The epidemiology, management, and the associated burden of mental health illness, atopic and autoimmune conditions, and common infections in alopecia areata: Statistical Analysis Plan

Dr Matthew Harries, Dr Abby E. Macbeth, Dr Susan Holmes, Professor Andrew R Thompson, Dr Wing Sin Chiu, Dr William Romero Gallardo, Professor Andrew G Messenger, Dr Christos Tziotzios, Professor Simon de Lusignan

Harries M, *et al. BMJ Open* 2021; 11:e045718. doi: 10.1136/bmjopen-2020-045718

Contents

List of abbreviations

Abbreviation	Full Form
AA	Alopecia Areata
BMI	Body Mass Index
CKD	Chronic Kidney Disease stage 3-5
COPD	Chronic Obstructive Pulmonary Disorder
HR	Hazard Ratio
IMD	Index of Multiple Deprivation
NIHR	National Institute for Health Research
OR	Odds Ratio
RA	Rheumatoid Arthritis
RCGP RSC	Royal College of General Practitioners Research and Surveillance Centre
SD	Standard Deviation
SES	Socio-Economic Status
SLE	Systemic Lupus Erythematosus
SSRI	Selective Serotonin Reuptake Inhibitor
T1DM	Type 1 Diabetes
T2DM	Type 2 Diabetes

1. Introduction

1.1 Background

Alopecia areata (AA) is a common cause of immune-mediated, non-scarring hair loss. (1) Clinical presentations of AA are heterogeneous, ranging from well-demarcated patches of hair loss on the scalp to total loss of hair on the scalp *(alopecia areata totalis*) or entire scalp and body (*alopecia areata universalis*).(2) Active AA is estimated to affect approximately 0.1-0.2% (2, 3)of the population worldwide, with a lifetime risk of 1.0 -2.1%. (4-6)

Following the initial episode, the majority of those with limited patchy AA will experience spontaneous remission, but over 80% will have subsequent relapses. (7) Spontaneous remission has been reported to occur in less than 10% of those with extensive AA. (7) A number of treatment options can induce hair growth in people with AA, but none have been shown to have a durable impact on remission. (8)

AA has been associated with several autoimmune conditions supporting the theory of an underlying autoimmune aetiology. (9, 10) Atopic conditions and mental health disorders have also been demonstrated to be more common in people with AA, (9-11) although the temporal relationship between psychiatric conditions and AA remains poorly researched. (12) The association between AA and common infections is not well studied.

1.2. Purposes of analyses

This statistical analysis plan covers four related AA studies.

Study 1: Epidemiology of AA

Population studies in the USA using the Rochester Epidemiology Project database estimate the lifetime risk of AA to be 2.1%. (4) The database contained information from all patients registered with a healthcare provider in Olmstead County, Minnesota from 1990-2009. Previous studies using this database from 1971-1974 estimated the overall prevalence of AA to be 0.1-0.2%, (13) with no published studies reporting prevalence estimates since this. Other epidemiological studies have assessed the incidence of AA in patients referred to dermatology services. Clinic-based studies from around the world have estimated the incidence of AA to be in the region of 0.57-3.8%. (14-18)

Worldwide, the majority of studies on AA have not found a significant difference in incidence between males and females. (5) AA onset occurs before the age of 40 in 82-86% of cases. (5) To date, there have been no large population-based studies of AA epidemiology in UK populations.

Study 2: Mental health comorbidities in patients with AA

It is well recognised by clinical experts in AA that the condition can have a profound psychological impact and a possible comorbid burden of mental health conditions. A recent USA cross-sectional database study comparing 5,605 patients with AA with matched controls found a higher prevalence of mental health diagnoses in the AA group (32.8% vs 20.0%). (19) In particular, higher rates of depression (OR 2.18, 95%CI 1.80-2.63), anxiety disorders (OR 2.46, 95%CI 1.91-3.16), attention-deficit disorders (OR 8.11, 95%CI 5.22-12.59), adjustment disorders (OR 4.80, 95%CI 2.58-8.91) and obsessive compulsive

disorders (OR 4.95, 95%CI 1.59-15.35) were observed when compared with controls. (19) These findings concur with previous work and a similar pattern is seen in other countries. (20-23)

Recently a British study identified a bi-directional positive association between major depressive disorder and AA. (24) In a sample of over 400,000 participants, major depressive disorder was seen to increase the risk of incident AA by 90% [HR 1.90, 95%CI 1.67-2.15] and AA was found to increase the risk of incident major depressive disorder by 34% [HR 1.34, 95%CI 1.23-1.46]. (24) The association of other common mental health conditions with AA however was not assessed.

Study 3: Autoimmune and atopic conditions in people with AA

A central theory in AA pathogenesis is that it is an autoimmune phenomenon resulting from a disruption of hair follicle 'immune privilege', where numerous immunological mechanisms fail to prevent cytotoxic immune attack on cells and antigens present at that site. (25)

A number of studies have reported autoimmune conditions to occur more frequently in those with AA than in the general population, however results are conflicting, and studies have been limited by potential recruiting, recall and reporting biases. (26) A Taiwanese retrospective cohort study found individuals with AA to have an overall increased risk of developing an autoimmune condition when compared with matched controls (HR 1.86, 95%CI 1.32-2.63). In particular, greater risks of incident systemic lupus erythematosus (SLE) (HR 5.02; 95%CI 2.08-12.51), psoriasis (HR 2.01; 95%CI 1.06-3.83) and rheumatoid arthritis (RA) (HR 1.79; 95%CI 1.07-3.00) were observed. (27) There was no relationship between AA and Hashimoto's thyroiditis or type 1 diabetes (T1DM). A large USA cross-sectional study of 3,658 individuals with AA, found that 14.6% had a diagnosis of thyroid disease, 11.1% had T1DM, 4.3% had inflammatory bowel disease, 2.8% had vitligo and 6.3% had psoriasis. (10) These figures were all higher than would be expected in the general population. A recent systematic review and meta-analysis found AA to be significantly associated with Graves' Disease (OR 2.07, 95%CI 1.80-2.38) and Hashimoto's Thyroiditis (OR 2.15; 95%CI 1.71-2.72) but not with SLE, RA, Psoriasis or T1DM. (9)

There has also been interest in the prevalence of co-morbid atopy in patients with AA. The prevalence of atopy in patients with AA has been reported to be in the range of 11-38.2% (5) representing cohorts from the USA, India and Singapore. (10, 14, 28, 29) A large meta-analysis found patients with AA had higher odds of having atopic dermatitis (OR 2.36; 95%CI 1.80-3.09) and allergic rhinitis (OR 1.33;95%CI 1.19-1.47) when compared with controls. (9) It is not however surprising from a pathogenesis perspective in terms of shared pathways of autoimmune susceptibility; IL-13 loci have been associated with autoimmune and atopic conditions, and a genome-wide association study established that IL-13 is also susceptible loci for AA. (30)

