al.* The development of Assessment of SpondyloArthritis
521 international Society classification criteria for axial spondyloarthritis (part II): validation and final selection.
522 *Annals of the Rheumatic Diseases* 2009; **68**: 777–83.
- 523 32 Wright GC, Kaine J, Deodhar A. Understanding differences between men and women with axial
524 spondyloarthritis. *Semin Arthritis Rheum* 2020; **50**: 687–94.
- 525 33 Scott IC, Whittle R, Bailey J, *et al.* Rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis
526 epidemiology in England from 2004 to 2020: An observational study using primary care electronic health
527 record data. *The Lancet Regional Health – Europe* 2022; **23**. DOI:10.1016/j.lanepe.2022.100519.
- 528 34 Honkanen H, Oikarinen S, Nurminen N, *et al.* Detection of enteroviruses in stools precedes islet
529 autoimmunity by several months: possible evidence for slowly operating mechanisms in virus-induced
530 autoimmunity. *Diabetologia* 2017; **60**: 424–31.
- 531 35 Versini M, Jeandel P-Y, Rosenthal E, Shoenfeld Y. Obesity in autoimmune diseases: Not a passive bystander.
532 *Autoimmunity Reviews* 2014; **13**: 981–1000.

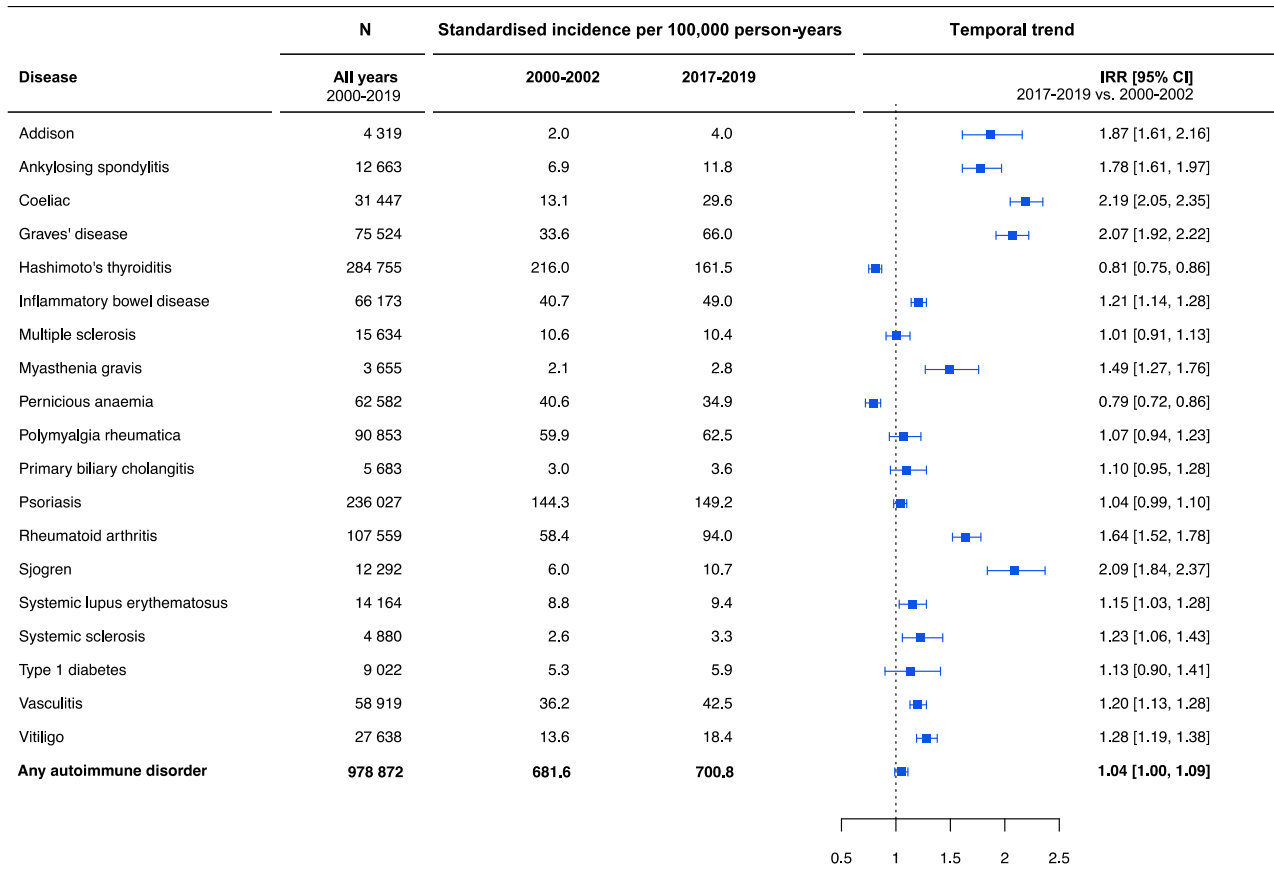
- 533 36 Ji J, Hemminki K, Sundquist J, Sundquist K. Ethnic Differences in Incidence of Type 1 Diabetes among
534 Second-Generation Immigrants and Adoptees from Abroad. *The Journal of Clinical Endocrinology &*
535 *Metabolism* 2010; **95**: 847–50.
