Cardiovascular risk factors and outcomes in early rheumatoid arthritis: a population-based study

Elena Nikiphorou,1,2 Simon de Lusignan,3,4 Christian D Mallen,5 Kaivan Khavandi,6 Gabriella Bedarida,6 Christopher D Buckley,7,8 James Galloway,1 Karim Raza7,9

ABSTRACT
Objective To assess the burden of cardiovascular disease (CVD) at and prior to diagnosis in people with early rheumatoid arthritis (RA) and subsequent CVD in these patients.

Methods A retrospective case–control study using a large English primary care database. People with RA (n=6591) diagnosed between 2004 and 2016 (inclusive) were identified using a validated algorithm, matched 1:1 by age and gender to those without RA (n=6591) and followed for a median of 5.4 years. We assessed differences in CVD at, before and after diagnosis, and the impact of traditional and RA-related risk factors (C reactive protein, RA-related autoantibodies and medication use) on incident CVD (a composite of myocardial infarction (MI), stroke or heart failure).

Results RA cases and their matched controls were both of mean age 58.7 (SD 15.5) at cohort entry, and 67.5% were female. Some CVD risk factors were more common at RA diagnosis including smoking and diabetes; however, total and low-density lipoprotein cholesterol were lower in patients with RA. CVD was more common in RA at cohort entry; stroke (3.9% vs 2.7%, p<0.001), heart failure (1.6% vs 1.0%, p=0.001), and non-significantly MI (3.1% vs 2.8%, p=0.092). Excess CVD developed in the 5 years preceding diagnosis. After adjustment for traditional and RA-related risk factors, RA was associated with greater risk of post-diagnosis CVD (HR 1.33, 95% CI 1.07 to 1.65, p=0.010).

Conclusions An excess of stroke and heart failure occurs before diagnosis of RA. There is excess risk for further cardiovascular events after diagnosis, which is not explained by differences in traditional CVD or RA-related risk factors at diagnosis.

INTRODUCTION
The long-term prognosis of rheumatoid arthritis (RA), in terms of joint damage and disability, has improved with the availability of conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs), biological DMARDs (bDMARDs) and targeted synthetic DMARDs1 2 in conjunction with improved monitoring and treatment adjustment to actively target disease remission.3 Despite these improvements in outcome, and possibly as a result of improved survival and greater detection, comorbidities are highly prevalent among patients with RA and are associated with premature mortality.4 8 Cardiovascular disease (CVD) is responsible for over 50% of premature deaths in RA with the risk of CVD in RA only partly explained by traditional CVD risk factors.9 10 Inflammation has been recognised as a central process in the development of CVD,11 and the excess inflammation present in RA has been hypothesised to be a determinant of the observed higher risk in RA either directly or through its impact on other CVD risk factors.12

The European League Against Rheumatism recommends annual CVD risk assessment for people with RA.13 These guidelines suggest using a multiplicative factor of 1.5 to calculate future cardiovascular risk when using a whole population derived cardiovascular risk calculator (eg, SCORE). Other cardiovascular risk assessment tools, such as the QRESEARCH Cardiovascular Risk Algorithm (QRisk) 2, has RA incorporated as a variable into the instrument (using a multiplication factor of 1.4).14 However, analyses using the RABBIT registry found a significantly greater all-cause mortality only in those with high disease activity.15 Long duration of disease may confer excess risk; however, atherosclerotic burden may precede the onset of clinically manifest RA, in parallel with inflammation.16 It therefore remains unclear whether all patients with RA are at higher CVD risk and if CVD risk has diverged from that of people without RA by the time of diagnosis.

The primary objective of this study was to describe the prevalence of established CVD risk factors at diagnosis in a contemporary RA population in England compared with age- and gender-matched controls. Secondary objectives were: (1) to explore the timeframe in which any excess events accumulate prior to diagnosis of RA when compared with controls, (2) whether RA is associated with a greater risk of CVD events within the first few years after diagnosis compared with controls, (3) to explore the association between traditional CVD risk factors, RA-related factors and cardiovascular outcomes comparing across RA cases and controls and (4) examine whether these factors account for any excess CVD risk in people with RA. We hypothesised that there would be a higher incidence of CVD in people with RA at, prior to and following diagnosis.