Study 4: Common infections in people with AA

It has been hypothesized that infections may act as triggers for AA; with particular interest in Helicobacter pylori, cytomegalovirus, Epstein-Barr virus, hepatitis B virus, hepatitis C virus, and human immunodeficiency virus but others have also been suggested.(9, 31-38) Some case reports has suggested improvement or resolution of AA with treatment of the associated infection.(33, 39). However, early small-scale studies have failed to find convincing evidence to support associations.(35, 36, 40)

Multiple genetic loci have been identified to be both protective from infection but increase the risk for autoimmune disease.(41) Examples include the HLA variants HLA-DQ2 and HLA-DQ8 which confer protection against viral infections including influenza but are associated with inflammatory bowel disease,(42, 43) and SH2B3 rs3184504*A gene variant protective against Escherichia coli infection but also detrimental for inflammatory bowel disease.(44) These loci which, in the past have conferred a genetic advantage, may still confer a reduced infection risk.(41) It is currently unclear if the genetic predisposition for AA is caused by genetic factors which also confer protection against any common infections. In other autoimmune conditions, infection risk is often increased as a result of the pathological changes resulting from the condition itself, the effects of immune-modulatory treatment, or through increased exposure to pathogens through healthcare contact.(45-51) An understanding of the relative infection rates in people with AA compared to those without is of particular importance with the emergence of Janus kinase inhibitors as a potential treatment option.(52) Whilst this medication class can produce impressive hair regrowth results even in long-standing, AA their use is associated with some increased infection risk; in particular, urinary tract infections, viral gastroenteritis fungal infections, and varicella zoster infection.(52)

2. Study objectives and endpoints

2.1 Study objectives

2.1.1 Study 1

The overall purpose of study 1 is to describe the epidemiology (incidence and prevalence) of AA and to assess the current level of primary care service utilisation and management patterns associated with people diagnosed with AA.

2.1.2 Study 2

The overall purpose of study 2 is to assess the prevalence and incidence of mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) in adult people diagnosed with AA relative to a control population of people without a diagnosis of AA. In addition, the study will determine the treatment burden (medications and psychological interventions used to treat mental health conditions), 'sick day' burden, and unemployment burden of AA.

2.1.3 Study 3

The overall purpose of study 3 is to assess the prevalence and incidence of atopic and autoimmune conditions in adult people diagnosed with AA relative to a control population of people without AA.

2.1.4 Study 4

The overall purpose of study 4 is to assess the incidence of common infections in people diagnosed with AA relative to a control population of people without AA.

2.2 Primary objectives

2.2.1 Study 1

2.2.1.1

Describe the epidemiology of AA: Incidence of AA over time and by age, sex, ethnicity, and sociodemographic factors, and point prevalence by age and sex.

2.2.1.2

Describe current primary care service utilisation and management patterns in AA.

2.2.2 Study 2

2.2.2.1

Describe the prevalence of common mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) in adult people diagnosed with AA in a contemporary real-world population compared with matched controls without AA.

2.2.2.2

Describe the incidence of common mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) in adult people diagnosed with AA in a contemporary real-world population compared with matched controls in the first two years after AA diagnosis.

2.2.2.3

Describe the mental health treatment burden of adult people diagnosed with AA.

2.2.3 Study 3

2.2.3.1

Describe the prevalence of atopic and autoimmune conditions in adult people diagnosed with AA in a contemporary real-world population at diagnosis compared with matched controls without AA.

2.2.3.2

Describe the incidence of atopic and autoimmune conditions in adult people with AA in a contemporary real-world population compared with matched controls in the first five years after AA diagnosis.

2.2.4 Study 4

2.2.4.1

Describe the overall rate of common infections in people with AA in a contemporary real-world population at diagnosis compared with matched controls in the first five years after AA diagnosis.

2.2.4.2

Describe the rates of individual common infections in people with AA in a contemporary real-world population at diagnosis compared with matched controls in the first five years after AA diagnosis.

2.3 Secondary objectives:

2.3.1 Study 1

2.3.1.1

Describe the geographic distribution of AA incidence within England. English regions comprise; Greater London, South East, South West, West Midlands, North West, North East, Yorkshire and the Humber, East Midlands, and East of England.

2.3.2 Study 2

2.3.2.1

Describe the prevalence of additional mental health conditions (adjustment disorder, agoraphobia, self-harm, and overdose/parasuicide attempts) in adult people diagnosed with AA and controls.

2.3.2.2

Describe the burden of 'sick days' in adult people diagnosed with AA and controls.

2.3.2.3

Describe the prevalence of unemployment in adult people diagnosed with AA and controls.

2.3.3 Study 3

None

2.3.4 Study 4

None

2.4 Primary endpoint

The primary endpoints for the four studies are listed in Table 1.

Study	Endpoints
1	Diagnosis of AA is the date of the first diagnosis code in the record and no alternative diagnosis that merits exclusion* diagnosed in the subsequent 365 days from date of AA diagnosis (see Supplementary Table S1.1, Additional file 2).
2	Diagnosis of AA is the date of the first diagnosis code in the record and no alternative diagnosis that merits exclusion* diagnosed in the subsequent 365 days from date of AA diagnosis (see Supplementary Table S1.1, Additional file 2). Diagnosis of a common mental health condition categorized as depressive episodes, recurrent depressive disorder or non-phobia related anxiety disorders at the time of AA diagnosis (See Supplementary Tables S2.1 - S2.4, Additional file 2). Time to diagnosis of a common mental health condition categorized as depressive episodes, recurrent depressive disorder or non-phobia related anxiety disorders in the first two years post AA diagnosis (See Supplementary Tables S2.1 -S2.4, Additional file 2)
3	Diagnosis of AA is the date of the first diagnosis code in the record and no alternative diagnosis that merits exclusion* diagnosed in the subsequent 365 days from date of AA diagnosis (see Supplementary Table S1.1, Additional file 2). Diagnosis of an atopic (atopic dermatitis, allergic rhinitis, asthma) or autoimmune condition (Supplementary Tables S4.1 – S4.16, Additional file 2) at the time of AA diagnosis. Time to diagnosis of an atopic (atopic dermatitis, allergic rhinitis, asthma) or autoimmune condition (Supplementary Tables S4.1 – S4.16, Additional file 2) in the first five years after AA diagnosis.
4	Diagnosis of AA is the date of the first diagnosis code in the record and no alternative diagnosis that merits exclusion* diagnosed in the subsequent 365 days from date of AA diagnosis (see Supplementary Table S1.1, Additional file 2). Diagnosis of a common infection (Upper and lower respiratory tract infection, Pneumonia, Acute bronchitis, Influenza, Skin infection, Urinary tract infection, Genital infections, Gastrointestinal infection, Herpes simplex , Herpes zoster, Meningitis and Covid-19 (depending on data availability)) (Supplementary Tables S5.1 – S5.12, Additional file 2) in the first five years after AA diagnosis.