- 536 37 Perricone C, Versini M, Ben-Ami D, *et al.* Smoke and autoimmunity: The fire behind the disease.
537 *Autoimmunity Reviews* 2016; **15**: 354–74.
- 538 38 Office for National Statistics. Adult smoking habits in England. 2020
539 [https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/datasets/adultsmokinghabitsinengland)
540 [/datasets/adultsmokinghabitsinengland](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/datasets/adultsmokinghabitsinengland) (accessed Nov 19, 2022).
- 541 39 Cohn A, Sofia MA, Kupfer SS. Type 1 Diabetes and Celiac Disease: Clinical Overlap and New Insights into
542 Disease Pathogenesis. *Curr Diab Rep* 2014; **14**: 517.
- 543 40 Seminara N m., Abuabara K, Shin D b., *et al.* Validity of The Health Improvement Network (THIN) for the
544 study of psoriasis. *British Journal of Dermatology* 2011; **164**: 602–9.
- 545 41 Prete A, Bancos I. Glucocorticoid induced adrenal insufficiency. *BMJ* 2021; **374**: n1380.
- 546 42 Conrad N, Verbeke G, Molenberghs G, *et al.* Autoimmune diseases and cardiovascular risk: a population-
547 based study on 19 autoimmune diseases and 12 cardiovascular diseases in 22 million individuals in the UK.
548 *Lancet (London, England)* 2022; **0**. DOI:10.1016/S0140-6736(22)01349-6.
- 549 43 Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General
550 Practice Research Database: a systematic review. *British journal of clinical pharmacology* 2010; **69**: 4–14.
- 551
- 552