METHODS
Data sources and cohort
The Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) database was used for this study. It comprises pseudonymised primary care records for individuals registered with
a large network of General Practice (GP) practices, providing a broadly representative sample of the English primary care population.\textsuperscript{17} Data were available from 164 GP practices across England with a registered population of 1475762 people. The dataset contains information on clinical diagnoses, anthropometric measurements (eg, body mass index (BMI)), laboratory tests (eg, C reactive protein (CRP)) and prescriptions, coded with the Read coding system (a thesaurus of clinical terms). UK general practice lends itself to this type of study because it is a registration-based system (each patient registers with a single GP); it has been computerised since the 1990s, with electronic laboratory links, and pay-for-performance data available from 2004 that has resulted in consistent high-quality clinical data entry about chronic disease.

We included adults with a diagnosis of RA between 1 January 2004 and 1 January 2017. RA was defined using criteria that have been validated in a similar UK database, the Clinical Practice Research Datalink.\textsuperscript{18} In brief, RA was considered to be confirmed by either two disease codes consistent with a diagnosis of RA (online supplementary appendix 1) and no alternative diagnosis (psoriatic arthritis, ankylosing spondylitis or other spondyloarthropathy) or one disease code and a prescription for a DMARD with no alternative indication recorded. Once RA was identified, the index date was the date of diagnosis of RA indicated by first diagnostic code in the clinical record. Patients were required to have a minimum of 3 months between registration with their current GP and RA diagnosis to be considered a true incident case (online supplementary appendix 2).

Patients with RA were age and gender-matched (using a 1:1 ratio; online supplementary appendix 2) at GP practice level with individuals without RA. The index date for controls was defined as the RA diagnosis date of their matched counterpart.

Primary objective
Our primary objective was to compare the prevalence of CVD and CVD risk factors at diagnosis between the cases and controls. CVD comprised an a priori composite of myocardial infarction (MI), stroke and heart failure.

Risk factors were selected for inclusion based on clinical expertise and existing literature demonstrating an established association with CVD (traditional risk factors) or evidence to support an association with CVD among individuals with RA (RA-related factors).\textsuperscript{9 19–22} We explored ethnicity, smoking status, BMI, cholesterol, blood pressure and comorbidities (chronic kidney disease [CKD] stages III–V, diabetes [with breakdown by diabetes type] and atrial fibrillation). Numerical data including cholesterol, CRP and blood pressure values were derived by taking the mean of the three most recent values in the 12 months before the index date. Where three values were unavailable, the mean of two or a single available value were used. For those missing data, multiple imputation was used as outlined below. Values recorded more than 12 months before the index date were not considered except for BMI where the most recently recorded value was used. RA seropositivity was defined as seropositivity for either anti-cyclic citrullinated peptide antibodies or rheumatoid factor where recorded. The presence of medical conditions prior to diagnosis was determined by the presence of indicative Read diagnosis codes in the electronic health record at any time prior to diagnosis. This includes conditions that had been diagnosed at the patients’ previous GP practice (if relevant) and transferred with their record.

We also report the use of prescribed non-steroidal anti-inflammatory drugs (NSAIDs), glucocorticoids and DMARDs at the time the diagnosis was first recorded in primary care. This was defined as the presence of an issued prescription in the 3 months preceding or on the date of the first diagnostic code for RA. For comparison, we also report the proportion of people with NSAID, glucocorticoid and DMARD prescriptions at any point during the 3 years post-diagnosis. Anti-platelet agents were defined as aspirin or ADP receptor inhibitors, and lipid-lowering therapy was defined as statins, fibrates or ezetimibe.

