


BMJ Open Epidemiology and management of atopic dermatitis in England: an observational cohort study protocol

Simon de Lusignan ^{1,2}, Helen Alexander,³ Conor Broderick,³ John Dennis,⁴ Andrew McGovern,⁴ Claire Feeney,⁵ Carsten Flohr³

To cite: de Lusignan S, Alexander H, Broderick C, *et al.* Epidemiology and management of atopic dermatitis in England: an observational cohort study protocol. *BMJ Open* 2020;**10**:e037518. doi:10.1136/bmjopen-2020-037518

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-037518>).

Received 06 February 2020

Revised 07 July 2020

Accepted 22 August 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

²Royal College of General Practitioners Research and Surveillance Centre, London, UK

³Unit for Population-Based Dermatology Research, St John's Institute of Dermatology, Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK

⁴Momentum Data, St Albans, UK

⁵Pfizer, Surrey, UK

Correspondence to
Professor Carsten Flohr;
carsten.flohr@kcl.ac.uk

ABSTRACT

Introduction Atopic dermatitis (AD) is one of the most common inflammatory skin conditions in both children and adults. Despite this, contemporary descriptions of the incidence, prevalence and current management of the condition in the UK are lacking.

Methods and analysis We will perform a series of retrospective studies using a large population-based cohort derived from the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) network database to explore two key research themes: AD epidemiology and AD management. In the epidemiology theme, we will describe the incidence and prevalence of AD in children and adults in England from 2009 to 2018 inclusive. We will stratify findings by age, national Index of Multiple Deprivation (IMD), ethnicity, urban-rural environment and geographic location; and explore independent associations of these features with AD in multivariable models.

In the management theme, we will explore healthcare utilisation and treatment in people with AD. Regarding healthcare utilisation, we will evaluate rates of AD-associated primary care visits and specialist dermatology referrals in people with AD. Rates will be stratified by age, gender, socioeconomic IMD quintile and ethnicity. We will describe contemporary treatment by estimating prescribing rates across medication classes used in AD (emollients, topical corticosteroids by potency, topical calcineurin inhibitors, topical antimicrobials, antihistamines, oral corticosteroids and systemic immunomodulatory therapies) overall, and by age and sociodemographic groupings. We will also examine trends in prescribing over the study period. In people first diagnosed with AD during the study period, we will describe differences in treatment escalation by sociodemographic factors using time-to-event analysis.

Ethics and dissemination The Health Research Authority decision tool classed this a study of 'usual practice', ethics approval was not required. Study approval was granted by the RCGP RSC Study Approval Committee. Results will be disseminated through peer-reviewed publications.

Trial registration number NCT03823794.

INTRODUCTION

Atopic dermatitis (AD), also known as 'atopic eczema' or 'eczema', poses the highest burden of all skin diseases in adults and

Strengths and limitations of this study

- We will use a large, nationally representative primary care database to explore the epidemiology and management of atopic dermatitis (AD) using a validated case definition.
- As the majority of AD is managed in primary care, this study will provide an accurate contemporary analysis of AD epidemiology in England in both children and adults.
- The study will also provide a large-scale analysis of AD treatment and management.
- The dataset has a high level of completeness on sociodemographic and geographic data and therefore enables an assessment of sociodemographic and geographic associations with AD epidemiology and treatment across England.
- Our study will not be able to detect subclinical AD, where people have not consulted their general practitioner for the condition and may under-represent some of the secondary care AD treatments, since we rely on medical records held in primary care.

children worldwide.¹ While many affected children will have resolution or improvement by their teenage years, a significant proportion will continue to have AD as an adult.^{2,3} Despite this, contemporary descriptions of the incidence and prevalence of AD in the UK are lacking and there are very little data on current management. In the UK, AD prevalence estimates vary considerably with questionnaire-based studies quoting prevalence rates in adults from 2.5% to 15%.^{4,5} Treatment recommendations have changed over the years, but there have been no recent studies of treatment in AD, and we therefore do not have a current picture of AD-related healthcare resource utilisation in the UK.

While severe cases may be referred to secondary care, the vast majority of AD in the UK is managed in general practice.⁶ UK primary care records therefore provide an excellent opportunity to study these unanswered questions. This study series, using a

large and contemporary population-based AD cohort, aims to provide an up-to-date picture of the epidemiology and current management of AD in England.