Table1: Baseline characteristics of patients with incident autoimmune disease during 2000-2019.

	All patients (N = 978 872)	Sex		Socioeconomic status quintile		Time period	
		Women (N = 625 879)	Men (N = 352 993)	Least deprived (N = 220 047)	Most deprived (N = 170 856)	2000-2002 (N = 141 951)	2017-2019 (N = 123 420)
Age at diagnosis (years)							
Mean (SD)	54.0 (21.4)	54.4 (21.1)	53.3 (22.0)	55.5 (21.0)	51.0 (22.0)	55.7 (20.7)	52.6 (22.0)
Women, N (%)	625 879 (63.9%)			139 694 (63.5%)	110 322 (64.6%)	95 166 (67.0%)	76 917 (62.3%)
Ethnicity							
African/Caribbean	17 257 (1.8%)	11 732 (1.9%)	5 525 (1.7%)	751 (0.4%)	7 652 (4.6%)	1 471 (1.1%)	2 970 (2.5%)
Asian	46 621 (5.0%)	29 458 (4.9%)	17 163 (5.2%)	6 695 (3.2%)	12 375 (7.5%)	3 686 (2.8%)	8 362 (7.0%)
Mixed/Other	45 363 (4.9%)	27 201 (4.5%)	18 162 (5.5%)	10 202 (4.9%)	7 925 (4.8%)	3 843 (2.9%)	8 267 (6.9%)
White	824 573 (88.3%)	533 414 (88.6%)	291 159 (87.7%)	189 790 (91.5%)	137 470 (83.1%)	121 283 (93.1%)	99 984 (83.6%)
Missing	45 058 (4.6%)	24 074 (3.8%)	20 984 (5.9%)	10 375 (5.1%)	5 434 (3.2%)	11 668 (8.2%)	3 837 (3.1%)
Socioeconomic status quintile							
1 (least deprived)	220 047 (22.5%)	139 694 (22.3%)	80 353 (22.8%)			32 134 (22.6%)	28 215 (22.9%)
2	204 731 (20.9%)	129 945 (20.8%)	74 786 (21.2%)			30 146 (21.2%)	25 591 (20.7%)
3	197 109 (20.1%)	125 845 (20.1%)	71 264 (20.2%)			28 870 (20.3%)	24 498 (19.8%)
4	186 129 (19.0%)	120 073 (19.2%)	66 056 (18.7%)			26 957 (19.0%)	23 421 (19.0%)
5 (most deprived)	170 856 (17.5%)	110 322 (17.6%)	60 534 (17.1%)			23 844 (16.8%)	21 695 (17.6%)
Number of autoimmune disorders							
1	833 157 (85.1%)	522 095 (83.4%)	311 062 (88.1%)	187 435 (85.2%)	145 738 (85.3%)	112 615 (79.3%)	112 712 (91.3%)
2	124 742 (12.7%)	87 678 (14.0%)	37 064 (10.5%)	28 024 (12.7%)	21 449 (12.6%)	24 147 (17.0%)	9 698 (7.9%)
3 or more	20 973 (2.1%)	16 106 (2.6%)	4 867 (1.4%)	4 588 (2.1%)	3 669 (2.1%)	5 189 (3.7%)	1 010 (0.8%)
Body mass index (kg/m2)							
Mean (SD)	27.8 (6.33)	27.8 (6.69)	27.7 (5.55)	27.1 (5.75)	28.5 (6.99)	27.3 (5.90)	28.1 (6.71)
Missing (%)	509 139 (52.0%)	316 940 (50.6%)	192 199 (54.4%)	120 581 (54.8%)	82 763 (48.4%)	102 926 (72.5%)	56 660 (45.9%)
Smoking status							
Current smoker	126 240 (20.7%)	77 046 (19.7%)	49 194 (22.6%)	18 463 (14.1%)	33 855 (30.3%)	10 029 (24.8%)	14 054 (18.0%)
Former smoker	167 829 (27.6%)	92 501 (23.6%)	75 328 (34.7%)	37 139 (28.4%)	28 250 (25.2%)	9 151 (22.6%)	23 176 (29.7%)
Never smoker	314 512 (51.7%)	221 793 (56.7%)	92 719 (42.7%)	75 190 (57.5%)	49 784 (44.5%)	21 274 (52.6%)	40 754 (52.3%)
Missing (%)	370 291 (37.8%)	234 539 (37.5%)	135 752 (38.5%)	89 255 (40.6%)	58 967 (34.5%)	101 497 (71.5%)	45 436 (36.8%)