Table 1: Primary endpoints. *Exclusion diagnoses are; any form of scarring alopecia, (53) traction alopecia, congenital alopecia, androgenetic alopecia, telogen effluvium, tinea capitis, trichotillomania, or secondary syphilis of the scalp (See Supplementary Table S1.1, Additional file 2)

2.5 Secondary endpoints

None

2.6 Other exploratory analyses

Study 1: Explore the incidence of AA by age (children) and age group (adults (18-30], (30-40], (50-60], (60,70], (70-80], ≥80). These age groups are provisional. If, during exploratory analysis, we find that these age groups obscure the true incidence of AA (for example if AA incidence rates change substantially across one of these age groups) then we may modify these groupings. Any changes will be clearly stated as such in a final manuscript.

3 General Study Design and Plan

The proposed studies will use the retrospectively collected data of all eligible patients within Oxford-Royal College of General Practitioners Research and Surveillance Centre database (RCGP RSC) at the date of data extraction. The RCGP RSC database incorporates pseudonymised primary care records from up to 200 GP practices distributed across England. Details of the cohort included in the database have been published previously. (54) The RCGP RSC cohort size (as of January 2019) is over 2.6 million actively registered patients, and historic records available for up to 5.2 million patients over the last 20 years. The cohort provides a broadly representative sample of the UK population.

The study period for study 2 and study 3 incidence analyses will be ten-years starting January 1, 2009 and ending December 31, 2018 inclusive.

For study 4, we will start data collection from the beginning of 2004 when the quality and outcomes framework was first introduced in primary care. The study period will end December 31, 2018.

3.1 Study Design

These are non-interventional studies. A range of study designs will be used across the studies as follows:

3.1.1 Study 1:

Cross-sectional design (incidence and prevalence of AA). Cohort design to assess healthcare utilisation in people with AA.

3.1.2 Study 2:

Case-control design for prevalence of mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) at AA diagnosis. Matched-cohort design for incidence of such conditions in the first two years after AA diagnosis in the subset of adult people diagnosed with AA without a mental health condition at baseline (and excluding controls if they have a baseline mental health condition).

3.1.3 Study 3:

Case-control design for prevalence of atopic and autoimmune conditions at AA diagnosis. Matched-cohort design for incidence of atopic and autoimmune conditions in the first five years after AA diagnosis in the subset of adult people diagnosed with AA without an atopic/autoimmune condition at baseline (and excluding controls if they have a baseline atopic/autoimmune condition).

3.1.4 Study 4:

Matched-cohort design for incidence of common infections in the first five years after AA diagnosis in the subset of people diagnosed with AA without a common infection at baseline (and excluding controls if they have a baseline common infection).

3.2 Data source and Read codes

The case and control cohorts will be identified from the RCGP RSC database. The RCGP RSC database contains complete data on all events and clinical entities coded in UK primary care. These include demographic information, clinical diagnoses, laboratory test results, primary care issued prescriptions, process of care codes (e.g. specialist dermatology reviews), and anthropometric measurements (e.g., body mass index (BMI)), and are coded using the Read Coding system. (55) Read code lists used to define variables have been developed in accordance with published guidelines. (55)

3.3 Data extraction

Individual patient data will be anonymised at the point of data extraction. All data will remain in anonymised form and will be held on a secure server at the University of Surrey. The data will not be used for any purposes other than for the research which is described in the respective protocols and which has been approved by the RCGP RSC Research Committee. The sponsor, Pfizer Ltd, will not have access to the individual anonymised patient data.
Queries will be executed to extract relevant data from the Structured Query Language (SQL) database using the Read code listed in Additional file 2.

3.4 Inclusion-Exclusion Criteria and General Study Population

3.4.1 AA definition

We will identify people diagnosed with AA using diagnosis Read codes which are specific to the condition (see Supplementary Table S1.1 (AA section), Additional file 2). A person will be considered to have a confirmed AA diagnosis if they have one or more of these diagnoses Read codes and no code for an alternative diagnosis in the subsequent 365 days. Alternative diagnoses which merit exclusion comprise any form of scarring alopecia, (53) traction alopecia, congenital alopecia, androgenetic alopecia, telogen effluvium, tinea capitis, trichotillomania, or secondary syphilis of the scalp (see Supplementary Table S1.1, Additional file 2). We will identify all incident cases of AA meeting this case definition over the study period which will form the case cohorts for the four studies. We will also extract data on the extent of the alopecia where this has been recorded (See Supplementary Table S1.2, Additional file 2).

Study 2 will identify three groups of common mental health conditions defined by the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD10) classification (56) (depressive episodes (F32), recurrent depressive disorder (F33), and non-phobia related anxiety disorders (F41)) identified using algorithms validated in UK primary care. (57)

Study 3 will identify atopic conditions and autoimmune conditions respectively (Supplementary Tables S4.1 – S4.16, Additional file 2) using diagnostic Read codes.

Autoimmune conditions were selected based on a combination of expert knowledge and a minimum background population prevalence of 0.3% in the general population (refer to section 4.1.3). Case definition using Read codes for atopic dermatitis has been previously validated. An ontologically driven case definition is used for both allergic rhinitis and asthma. (55)

Study 4 will identify common infections using diagnostic Read codes (Supplementary Tables S5.1 – S5.12, Additional file 2).

3.5 Baseline socio-economic characteristics

Age, sex, socio-economic status (SES), and ethnicity will comprise the socio-economic factors used for stratification of the outcome measures. Ethnicity will be categorized as follows in accordance with the major UK census categories: (58, 59) White, Asian, Black African/Caribbean, Mixed, Other and not recorded. Deprivation will be defined using the official national deprivation measure; index of multiple deprivation (IMD). (60) This will be calculated at the point of data extraction, using patient postcode, with the resultant scores stratified by deprivation quintile according to the national distribution.

3.6 Epidemiology of AA

This will comprise of the incidence of AA over the study period (stratified by age group, sex ethnicity, and SES) and point prevalence (stratified by age and sex). The study will also describe the geographic distribution of AA by region across England.

3.7 Current service utilisation and management patterns

Current service utilisation will comprise of annual primary care visit rates, in Study 1 we will compare primary care visit frequency in people with AA and matched controls without AA.