Secondary objectives
For our secondary objectives, we compared: (1) the incidence of CVD in cases and controls before RA diagnosis, within 5 years, 5–10 years or more than 10 years, (2) time to the CVD composite, and each component endpoint, in cases and controls after the diagnosis date, (3) the associations between traditional and possible RA-specific CVD risk factors (see below) on the CVD composite after diagnosis and (4) whether these CVD risk factors explain any differences between CVD risk in those with and without RA. For secondary objectives 2–4, we used a matched cohort design, using the same cases and controls as in the earlier study components. Data on cause-specific mortality were not available; therefore, cardiovascular mortality was not included in the composite endpoint.

Statistical analyses
We evaluated differences in baseline characteristics between those with and without RA using McNemar’s or McNemar-Bowker tests for categorical variables and the paired t-test for continuous data. In the matched cohort components of the study, event rates were calculated as the number of events divided by the total person-years of follow-up (expressed per 1000 person-years). The follow-up period started at RA diagnosis (or date of case diagnosis for matched controls) and extended to the first of: 1 January 2017, loss to follow-up or death. Adjusted HRs and 95% CIs were estimated using Cox regression. Cumulative incidence curves were used to compare the unadjusted incidence of the composite and component outcomes in patients with and without RA.

The influence of traditional cardiovascular risk factors (age, sex, ethnicity, systolic blood pressure, low-density lipoprotein [LDL] cholesterol, diabetes and CKD stage III–V) and RA-specific factors (glucocorticoid use, disease-modifying antirheumatic drug [DMARD] use, seropositivity and CRP) on the composite CVD outcome was compared between RA and non-RA groups using separate models for the case and control cohorts. To explore possible heterogeneity in CVD risk factors between cases and controls post hoc, we generated a combined model with cases and controls and additional interaction terms between RA status and risk factors associated with CVD in either of the separate models. Two sensitivity analyses were performed for the separate models: First, by including existing CVD at baseline as binary variables to determine any impact on other risk factors. Second by using an extended composite endpoint of MI, stroke, heart failure, coronary artery disease, percutaneous coronary intervention and angina.

To explore whether differences in traditional CVD risk factors or RA-specific CVD risk factors explain any of the excess cardiovascular risk in people with RA a model with cases and controls (combined) was performed. This was done by calculating the association between RA and CVD both before and after adjustment for the traditional and RA-specific CVD risk factors described above, with stratification by matched set. For models that included incomplete data, we used multiple
Special populations

Table 1  Baseline characteristics among patients with RA at diagnosis and age-matched and sex-matched controls