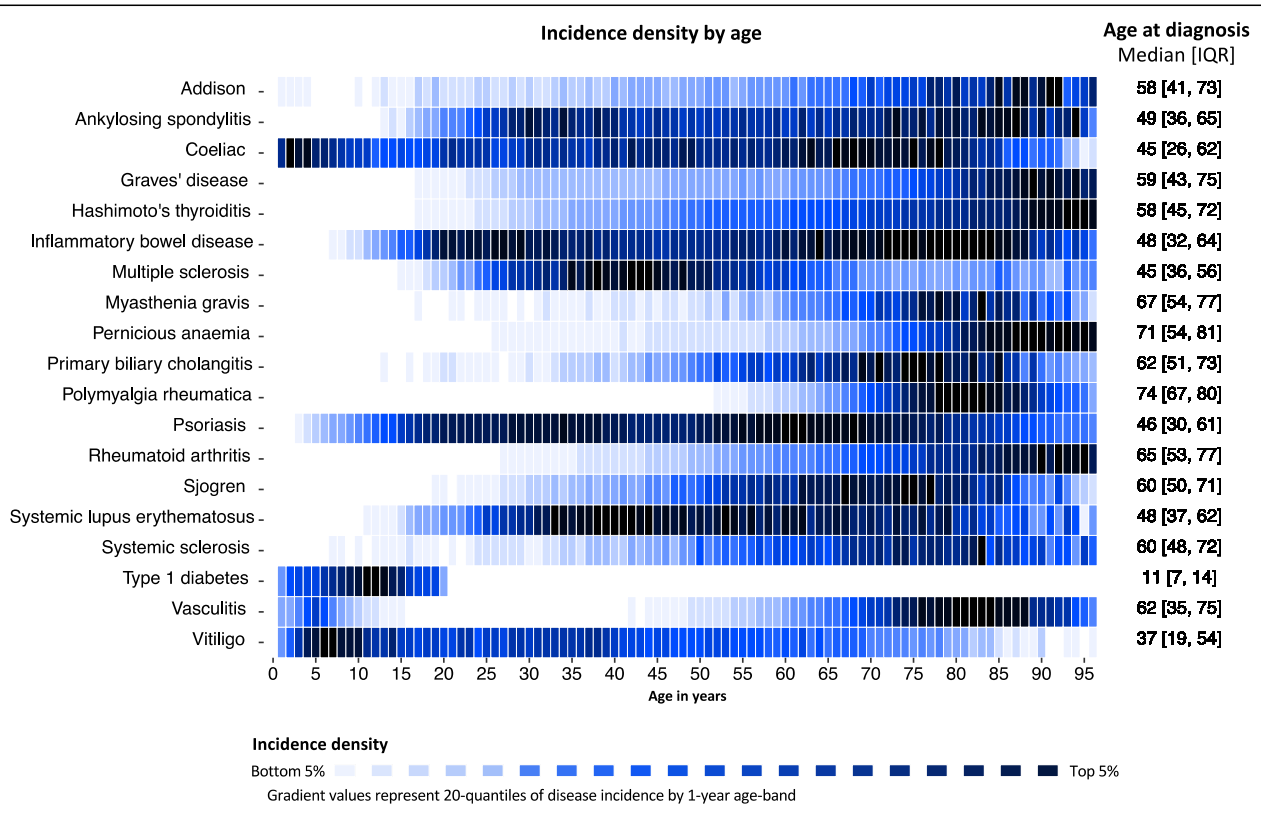
Socioeconomic status was defined as the Index of Multiple Deprivation (IMD) 2015 quintile, with 1 referring to the most affluent and 5 to the most deprived socioeconomic quintile. For variables with missing entries, category percentages refer to complete cases, and are presented alongside the percentage of missing values.

Figure 1: Incidence of autoimmune disorders over time from 2000-2019.