The management patterns will comprise of an assessment of the percentage of the people with AA having a secondary care review (by dermatology (study 1) or mental health services (study 2)), and an assessment of primary care medication prescriptions issued to treat AA (Supplementary Table S1.4, Additional file 2). Primary care visits, secondary care reviews and medication to treat AA will all be limited to those within 365 days of AA diagnosis.

3.8 Mental health treatment, sick day, and unemployment burden

This will comprise of the rates for medications and psychological interventions used to treat mental health conditions in people diagnosed with AA. Antidepressant medication classes to be examined comprise; selective serotonin reuptake inhibitors (SSRIs) and related medications (serotonin and norepinephrine reuptake inhibitors; SNRIs), tricyclic antidepressants (TCAs) and related medications (tetracyclic antidepressant TeCAs), and monoamine oxidase inhibitors (MOAIs). Anxiolytic medications to be examined comprise all benzodiazepines and other related medications indicated for use in anxiety states. Psychological interventions comprise of will comprise; referral rates for counselling, secondary care psychiatrist referrals, including the referrals to the clinical psychology services/specialists and UK Improving Access to Psychological Therapies (IAPT) program. (61) (Supplementary Table S2.10, Additional file 2).

Additional mental health conditions consist of adjustment disorder, agoraphobia, self-harm, and overdose/parasuicide attempts and will be identified using the presence of one or more diagnosis codes appearing in the clinical record (Supplementary Tables S2.5 – S2.9, Additional file 2).

Unemployment will be identified using Read codes relating to unemployment recorded in the primary care record or the issuing of IB113 or ESA113 forms (Supplementary Table S3.1, Additional file 2). Sick days will be indicated by the issuing of Med 3 certification from primary care (Statement of Fitness for Work certification) (Supplementary Table S3.2, Additional file 2).

4 Sample size calculations

In the controls, the prevalence of common conditions (e.g. depressive episodes) is assumed to be 10%, less common conditions are assumed to be 1% and rare conditions (e.g. multiple sclerosis) assumed to be 0.1%. Sample size calculations were performed in OpenEpi, (62) results are presented using methods of Kelsey.(63)

4.1.1 Study 1

The analysis of the primary endpoint is a descriptive analysis of the incidence and prevalence of AA, and no sample size calculation is therefore applicable.

Our inclusion criteria specify all people with a confirmed diagnosis of AA will be included for analysis. The RCGP RSC database contains information on around 2 million actively registered patients. Based on an AA prevalence of 0.2% (13), we would anticipate a sample size of around 4,000 people diagnosed with AA for this study.

4.1.2 Study 2

Assuming 80% power, a 5% level of statistical significance, and a background population prevalence of 15% for anxiety and 10% for depression, (64) our anticipated sample size for AA (n = 4,000) would be sufficient to detect a prevalence difference of 1.8% in anxiety between those with and without AA, and 1.5% difference in depression between those with and without AA.

4.1.3 Study 3

A sample size calculation is applicable to estimate the incidence of atopic and autoimmune conditions in people diagnosed with AA. There is, however, a paucity of large-scale epidemiological data exploring the association between AA and either atopic or autoimmune conditions on which to base the sample size calculations for this study.

A large population-based study based in Taiwan (65) found associations between the presence of AA and the presence of atopic dermatitis (OR 2.24, 95% CI 1.95-2.58) and allergic rhinitis (OR 1.29, 95% CI 1.18-1.41), and with the autoimmune condition psoriasis (OR 2.80, 95% CI 2.24-3.50). No associations were found in either asthma (atopic) or RA (autoimmune). Applying an estimate informed by this study of OR 2.0 for atopic conditions and OR 2.8 for autoimmune conditions, assuming 80% power, a 5% level of statistical significance, a 4:1 ratio of controls to cases and a probability of outcome in the control cohort of 1% (less common condition prevalence) for both atopic and autoimmune conditions, a total (both cases and controls) sample size of 6,034 for atopic conditions (1,207 cases and 4,827 controls) and 2,133 for autoimmune conditions (427 cases and

1,706 controls) would be required to detect statistically significant differences between people diagnosed with AA and the control populations.

For the assessment of prevalence of co-existing autoimmune conditions, assuming 80% power, a 5% level of statistical significance, our anticipated sample size for AA (n = 4,000) and an OR 2.0 effect, we would have sufficient power to detect a difference between those with and without AA in autoimmune conditions having >= 0.3% background population prevalence.

4.1.4 Study 4

A sample size calculation is applicable to estimate the incidence of common infections in people diagnosed with AA. There is, however, no previous data exploring the association between AA and common infections on which to base the sample size calculations for this study. We therefore estimate the number of people diagnosed with AA we will have for this study (n = 4,000) and use this number to estimate the minimum effect size we are powered to detect. Based on a Cox PH model, 80% power, a 5% level of statistical significance and a 4:1 ratio of controls to cases, we would have sufficient power to detect a difference between those with and without AA in common infections having >= 0.46% background population prevalence.

5 Populations to be analysed

5.1 Inclusion Criteria

The cohorts for studies 1, 2 and 3 (cases and controls) will consist of patients contributing to RCGP RCS database during the study period (between January 1, 2009 and December 31, 2018 inclusive). The cohorts for study 4 (cases and controls) will consist of patients contributing to RCGP RCS database between January 1, 2004 and December 31, 2018 inclusive.

The AA prevalence population consists of patients actively registered and alive at the end date of the study period (December 31, 2018).

The AA incidence population consists of patients registered at any point during the study period but without a prior recorded AA diagnosis. The follow-up period for incident cases will begin on the latest of the date of diagnosis indicated by the first diagnostic AA Read code. Follow-up will end at the earliest of the study end-date (December 31, 2018), the date of patient transfer from an included practice or date of death.

Studies 1 and 4 will include both children and adults. Studies 2 and 3 will include only people aged ≥18 (as at date of AA diagnosis for cases or start of follow-up for controls). Only people aged ≥ 18 and ≤ 65 will be included in the unemployment and sick day analysis (Study 2).

5.2 Exclusion Criteria

The following exclusion criteria will be applied:

- People with the alternative non-AA diagnoses (section 3.4.1) (AA incidence cohort).
- People with AA diagnosis within 6 months of registration (AA incidence cohort).
- People with less than 1 year of follow up within the dataset (unless under 1 year old) (AA incidence and AA prevalence cohorts).

5.3 Matching process

The control cohorts will be defined by matching cases with people who have never been diagnosed with AA either prior to or during the study period, by age, sex, and time since practice registration, at GP practice level. After matching an assessment of sociodemographic differences using IMD quintiles will be performed. If there are substantial differences in socioeconomic status between cases and controls, re-matching with IMD as an additional matching criterion will be considered

Controls will require at least one year of follow-up when matched to minimize the risk of a non-recorded existing diagnosis of AA. The follow-up period for each matched control will begin at the start of follow-up of their matched case and have the same follow-up time criteria applied (where applicable).