Current knowledge

Epidemiology

The relapsing and remitting nature of AD, coupled with methodological issues and varying case definitions, makes accurate estimates of AD prevalence challenging. Recent work using the UK Clinical Practice Research Data-link reported an approximate UK prevalence of 10%. However, this study was not specifically designed to report AD prevalence and only included adults.⁷ Another study reported AD incidence rates, but only in children.⁸

Treatment

AD treatment takes a multistep approach. Trigger avoidance and daily topical emollients form the mainstay of maintenance therapy.⁹ Topical corticosteroids (TCS) can be used to induce better disease and flare control, with steroid potency determined by site and severity.^{9,10} Topical calcineurin inhibitors (TCI) are sometimes added in for more severe disease, especially where TCS are contraindicated, not working or in body sites with a predilection for skin thinning.^{9,10} Antimicrobials and antihistamines may be required for infected eczema or pruritus, respectively.¹⁰ Moderate-severe AD is also sometimes treated with phototherapy.

AD cases recalcitrant to topical treatment and/or phototherapy often require systemic immunomodulatory treatment¹¹ with conventional immunosuppressive therapies, such as methotrexate, ciclosporin, azathioprine or mycophenolate mofetil. Oral corticosteroids are also used, in particular in the context of severe acute flares.¹² In addition, dupilumab (an anti-interleukin-4 receptor α monoclonal antibody) is now licensed in several major markets and has shown efficacy in randomised trials and real-world studies in moderate to severe AD.^{13–15} It is the first biological therapy available for the treatment of AD in adult and adolescent patients where conventional systemic medication has been unsuccessful or is contraindicated.¹⁶

While it has been noted that primary care attendance rates for eczema are high, with 96% of children with AD seeing their general practitioner (GP) within the last 12 months,⁶ there have been no recent large-scale analyses of healthcare resource utilisation in AD.

Aims and objectives

We aim to perform a series of retrospective studies using a large population-based cohort in England exploring AD in two theme areas: epidemiology and management.

Within the AD epidemiology theme, we aim to provide a comprehensive description of the incidence and prevalence of active AD throughout the life course (in both children and adults). We will explore factors associated with both AD incidence and prevalence at different life stages and describe geographic differences in incidence

and prevalence across England. In those with incident AD we will also describe the proportion of individuals who develop moderate and severe AD after diagnosis and will explore factors associated with more severe disease.

Within the AD management theme, we will describe the rates of primary care attendance and specialist referrals in those with active AD and examine prescribing by medication class. We will also analyse the association of sociodemographic factors on these components, trends in prescribing over the last decade and treatment intensification.

METHODS AND ANALYSIS

Study design

This study will comprise a series of retrospective cohort analyses over a 10-year period covering the epidemiology and management of AD. Two study cohorts will be defined to enable multiple research questions to be answered under each theme: an incident cohort and an active AD (prevalent) cohort. These cohorts are described in detail in the study protocol below.

Data source

We will use routinely collected data extracted from the Oxford-Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) network database to provide a representative population-based sample. The RCGP RSC database contains the pseudonymised primary care computerised medical records (CMRs) of all individuals registered with a large number of affiliated GP practices. Data are extracted using Structured Query Language (SQL) queries which are used to filter and retrieve specific patient information relevant to the study from the database. RCGP RSC practices have practice visits and receive feedback via a dashboard to improve data quality.¹⁷ At the time of this study, the database contains data from 3.85 million people registered with 293 practices across England providing a broadly representative sample of the English primary care population.¹⁸

RCGP RSC CMRs include data on patient demographics, diagnoses, test results, prescriptions and 'processes of care' such as clinic visits and specialist referrals. Nearly all prescribing takes place in primary care, though a few high-cost drugs may be prescribed in secondary care. These data are coded using the Read code system.¹⁹ Data from UK primary care are useful for studies like this as each patient can only be registered with one GP, CMRs have been used since the 1990s, and pay-for-performance targets, present since 2004, ensure consistent high-quality data about chronic diseases.²⁰ Studies in several chronic diseases such as diabetes, atrial fibrillation, chronic kidney disease, asthma and chronic obstructive pulmonary disease have been published using RCGP RSC data.^{21–24}

Study participants' eligibility

All children and adults registered with an RCGP RSC affiliated practice between 1 January 2009 and 1 January 2019 will be extracted from the RCGP RSC database and eligible for inclusion in the studies. People will be excluded if they have less than 12 months of follow-up in the database (unless aged <1 year) and if they have not had at least one consultation during the study period. People who opted out of record sharing were excluded (approximately 1.8% of the adult population). End of follow-up is the earliest of the study end date (1 January 2019), the date of patient transfer from an included practice or date of death. Transfer from a practice includes deregistration where people have moved abroad.