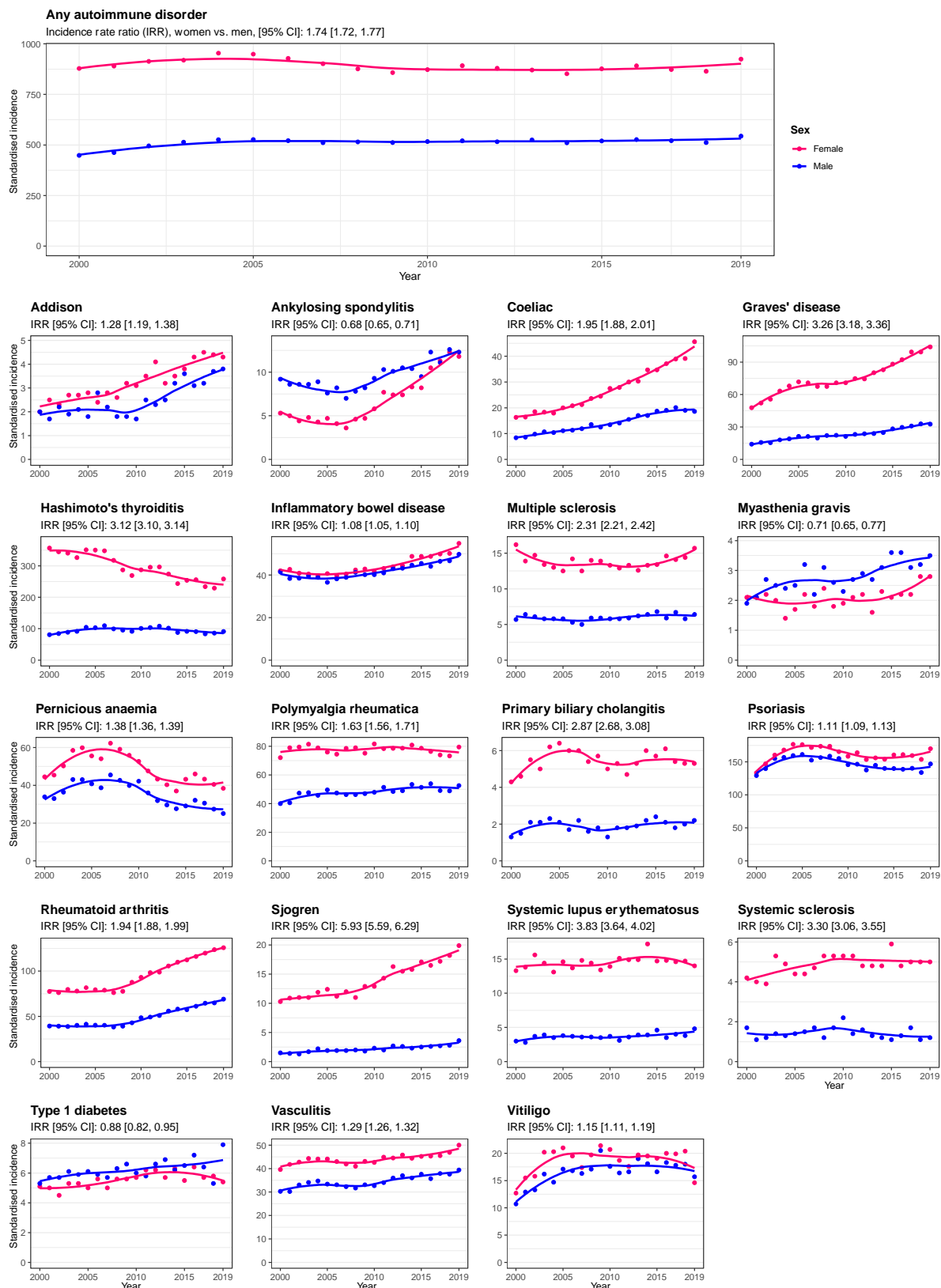
Incidence rates are presented as incidence rates per 100 000 person-years at risk and are age-sex-standardised to the 2013 European Standard Population. 'Any autoimmune disorder' refers to the primary incidence of the 19 autoimmune disorders investigated in this study (that is the number of patients first diagnosed with one autoimmune disease). 'N' refers to the number of patients newly diagnosed with autoimmune disease during the study period.