Each case is matched with four controls using a nearest neighbour matching algorithm (66) to create a baseline matched dataset.

Study 1 will include all matched patients to compare rates of primary care visits. Study 2 (incidence of mental health after AA diagnosis) and study 3 (incidence of atopic and autoimmune conditions) will both use a subset of the baseline matched dataset; cases will be people diagnosed with AA without the relevant condition at their start of follow up, controls will be people not diagnosed with AA and without the relevant condition at their start of follow up.

The matching for the incidence of common infections (study 4) will use a different approach to studies 2 and 3: cases will be people diagnosed with AA on or before 1st January 2014 (exact date will be dependent on data availability) without the relevant infection, controls will be people not diagnosed with AA prior to this date and without the relevant infection. People who develop AA during the follow up will be excluded from the pool of available cohorts prior to matching. This approach will mitigate systematic recording bias that is likely due to an increased propensity of people with incident AA to consult within the months following diagnosis.

The case-control design studies (prevalence of mental health at AA diagnosis (study 2), and prevalence of atopic and autoimmune conditions at AA diagnosis (study 3)) will both define cases as people newly diagnosed with AA over the study period. For each case, the prevalence of the conditions of interest as at the date of AA diagnosis will be retrospectively established. Controls will be people not diagnosed with AA. For each control, the date of AA diagnosis of their matched case counterpart will then be used as the baseline date and the prevalence of the conditions of interest on that date will be retrospectively established.

In Study 2 and Study 3, the same baseline cases and control cohorts will be used for both the case-control and matched-cohort analyses.

5.4 AA definition exploratory analyses

To check the validity of our AA definition (section 3.4.1), annual incident trends of specified alopecia (AA), (Supplementary Table S1.1, Additional file 2), non-specific alopecia (Supplementary Table S1.1, Additional file 2) and other specified alopecia (Supplementary Table S1.1, Additional file 2) will be analysed for the purpose of identifying changes in coding practice. The AA case definition may be amended based on the results of this analysis if considered appropriate and necessary. Any changes required will be reported in full in a final study publication.

6 Statistical Analysis

6.1 Statistical principles

The mean, standard deviation (sd) and any other statistical measures, will be reported to one decimal place. Continuous data will be summarised in the form of means, sd, median, interquartile range (IQR) and range as appropriate. Categorical data will be summarised using frequencies and proportions. Chi-squared and t-tests will be performed to compare the frequency of dichotomous variables and the values of continuous variables where relevant. 95% CIs will be reported for main effect sizes. Statistical significance will be assessed using $p<0.05$. Actual p-values will be reported except for p-values less than 3 decimal places which will be reported as "<0.001". P-values >0.1 will be reported to one decimal place, p-values between 0.01 and 0.1 will be reported to 2 decimal places, and p-values between 0.001 and 0.01 will be reported to three decimal places.

The assessment of any associations with baseline characteristics and the outcomes of interest (mental health conditions (study 2), autoimmune and atopic conditions (study 3) and common infections (study 4)) will be assessed using logistic regression, Cox proportional hazards models, Poisson regression or zero-inflated Poisson regression as appropriate depending on the characteristics of the outcome variable. A standard set of features will be used in all multivariable models: age, sex, ethnicity, socioeconomic status, BMI, smoking status, alcohol status, and historical propensity to consult (defined as the number of primary care visits an individual has in the year prior to their study start date). Specific co-morbidity adjustment will be specific to each analysis.

Assumptions of statistical approaches will be assessed, and transformations used if necessary (although it is not anticipated that the data will violate model assumptions).

6.2 Primary Endpoint Analysis

6.2.1 Study 1

AA incidence: Incident cases will be defined as people with a first ever diagnostic Read code for AA during the study period. To ensure cases are truly incident, people with an AA diagnosis within six months of registering with a practice will be excluded from the analysis (unless under one year old). Incidence of new onset AA will be calculated by dividing the total number of new cases by the total person-years of follow-up for the total eligible population over the study period. Incidence rates will be reported for the complete cohort and stratified by calendar year, sex and age (categorised), and SES factors. Adjusted incidence rates by socio-demographic factors (including geographic area classification) will be estimated using Poisson regression.

AA point prevalence: Prevalence cases will be defined as all people with a historical diagnostic Read code for AA and no exclusion code within one year of AA diagnosis. For this prevalence analysis all people will be included regardless of the date of AA diagnosis in relation to the date of registration with a practice. Cases with < 365 days of follow up from date of registration will be excluded from the analysis (unless under one year old). Point prevalence will be calculated by dividing the number of prevalent cases by the total number of eligible people in the study population at the study end date (December 31, 2018). Point prevalence will be reported for the complete cohort and stratified by sex and age (categorised).

We anticipate some attrition of recording of AA in the historic records. This is particularly likely to be the case in records preceding the computerisation of primary care which occurred in the 1990s. If we observe evidence of this being a major source of error in our point calculations (as evidenced by a falling point incidence in the older adult population despite there being a substantial number of incident diagnoses in this age group) then we will consider using the actuarial life table method to estimate lifetime AA risk (67).

Service utilisation and management patterns: Service utilisation will be evaluated by calculation of the annual rates of primary care visits for people diagnosed with AA in the year following diagnosis. Primary care visits will be identified from the patient's electronic health record.

Management patterns will be evaluated by calculation of annual proportion of people referred for dermatology review in people diagnosed with AA in the year following diagnosis. Secondary care dermatology reviews will be identified from Read codes.

In addition, management patterns will also comprise of the percentage of people diagnosed with AA and prescribed AA medication, in primary care, grouped by medication class during the year following diagnosis. Prescription information will be identified using Read and EMIS codes.

6.2.2 Study 2

Prevalence of mental health conditions: Prevalence will be calculated by dividing the number of prevalent people at the study start date by the total number of eligible people in the study population at the same time point.

The prevalence of each common mental health condition (depressive episodes, recurrent depressive disorder and anxiety disorder) category for people diagnosed with AA will be reported in Study 2 and compared to the prevalence in the control population without AA. The prevalence of mental health conditions by socio-demographic subgroups will also be reported.

Incidence of mental health conditions: Incident cases will be defined as in Study 1. Incidence rates of mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) within two years of diagnosis date (matched follow-up start date for controls) will be calculated in both cases and controls and compared.