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. Some individuals may need to be excluded due to incomplete data, although in similar studies the exclusion rate due to incomplete data has been low.^{22–24}

Case detection and definition of variables

Definition of AD

Individuals with AD over the study period will be identified using clinical codes and prescription records in CMR systems by applying a recently validated algorithm developed in UK real-world data²⁵ and used in other recent UK studies of AD.^{7,8} In short, AD cases are identified by the presence of one diagnostic code and with at least two AD-related treatment codes on separate days. The positive predictive value of this algorithm in identifying AD has been shown to be 90% (95% CI 80% to 91%) in children and 82% (95% CI 73% to 89%) in adults.

People with new-onset (incident) AD meeting these criteria during the study period (1 January 2009 and 1 January 2019) will form an incident AD cohort. The date of onset will be defined as the date of the first AD diagnosis Read code. Patients with a diagnosis of AD prior to the study period will be excluded from this cohort. This comprises any patient with AD recorded in their record at any time prior to the study start date; we will use the complete historic record to look for any prior recording of AD. To increase certainty an AD diagnosis is incident, individuals with a diagnosis within 1 year of registering with a practice will also be excluded from the cohort, unless under 1 year old.

Definition of active AD

In a similar AD epidemiological study, Silverwood *et al.*⁷ defined active AD as the latter of two AD records appearing within any 12-month period. Active AD was then presumed to last for 1 year, unless another AD record appeared, in which case the active duration was prolonged for a further 12 months. We will modify this approach by using the first of two codes (rather than the

second as in their work) as this has been shown to have reliable agreement with physician-confirmed disease onset.²⁵

Using these criteria, a cohort of people with prevalent active AD will be defined as those fulfilling the diagnostic criteria for active AD at any point during the calendar year of analysis. Given that treatment indications are not included in the database, individuals with potentially confounding comorbidities will be excluded from the active AD cohort, in line with previous work²⁵: psoriasis, contact dermatitis, photodermatitis, ichthyosis, inflammatory bowel disease, rheumatoid arthritis or a history of organ transplantation.

Classification of AD severity

AD severity will be classified in accordance with the epidemiology work by Silverwood *et al.*⁷ AD will be classified as moderate severity when prescribed a second potent TCS within 1 year, or a first TCI. AD will be classified as severe at the start of referral to secondary care dermatology services or the start of phototherapy or systemic immunomodulatory therapy.

Sociodemographic factors

Age, gender, Index of Multiple Deprivation (IMD) and ethnicity will comprise the sociodemographic factors used for stratification of the outcome measures. Age will be initially grouped as follows: 0–1, 2–6, 7–11, 12–17, 18–29, 30–39, 40–49, 50–59, 60–69, 70–79 and ≥80 years. To examine variation across other sociodemographic factors when stratified by age, broader age groups will be used: <2, 2–11, 12–17, 18–49, ≥50. Ethnicity will be categorised into major ethnic groups as used in the 2011 census: White, Asian, Black African/Caribbean, Mixed and Other. We use an ontology to extend our ability to categorise people by ethnicity.²⁶ IMD is the official national socioeconomic status measure.²⁷ This will be derived from the patient's postcode at the point of data extraction. The resultant scores will then be stratified by deprivation quintile according to the national distribution of IMD scores.

Definition of treatments and healthcare utilisation

We will assess healthcare utilisation in AD. A primary care attendance for AD will be defined as any primary care episode associated with a diagnostic code for AD or AD medication prescription on the same day. Specialist referrals will be identified using Read codes for referral to either a dermatologist, a GP with specialist interest in dermatology or a dermatology specialist nurse.