Figure 2: Incidence of autoimmune disorders by age



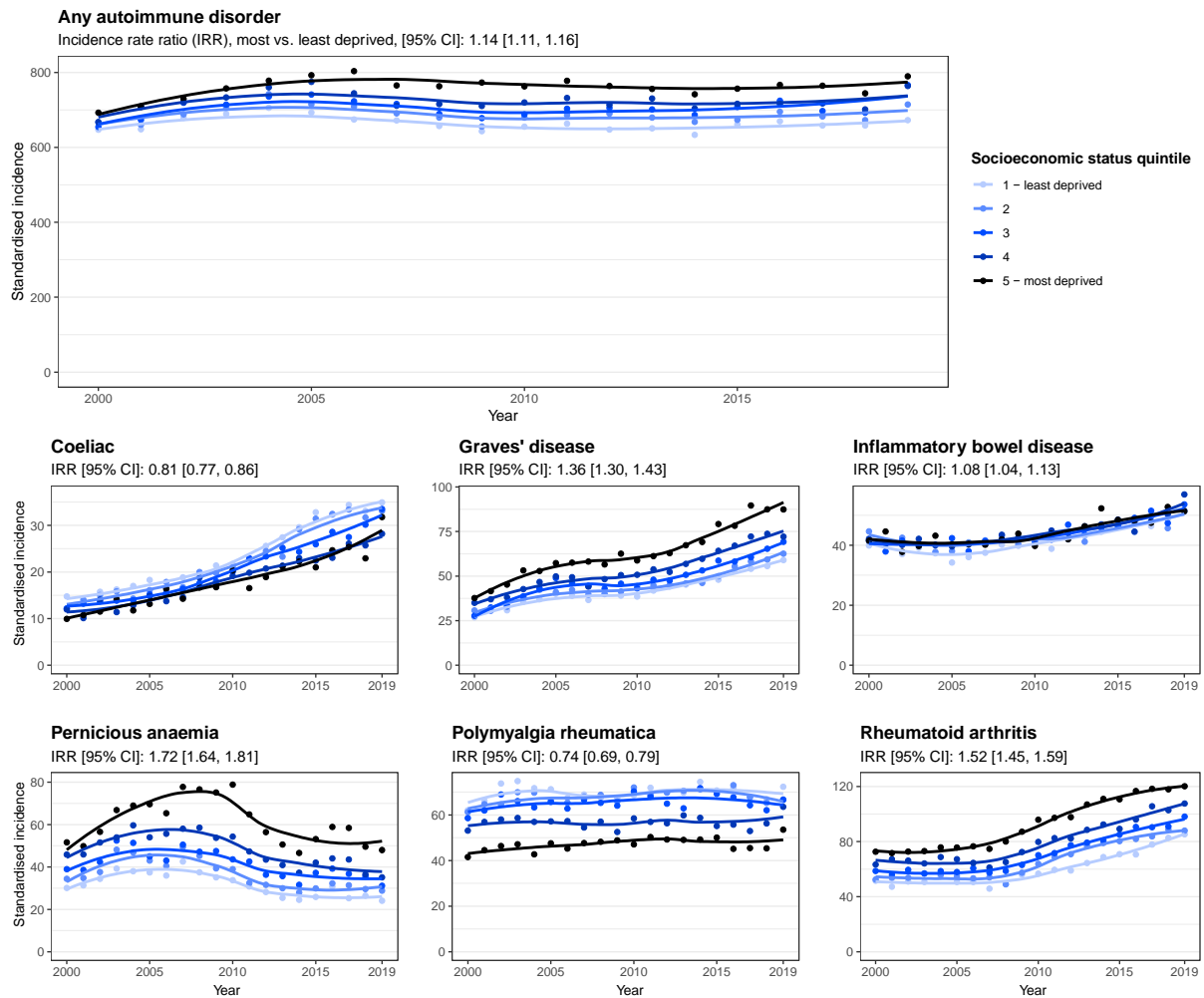
Incidence rates were calculated per 1-year age band and divided into a colour-gradient of 20-quantiles to reflect incidence density by age. The definition of type 1 diabetes used in this study refers to childhood-onset type 1 diabetes and hence to individuals aged 19 years or less at the time of diagnosis. IQR = interquartile range.

Figure 3: Incidence of autoimmune disorders by sex and over time.