Risk of developing a mental health condition (grouped by mental health condition category) will be examined using time to failure analysis. Initially, unadjusted Cox proportional hazards models, stratified by matched set (AA versus non-AA), will be used to provide overall hazard ratios (HRs) as summary estimates for the association of the presence of AA with the time to each mental health condition category in separate models. Models will be subsequently adjusted for standard set factors (section 6.1) and comorbidities (type 2 diabetes (T2DM), hypertension, atrial fibrillation, angina, acute myocardial infarction, stroke, heart failure, chronic liver disease, dementia, RA, asthma, Chronic obstructive pulmonary disease, Chronic Kidney Disease stage 3-5 (CKD), malignancy and inflammatory bowel disease) in multivariable analysis.

Mental health treatment and management patterns: The burden of treatment for mental health conditions will be evaluated as the total percentage of people diagnosed with

AA and prescribed, within one year of AA diagnosis, a medication used to treat mental health conditions (section 3.8) (e.g. SSRIs), grouped by mental health condition category.

Management patterns will be evaluated by calculation of annual proportion of people referred for psychiatrist referral and counselling in people diagnosed with AA in the year following diagnosis (Supplementary Table S2.10, Additional file 2).

6.2.3 Study 3

Prevalence of atopic and autoimmune conditions: The prevalence of atopic and autoimmune conditions in both cases and controls will be reported and compared.

Incidence of atopic and autoimmune conditions: Incidence rates of atopic and autoimmune conditions within five years of diagnosis date (matched follow-up start date for controls) will be calculated in both cases and controls and compared.

Risk of developing an atopic or autoimmune condition will be examined using time to failure analysis. Initially, unadjusted Cox proportional hazards models, stratified by matched set (AA versus non-AA), will be used to provide overall HRs as summary estimates for the association of the presence of AA with the time to each of atopic and autoimmune conditions in separate models. Models will be subsequently adjusted for standard set factors (section 6.1) and comorbidities (hypertension, hyperlipidaemia, T2DM, peripheral arterial disease, atrial fibrillation, angina, myocardial infarction, stroke, heart failure, CKD, chronic obstructive pulmonary disorder (COPD), chronic liver disease, malignancy, dementia, fracture history, depression) in multivariable analysis.

6.2.4 Study 4

Incidence of common infections: Incidence rates of common infections within five years of diagnosis date (matched follow-up start date for controls) will be calculated at both overall and individual infection levels in both cases and controls and compared. : Initially, infection incidence rates in people with and without AA will be estimated by dividing the number of incident common infections by the sum of person-years of follow-up for each cohort over the period of interest. Unadjusted Poisson regression models, stratified by matched set (AA versus controls), will be used to provide overall incident rate ratios as summary estimates for the association between rates of common infections and the presence of AA. Models will subsequently be adjusted for standard set factors (section 6.1) and comorbidities (hypertension, hyperlipidaemia, T2DM, peripheral arterial disease, atrial fibrillation, angina, myocardial infarction, stroke, heart failure, CKD, COPD, chronic liver disease, malignancy, dementia, RA, depression).

6.3 Secondary Objective Analysis

6.2.1 Study 1

Geographical distribution of AA: The geographic distribution of people with AA will be calculated from the disease incidence. The adjusted incident rate ratios will be reported. Models will be adjusted for age category, sex, ethnicity, IMD and geographic area classification derived using patient postcode at the point of data extraction.

6.2.2 Study 2

Stratified prevalence of mental health conditions: The prevalence of mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) by sociodemographic subgroups will also be reported in both cases and controls.

Additional mental health conditions: The prevalence of additional mental health conditions (adjustment disorder, agoraphobia, self-harm, and overdose/parasuicide attempts) will be estimated as the total percentage of people diagnosed with AA and with the presence of one or more diagnosis codes appearing in the clinical record

Sick day and unemployment burden: The burden of sick days will be estimated as the total percentage of people diagnosed with AA and issued one or more Med 3certificate. The burden of unemployment will be estimated as the total percentage of people diagnosed with AA and with a Read code relating to unemployment recorded in the primary care record or an IB113 or ESA113 form issued.

6.2.3 Study 3

Not applicable.

6.2.4 Study 4

Not applicable.

6.4 Sensitivity analysis

Study 2: Sensitivity analyses will be performed to compare the baseline prevalence and incidence over two years post alopecia diagnosis of the three primary mental health conditions (depressive episodes, recurrent depressive disorder and anxiety disorder) in people with non-specific alopecia codes only (i.e. no prior or subsequent codes specifying alopecia type) using the same matching process used for the primary analyses.

Study 3: We will replicate this sensitivity analysis to compare the baseline prevalence and incidence of atopic and autoimmune conditions in people with non-specific alopecia codes only.

Study 4: In the common infections theme, we will use a survival model to investigate time to first infection. We will also compare the results of the study outcomes when 1) we increase the number of cases to include people who develop AA during the follow up and 2) use a similar matching approach to the other studies: perform matching using the time of AA diagnosis and follow for up to five years after diagnosis (or matched date for controls). Both of these approaches will be used to investigate the effect of bias.

6.5 Missing data

These studies will use the missing indicator variable method as missing data are considered likely not to be missing at random, meaning multiple imputation approaches will lack validity.

Patient will need to be excluded from analyses based on a matched design if they have incomplete data in the fields (age, sex) required to run the matching process. In similar studies, the exclusion rate due to incomplete data has been low (68, 69).

7 Study Limitations

These are discussed in the associated protocol manuscript.

8 Statistical software

All statistical analyses will be performed using R software.

9 Ethics

Study approval will be requested from the Research Committee of the RCGP RSC. These studies do not meet the requirements for formal ethics board review in their current form as defined using the NHS Health Research Authority research decision tool (http://www.hra-decisiontools.org.uk/research/).

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigor and follow generally accepted research practices described in Guidelines for Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE), and Good Epidemiological Practice (GEP) guidelines issued by the International Epidemiological Association (IEA). Study reporting will be conducted in accordance with the relevant EQUATOR (Enhancing the QUAlity and Transparency Of health Research) guidelines.

10 Acknowledgements

Medical writing support was provided by Louise Jordan at Momentum Data and was funded by Pfizer. Editorial support was provided by Anita Lynam, Andrew McGovern, and John Dennis at Momentum Data and was funded by Pfizer. The authors would like to acknowledge project management support from Filipa Ferreira of the University of Oxford and University of Surrey. Matthew Harries was supported by the NIHR Manchester.