We will extract prescription data for the following therapy classes: emollients and soap substitutes, TCS, TCI, topical antimicrobials, oral antihistamines and oral systemic treatments (ciclosporin, azathioprine, methotrexate, mycophenolate and oral corticosteroids). Given the limited recording of phototherapy use and dupilumab prescribing in primary care, rates of use of these treatments will not be explored in this study. Where

recorded in the primary care CMR these therapies will be included to identify severe AD. Differences in phototherapy between practices will be explored, and a sensitivity analysis in a subgroup of practice with high-quality recording will be performed if relevant recording difference in phototherapy is identified.

Statistical analysis: epidemiology theme

Incidence of AD

Incident cases will be defined as individuals with a first ever AD diagnostic clinical code (in the UK recorded at the time using Read code) during the study period. Those individuals with an AD diagnosis between their GP registration date and their follow-up start date will be excluded. Furthermore, to ensure AD diagnoses were new diagnoses and not new recordings of existing disease, diagnoses within 1 year of an individual's registration with a practice will not be included in the incidence calculation (unless <1 year of age). Incidence rates (per 100 person-years) will be calculated by dividing the number of new cases by the total person-years at risk of the population, to ensure we take into account the precise follow-up time of each individual in the denominator.²⁸ Incidence rates will then be stratified by age, gender, ethnic group and IMD.

Point prevalence of active AD

We will estimate the point prevalence of active AD for each calendar year across the study period. Active AD cases will be defined as those fulfilling the diagnostic criteria for active AD on 31 December of the year in question. For each calendar year, point prevalence will be calculated by dividing the number of people with active AD by the total number of actively registered patients on 31 December.

Using the 2018 data, active AD prevalence will be stratified by age, gender, ethnicity, IMD, urban/rural habitation and geographical region. The unadjusted and multivariable adjusted odds of AD by the same factors will be estimated using logistic regression and comparisons will be made to the reference category.

Statistical analysis: healthcare utilisation and treatment theme

The active AD 2018 prevalence cohort (as described above) will be used for all the following three analyses under the healthcare utilisation and treatment theme.

Primary care visits and specialist referrals

Within the active AD cohort, we will calculate annual rates of family practice attendances and specialist referrals by calendar year. Using the 2018 data, we will stratify this by age, gender, ethnic group and IMD score.

Prescribing in AD

In those with active AD, we will describe AD therapy prescriptions by medication class. The proportion of those with active AD receiving each class of medication will be calculated by dividing the number of patients with active AD prescribed at least one medication within

a particular class during the year by the total number of patients with active AD included in the study group during that calendar year.

Treatment escalation

Treatment intensification will be explored in a cohort of individuals with incident (newly diagnosed) AD over the entire study period, meaning the maximum possible follow-up for an individual in this analysis will be 10 years. In this cohort, we will analyse treatment escalation in AD at three junctures using time-to-event analysis, separately analysing children and adults. First, we will analyse the time to treatment suggestive of moderate AD using Kaplan-Meier estimator; the prescription of a second potent TCS treatment within 1 year or a first prescription for a TCI. Second, we will analyse the time to treatment suggestive of severe AD using multivariable Cox proportional hazards models, with adjustment for sex, ethnicity and deprivation quintile; the prescription of systemic immunomodulatory therapy (see above), phototherapy (if reliably recorded) or a dermatology referral. HRs and corresponding 95% CIs will be reported. Individuals with evidence of moderate or severe disease at diagnosis will be excluded from analysis of these treatment escalation endpoints. As a sensitivity analysis we will analyse the time to treatment with a systemic immunomodulatory therapy only.

Strengths and limitations of the research methods

Given that the majority of patients with eczema are seen and treated principally in primary care in the UK, primary care CMRs provide the richest existing data source for this study type. An ontological approach to AD case detection improves accuracy compared with the use of diagnostic codes alone.²³ The approach to identify AD to be used here has been validated in other similar datasets and has been shown to have an excellent positive predictive value.²⁵ Exclusion of individuals with potentially confounding conditions will also improve the accuracy of our estimates. Although the RCGP RSC database contains only data from England, it has been shown to provide a representative sample of the UK population and hence our results are probably broadly generalisable to the UK as a whole.¹⁸ General practices in the RCGP RSC network also receive feedback about data quality, making data quality as good as it gets in primary care.¹⁷ However, our results should not be extended to populations which are dissimilar.