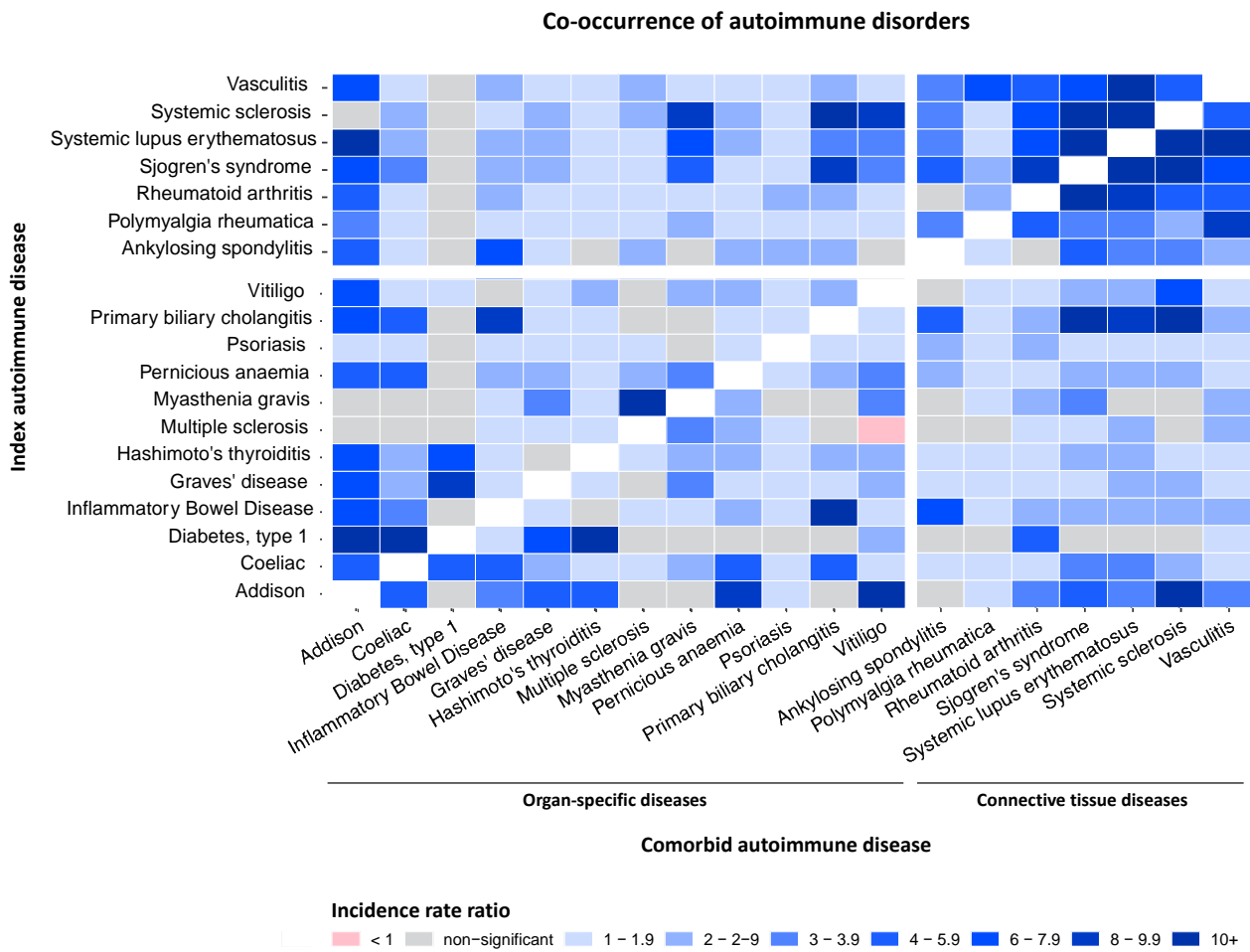
Age-standardised incidence rates by sex. 'Any autoimmune disorder' refers to the primary incidence of the 19 autoimmune disorders investigated in this study. The definition of type 1 diabetes used in this study refers to childhood-onset type 1 diabetes. Denominator populations include all age groups. Yearly incidence estimates were smoothed using loess (locally estimated scatterplot smoothing) regression lines. IRR refers to the incidence rate ratio of women compared to men and CI to the confidence interval.

Figure 4: Incidence of autoimmune disorders by socioeconomic quintile for selected disorders.



Age-sex-standardised incidence rates by socioeconomic status quintile (Index of Multiple Deprivation 2015) for selected autoimmune disorders. 'Any autoimmune disorder' refers to the primary incidence of the 19 autoimmune disorders investigated in this study. Yearly incidence estimates were smoothed using loess (locally estimated scatterplot smoothing) regression lines. IRR refers to the incidence rate ratio of the most deprived socioeconomic quintile compared to the least deprived quintile and CI to the confidence interval. Individual plots for each of the diseases investigated are presented in the **appendix**.

Figure 5: Incidence rate ratios for development of comorbid autoimmune disease among index populations with autoimmune disease compared with the general population



Incidence rate ratios were calculated for development of a second (comorbid) autoimmune disease among index populations with pre-existing autoimmune disease compared with the general population using negative binomial regression models adjusted for age and sex. Type 1 diabetes refers to childhood-onset type 1 diabetes, that is among people aged 20 years or less at the time of diagnosis.