References

1. Hordinsky MK. Overview of alopecia areata. J Investig Dermatol Symp Proc. 2013;16(1):S13-5.
2. Alkhalifah A, Alsantali A, Wang E, McElwee KJ, Shapiro J. Alopecia areata update: part I. Clinical picture, histopathology, and pathogenesis. Journal of the American Academy of Dermatology. 2010;62(2):177-88, quiz 89-90.
3. Madani S, Shapiro J. Alopecia areata update. Journal of the American Academy of Dermatology. 2000;42(4):549-66; quiz 67-70.
4. Mirzoyev SA, Schrum AG, Davis MDP, Torgerson RR. Lifetime incidence risk of alopecia areata estimated at 2.1% by Rochester Epidemiology Project, 1990-2009. J Invest Dermatol. 2014;134(4):1141-2.
5. Villasante Fricke AC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. Clin Cosmet Investig Dermatol. 2015;8:397-403.
6. Richard MA, Corgibet F, Beylot-Barry M, Barbaud A, Bodemer C, Chaussade V, et al. Sex- and age-adjusted prevalence estimates of five chronic inflammatory skin diseases in France: results of the << OBJECTIFS PEAU >> study. J Eur Acad Dermatol Venereol. 2018;32(11):1967-71.
7. Harries MJ, Sun J, Paus R, King LE. Management of alopecia areata. BMJ. 2010;341:c3671.
8. Cochrane. Treatments for alopecia aerate, alopecia totalis and alopecia universalis. 2008.
9. Lee S, Lee H, Lee CH, Lee WS. Comorbidities in alopecia areata: A systematic review and meta-analysis. Journal of the American Academy of Dermatology. 2019;80(2):466-77.e16.
10. Huang KP, Mullangi S, Guo Y, Qureshi AA. Autoimmune, atopic, and mental health comorbid conditions associated with alopecia areata in the United States. JAMA Dermatol. 2013;149(7):789-94.
11. Montgomery K, White C, Thompson A. A mixed methods survey of social anxiety, anxiety, depression and wig use in alopecia. BMJ open. 2017;7(4):e015468.
12. Hunt N, McHale S. The psychological impact of alopecia. BMJ. 2005;331(7522):951-3.
13. Safavi K. Prevalence of alopecia areata in the First National Health and Nutrition Examination Survey. Arch Dermatol. 1992;128(5):702.
14. Sharma VK, Dawn G, Kumar B. Profile of alopecia areata in Northern India. Int J Dermatol. 1996;35(1):22-7.
15. Tan E, Tay YK, Goh CL, Chin Giam Y. The pattern and profile of alopecia areata in Singapore--a study of 219 Asians. Int J Dermatol. 2002;41(11):748-53.
16. Yang S, Yang J, Liu JB, Wang HY, Yang Q, Gao M, et al. The genetic epidemiology of alopecia areata in China. Br J Dermatol. 2004;151(1):16-23.
17. Guzmán-Sánchez DA, Villanueva-Quintero GD, Alfaro Alfaro N, McMichael A. A clinical study of alopecia areata in Mexico. Int J Dermatol. 2007;46(12):1308-10.
18. Furue M, Yamazaki S, Jimbow K, Tsuchida T, Amagai M, Tanaka T, et al. Prevalence of dermatological disorders in Japan: a nationwide, cross-sectional, seasonal, multicenter, hospital-based study. J Dermatol. 2011;38(4):310-20.
19. Singam V, Patel KR, Lee HH, Rastogi S, Silverberg JI. Association of alopecia areata with hospitalization for mental health disorders in US adults. Journal of the American Academy of Dermatology. 2019;80(3):792-4.
20. Sellami R, Masmoudi J, Ouali U, Mnif L, Amouri M, Turki H, et al. The relationship between alopecia areata and alexithymia, anxiety and depression: a case-control study. Indian J Dermatol. 2014;59(4):421.
21. Ruiz-Doblado S, Carrizosa A, García-Hernández MJ. Alopecia areata: psychiatric comorbidity and adjustment to illness. Int J Dermatol. 2003;42(6):434-7.
22. Baghestani S, Zare S, Seddigh SH. Severity of Depression and Anxiety in Patients with Alopecia Areata in Bandar Abbas, Iran. Dermatol Reports. 2015;7(3):6063.

23. Chu SY, Chen YJ, Tseng WC, Lin MW, Chen TJ, Hwang CY, et al. Psychiatric comorbidities in patients with alopecia areata in Taiwan: a case-control study. Br J Dermatol. 2012;166(3):525-31.
24. Vallerand IA, Lewinson RT, Parsons LM, Hardin J, Haber RM, Lowerison MW, et al. Assessment of a Bidirectional Association Between Major Depressive Disorder and Alopecia Areata. JAMA Dermatol. 2019.
25. Paus R, Ito N, Takigawa M, Ito T. The hair follicle and immune privilege. J Investig Dermatol Symp Proc. 2003;8(2):188-94.
26. Barahmani N, Schabath MB, Duvic M. History of atopy or autoimmunity increases risk of alopecia areata. Journal of the American Academy of Dermatology. 2009;61(4):581-91.
27. Chen C-H, Wang K-H, Lin H-C, Chung S-D. Follow-up study on the relationship between alopecia areata and risk of autoimmune diseases. The Journal of Dermatology. 2016;43(2):228-9.
28. Tosti A. Practice gaps. Alopecia areata and comorbid conditions. JAMA Dermatol. 2013;149(7):794.
29. Tan E, Tay YK, Giam YC. A clinical study of childhood alopecia areata in Singapore. Pediatr Dermatol. 2002;19(4):298-301.
30. Jagielska D, Redler S, Brockschmidt FF, Herold C, Pasternack SM, Garcia Bartels N, et al. Follow-up study of the first genome-wide association scan in alopecia areata: IL13 and KIAA0350 as susceptibility loci supported with genome-wide significance. J Invest Dermatol. 2012;132(9):2192-7.
31. Bhardwaj EK, Trüeb RM. Acute diffuse and total alopecia of the female scalp associated with borrelia-infection. International journal of trichology. 2015;7(1):26-8.
32. Cho M, Cohen PR, Duvic M. Vitiligo and alopecia areata in patients with human immunodeficiency virus infection. Southern medical journal. 1995;88(4):489-91.
33. Gil Montoya JA, Cutando Soriano A, Jimenez Prat J. Alopecia areata of dental origin. Medicina oral : organo oficial de la Sociedad Espanola de Medicina Oral y de la Academia Iberoamericana de Patologia y Medicina Bucal. 2002;7(4):303-8.
34. Ito T, Tokura Y. Alopecia areata triggered or exacerbated by swine flu virus infection. The Journal of dermatology. 2012;39(10):863-4.
35. Jackow C, Puffer N, Hordinsky M, Nelson J, Tarrand J, Duvic M. Alopecia areata and cytomegalovirus infection in twins: genes versus environment? Journal of the American Academy of Dermatology. 1998;38(3):418-25.
36. Offidani A, Amerio P, Bernardini ML, Feliciani C, Bossi G. Role of cytomegalovirus replication in alopecia areata pathogenesis. Journal of cutaneous medicine and surgery. 2000;4(2):63-5.
37. Paoletti V, Mammarella A, Basili S, Paradiso M, Di Franco M, De Matteis A, et al. Prevalence and clinical features of skin diseases in chronic HCV infection. A prospective study in 96 patients. Panminerva medica. 2002;44(4):349-52.
38. Richardson CT, Hayden MS, Gilmore ES, Poligone B. Evaluation of the Relationship between Alopecia Areata and Viral Antigen Exposure. American journal of clinical dermatology. 2018;19(1):119-26.
39. Kartal ED, Alpat SN, Ozgunes I, Usluer G. Reversible alopecia universalis secondary to PEG-interferon alpha-2b and ribavirin combination therapy in a patient with chronic hepatitis C virus infection. European journal of gastroenterology & hepatology. 2007;19(9):817-20.
40. Abdel Hafez HZ, Mahran AM, Hofny EM, Attallah DAA, Sayed DS, Rashed H. Alopecia areata is not associated with Helicobacter pylori. Indian journal of dermatology. 2009;54(1):17-9.
41. Domínguez-Andrés J, Netea MG. Impact of Historic Migrations and Evolutionary Processes on Human Immunity. Trends in Immunology. 2019;40(12):1105-19.
42. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet. 2010;42(4):295-302.