There are also several limitations of the study design. This study does not attempt to detect undiagnosed cases of AD and hence may underestimate the true burden of disease, especially very mild AD. Furthermore, only coded data in the database can be used and hence we rely on accurate coding. Data on disease severity are not readily available from UK primary care CMR, meaning we will use a previously developed classification for primary care based on the surrogate marker of medication use,⁷ which may not fully reflect disease severity.

While several validated severity scales exist (such as the Eczema Area and Severity Index), the current primary care health record coding hierarchy does not allow these to be captured. Another limitation is the possibility of missing treatment information from secondary care. Secondary care records in the UK are frequently not computerised and prescriptions are still commonly hand-written. At present, there is no mechanism for universally capturing secondary care prescriptions and linking them to primary care records. We are therefore likely to underestimate prescription rates, particularly of systemic agents, for those with the most severe disease who are seen frequently in secondary care.

Data management

Individual patient data are pseudonymised at the point of data extraction. All data will remain in pseudonymised form and will be held on a secure network at the University of Surrey, no individual data leave this network. This network is compliant with NHS Digital Data Security and Privacy toolkit standards.

Queries will be executed to extract relevant data from the SQL database. All statistical analyses will be performed using R software.

Patient and public involvement

The design of this protocol did not involve patients or the public.

ETHICS AND DISSEMINATION

Study approval has been granted by the RCGP RSC Study Approval Committee. The study did not meet the requirements for formal ethics board review as defined using the NHS Health Research Authority research decision tool. This observational study has been registered with ClinicalTrials.gov. 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 Study Approval Committee.

We plan to publish the final study results in peer-reviewed journals and present early findings at scientific meetings.

CONCLUSIONS

This study series aims to provide a comprehensive description of the incidence, prevalence and health resource utilisation in AD in England. We anticipate that the results from these studies will be clinically beneficial for monitoring trends in AD and AD treatment, as well as determining health resource allocation and in policy development.

Twitter Simon de Lusignan @Lusignan_S

Acknowledgements The authors would like to acknowledge the project management support from Filipa Ferreira of the University of Oxford and University of Surrey.

Contributors The study concept and design was conceived and developed by CIF, CaF, HA, JD, AM and SdL. The protocol writing was directed and critically reviewed by SdL, CIF, CB, HA, JD, AM and CaF with medical writing provision from Louise Jordan at Momentum Data.

Funding This work was supported by Pfizer UK. Pfizer provided financial support to an author who is an employee of a vendor for the development of the manuscript. Editorial support was provided by Stephanie Lisa and Steve Melvin at ApotheCom, San Francisco, CA, USA, and was funded by Pfizer. HA, CB and CaF are supported by the NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

Competing interests SdL is director of RCGP RSC, and had received funding for projects from Eli Lilly, AstraZeneca, GSK, Seqirus and Takeda—all through his universities and none related to this study. JD and AM are employees of Momentum Data who were paid consultants to Pfizer in connection with the development of this manuscript. CIF is an employee of Pfizer. CaF is chief investigator of the UK National Institute for Health Research-funded TREAT (ISRCTN15837754) and SOFTER (ClinicalTrials.gov: NCT03270566) trials as well as the UK-Irish Atopic Eczema Systemic Therapy Register (A-STAR; ISRCTN11210918) and a principal investigator in the European Union Horizon 2020-funded BIOMAP Consortium (<http://www.biomap-imi.eu/>). His department has also received funding from Sanofi-Genzyme.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Simon de Lusignan <http://orcid.org/0000-0001-5613-6810>