43. Hung SC, Hou T, Jiang W, Wang N, Qiao SW, Chow IT, et al. Epitope Selection for HLA-DQ2 Presentation: Implications for Celiac Disease and Viral Defense. J Immunol. 2019;202(9):2558-69.
44. Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, Dubois PC, et al. Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. Am J Hum Genet. 2010;86(6):970-7.
45. Kucharzik T, Maaser C. Infections and Chronic Inflammatory Bowel Disease. Viszeralmedizin. 2014;30(5):326-32.
46. Toruner M, Loftus EV, Jr., Harmsen WS, Zinsmeister AR, Orenstein R, Sandborn WJ, et al. Risk factors for opportunistic infections in patients with inflammatory bowel disease. Gastroenterology. 2008;134(4):929-36.
47. Brassard P, Bitton A, Suissa A, Sinyavskaya L, Patenaude V, Suissa S. Oral Corticosteroids and the Risk of Serious Infections in Patients With Elderly-Onset Inflammatory Bowel Diseases. The American journal of gastroenterology. 2014;109:1795.
48. Lichtenstein GR, Feagan BG, Cohen RD, Salzberg BA, Diamond RH, Chen DM, et al. Serious infections and mortality in association with therapies for Crohn's disease: TREAT registry. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2006;4(5):621-30.
49. Lichtenstein GR, Feagan BG, Cohen RD, Salzberg BA, Diamond RH, Price S, et al. Serious infection and mortality in patients with Crohn's disease: more than 5 years of follow-up in the TREAT registry. The American journal of gastroenterology. 2012;107(9):1409-22.
50. Fardet L, Petersen I, Nazareth I. Common Infections in Patients Prescribed Systemic Glucocorticoids in Primary Care: A Population-Based Cohort Study. PLOS Medicine. 2016;13(5):e1002024.
51. Nikiphorou E, de Lusignan S, Mallen C, Khavandi K, Roberts J, Buckley CD, et al. Haematological abnormalities in new-onset rheumatoid arthritis and risk of common infections: a population-based study. Rheumatology. 2019.
52. Gilhar A, Keren A, Paus R. JAK inhibitors and alopecia areata. The Lancet. 2019;393(10169):318-9.
53. Olsen EA, Bergfeld WF, Cotsarelis G, Price VH, Shapiro J, Sinclair R, et al. Summary of North American Hair Research Society (NAHRS)-sponsored Workshop on Cicatricial Alopecia, Duke University Medical Center, February 10 and 11, 2001. Journal of the American Academy of Dermatology. 2003;48(1):103-10.
54. Correa A, Hinton W, McGovern A, van Vlymen J, Yonova I, Jones S, et al. Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile. BMJ open. 2016;6(4):e011092.
55. de Lusignan S, Liaw ST, Michalakidis G, Jones S. Defining datasets and creating data dictionaries for quality improvement and research in chronic disease using routinely collected data: an ontology-driven approach. Informatics in primary care. 2011;19(3):127-34.
56. World Health Organisation. International Statistical Classification of Diseases and Related Health Problems 10th Revision2016 2nd January 2020. Available from: http://apps.who.int/classifications/icd10/browse/2016/en.
57. John A, McGregor J, Fone D, Dunstan F, Cornish R, Lyons RA, et al. Case-finding for common mental disorders of anxiety and depression in primary care: an external validation of routinely collected data. BMC Medical Informatics and Decision Making. 2016;16(1):35.
58. Office for National Statistics. Ethnicity2018 2nd January 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity.
59. Tippu Z, Correa A, Liyanage H, Burleigh D, McGovern A, Van Vlymen J, et al. Ethnicity Recording in Primary Care Computerised Medical Record Systems: An Ontological Approach. J Innov Health Inform. 2017;23(4):920.
60. Ministry of Housing CLG. English indices of deprivation 2015: technical report2015 2nd Jan 2020. Available from: https://www.gov.uk/government/publications/english-indices-of-deprivation-2015-technical-report.

61. Clark DM. Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. Annu Rev Clin Psychol. 2018;14:159-83.
62. Dean AG, Sullivan KM, Soe MM. OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version [updated 06/04/2013. Available from: www.OpenEpi.com.
63. Kelsey JL, Whittemore AS, Evans AS, Thompson WD. Methods in Observational Epidemiology 2nd Edition, Table 12-15. New York: Oxford University Press; 1996.
64. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archives of general psychiatry. 2005;62(6):617-27.
65. Chu SY, Chen YJ, Tseng WC, Lin MW, Chen TJ, Hwang CY, et al. Comorbidity profiles among patients with alopecia areata: the importance of onset age, a nationwide population-based study. Journal of the American Academy of Dermatology. 2011;65(5):949-56.
66. Ho D, Imai K, King G, Stuart E. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software, Articles. 2011;42(8):1 - 28.
67. Schouten LJ, Straatman H, Kiemeney LA, Verbeek AL. Cancer incidence: life table risk versus cumulative risk. J Epidemiol Community Health. 1994;48(6):596-600.
68. Kumar S, de Lusignan S, McGovern A, Correa A, Hriskova M, Gatenby P, et al. Ischaemic stroke, haemorrhage, and mortality in older patients with chronic kidney disease newly started on anticoagulation for atrial fibrillation: a population based study from UK primary care. Bmj. 2018;360:k342.
69. Rayner LH, Mcgovern A, Sherlock J, Gatenby P, Correa A, Creagh-Brown B, et al. The impact of therapy on the risk of asthma in type 2 diabetes. Clin Respir J. 2019;13(5):299-305.