REFERENCES

- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1789–858.
- Illi S, von Mutius E, Lau S, et al. The natural course of atopic dermatitis from birth to age 7 years and the association with asthma. *J Allergy Clin Immunol* 2004;113:925–31.
- Garmhausen D, Hagemann T, Bieber T, et al. Characterization of different courses of atopic dermatitis in adolescent and adult patients. *Allergy* 2013;68:498–506.
- Barbarot S, Auziere S, Gadkari A, et al. Epidemiology of atopic dermatitis in adults: results from an international survey. *Allergy* 2018;73:1284–93.
- Kantar Research Data. *Atopic dermatitis prevalence - UK*, 2017.
- Emerson RM, Williams HC, Allen BR. Severity distribution of atopic dermatitis in the community and its relationship to secondary referral. *Br J Dermatol* 1998;139:73–6.
- Silverwood RJ, Forbes HJ, Abuabara K, et al. Severe and predominantly active atopic eczema in adulthood and long term risk of cardiovascular disease: population based cohort study. *BMJ* 2018;361:k1786.
- Ban L, Langan SM, Abuabara K, et al. Incidence and sociodemographic characteristics of eczema diagnosis in children: A cohort study. *J Allergy Clin Immunol* 2018;141:1927–9.
- McAleer MA, Flohr C, Irvine AD. Management of difficult and severe eczema in childhood. *BMJ* 2012;345:e4770.
- National Institute of Health and Care Excellence (NICE). *Atopic eczema in under 12s: diagnosis and management. Clinical guideline [CG57]*, 2007.
- Proudfoot LE, Powell AM, Ayis S, et al. The European treatment of severe atopic eczema in children Taskforce (treat) survey. *Br J Dermatol* 2013;169:901–9.
- Flohr C, Irvine AD. Systemic therapies for severe atopic dermatitis in children and adults. *J Allergy Clin Immunol* 2013;132:774–774.e6.
- Simpson EL, Bieber T, Guttman-Yassky E, et al. Two phase 3 trials of Dupilumab versus placebo in atopic dermatitis. *N Engl J Med* 2016;375:2335–48.

- 14 Silverberg JI, Yosipovitch G, Simpson EL, *et al.* Dupilumab treatment results in early and sustained improvements in itch in adolescents and adults with moderate to severe atopic dermatitis: analysis of the randomized phase 3 studies solo 1 and solo 2, AD ADOL, and CHRONOS. *J Am Acad Dermatol* 2020;82:1328–36.
- 15 Ribero S, Giura MT, Viola R, *et al.* Effectiveness and safety of dupilumab for the treatment of atopic dermatitis in adult cohort: a real-life Italian tertiary centre experience. *J Eur Acad Dermatol Venereol* 2020;34:e380–3.
- 16 National Institute of Health and Care Excellence (NICE). *Dupilumab for treating moderate to severe atopic dermatitis. Technology appraisal guidance [TA534]*, 2018.
- 17 Pathirannehelage S, Kumarapeli P, Byford R, *et al.* Uptake of a Dashboard designed to give RealTime feedback to a sentinel network about key data required for influenza vaccine effectiveness studies. *Stud Health Technol Inform* 2018;247:161–5.
- 18 Correa A, Hinton W, McGovern A, *et al.* Royal College of general practitioners research and surveillance centre (RCGP RSC) sentinel network: a cohort profile. *BMJ Open* 2016;6:e011092.
- 19 NHS Digital. *Read codes*, 2018.
- 20 de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006;23:253–63.
- 21 Woodmansey C, McGovern AP, McCullough KA, *et al.* Incidence, demographics, and clinical characteristics of diabetes of the exocrine pancreas (type 3C): a retrospective cohort study. *Diabetes Care* 2017;40:1486–93.
- 22 Kumar S, de Lusignan S, McGovern A, *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.
- 23 Rayner L, Sherlock J, Creagh-Brown B, *et al.* The prevalence of COPD in England: an ontological approach to case detection in primary care. *Respir Med* 2017;132:217–25.
- 24 Rayner LH, McGovern A, Sherlock J, *et al.* The impact of therapy on the risk of asthma in type 2 diabetes. *Clin Respir J* 2019;13:299–305.
- 25 Abuabara K, Magyari AM, Hoffstad O, *et al.* Development and validation of an algorithm to accurately identify atopic eczema patients in primary care electronic health records from the UK. *J Invest Dermatol* 2017;137:1655–62.
- 26 Tippu Z, Correa A, Liyanage H, *et al.* Ethnicity recording in primary care computerised medical record systems: an ontological approach. *J Innov Health Inform* 2017;23:799.
- 27 Department for Communities and Local Government. *English indices of deprivation*, 2015.
- 28 Spronk I, Korevaar JC, Poos R, *et al.* Calculating incidence rates and prevalence proportions: not as simple as it seems. *BMC Public Health* 2019;19:512.