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>RA (n=6591)</th>
<th>No RA (n=6591)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>58.67 (15.54)</td>
<td>58.67 (15.54)</td>
<td>NA</td>
</tr>
<tr>
<td>Male sex</td>
<td>2142 (32.5)</td>
<td>2142 (32.5)</td>
<td>NA</td>
</tr>
<tr>
<td>Ethnicity*</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>4883 (91.6)</td>
<td>4418 (94.6)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>255 (4.8)</td>
<td>147 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>128 (2.4)</td>
<td>69 (1.5)</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>26 (0.5)</td>
<td>18 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>36 (0.7)</td>
<td>18 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Smoking status*</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>1853 (30.0)</td>
<td>2247 (35.8)</td>
<td></td>
</tr>
<tr>
<td>Active smoker</td>
<td>1348 (21.9)</td>
<td>1187 (18.9)</td>
<td></td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>2968 (48.1)</td>
<td>2383 (45.2)</td>
<td></td>
</tr>
<tr>
<td>Clinical measurements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)*</td>
<td>27.67 (5.97)</td>
<td>27.23 (5.56)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Systolic BP, mm Hg*</td>
<td>130.86 (16.19)</td>
<td>131.10 (15.95)</td>
<td>0.084</td>
</tr>
<tr>
<td>Diastolic BP, mm Hg*</td>
<td>76.82 (9.68)</td>
<td>77.31 (9.46)</td>
<td>0.003</td>
</tr>
<tr>
<td>Total cholesterol (mmol/L)*</td>
<td>5.07 (1.13)</td>
<td>5.20 (1.12)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LDL cholesterol (mmol/L)*</td>
<td>2.96 (1.00)</td>
<td>3.05 (1.01)</td>
<td>0.005</td>
</tr>
<tr>
<td>HDL cholesterol (mmol/L)*</td>
<td>1.47 (0.45)</td>
<td>1.50 (0.44)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Positive RF or anti-CCP antibodies*</td>
<td>1885 (28.1)</td>
<td>159 (2.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CRP (mg/L)*</td>
<td>20.0 (32.7)</td>
<td>7.4 (16.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic kidney disease (stage III–V)</td>
<td>618 (9.4)</td>
<td>614 (9.3)</td>
<td>0.923</td>
</tr>
<tr>
<td>Diabetes (any)</td>
<td>735 (11.2)</td>
<td>635 (9.6)</td>
<td>0.004</td>
</tr>
<tr>
<td>Type 1</td>
<td>44 (0.7)</td>
<td>26 (0.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Type 2</td>
<td>691 (10.5)</td>
<td>609 (9.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>237 (3.6)</td>
<td>185 (2.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>202 (3.1)</td>
<td>169 (2.6)</td>
<td>0.079</td>
</tr>
<tr>
<td>Stroke</td>
<td>254 (3.9)</td>
<td>180 (2.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Heart failure</td>
<td>108 (1.6)</td>
<td>65 (1.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>Medication use†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE inhibitor</td>
<td>1415 (21.5)</td>
<td>1359 (20.6)</td>
<td>0.203</td>
</tr>
<tr>
<td>Angiotensin receptor blocker</td>
<td>536 (8.1)</td>
<td>528 (8.0)</td>
<td>0.816</td>
</tr>
<tr>
<td>Calcium channel blocker</td>
<td>1297 (19.7)</td>
<td>1184 (18.0)</td>
<td>0.007</td>
</tr>
<tr>
<td>Beta-blocker</td>
<td>1260 (19.1)</td>
<td>1294 (19.6)</td>
<td>0.452</td>
</tr>
<tr>
<td>Anti-platelet agents</td>
<td>1374 (20.8)</td>
<td>1371 (20.8)</td>
<td>0.962</td>
</tr>
<tr>
<td>Lipid lowering medication</td>
<td>1811 (27.5)</td>
<td>1731 (26.3)</td>
<td>0.078</td>
</tr>
<tr>
<td>NSAID</td>
<td>1962 (29.8)</td>
<td>430 (6.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Glucocorticoids</td>
<td>2001 (30.4)</td>
<td>899 (13.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>992 (15.1)</td>
<td>8 (0.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>csDMARD†</td>
<td>1263 (19.2)</td>
<td>90 (1.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>bDMARD†</td>
<td>20 (0.3)</td>
<td>0 (0.0)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Data are mean (SD) or n (%). P values from χ² test or t test are provided.

* The following variables had missing values: ethnicity in RA: n=1263; no RA: n=1921; smoking status in RA: n=422; no RA: n=319; BMI in RA: n=520; no RA: n=701; blood pressure in RA: n=191; no RA: n=221; cholesterol in RA: n=1678; no RA: n=2162; CRP in RA: n=2172; no RA: n=4499; one or more RA antibodies were measured in RA: n=3550; no RA: n=992.
†Any prescription on or within 3 months prior to the first diagnosis code in primary care.

ethical considerations

The study did not require formal research ethics committee as it used anonymised routinely collected healthcare data. This was based on the outputs from the National Health Service (NHS) Health Research Authority research decision tool (http://www.hra-decisiontools.org.uk/research/) and the University of Surrey owned Self-Assessment for Governance and Ethics tool. No patient identifiable information was available to researchers. All participating practices display information to patients regarding data use through the RCGP RSC network. All patients who choose to opt out of data sharing do not have their data processed. Approval for data access and use was granted by the RCGP RSC Research Committee.

RESULTS

The prevalence of CVD at diagnosis

Following matching, data for 13 182 individuals (6591 with RA, 6591 without RA) were included. At diagnosis, people with RA were more likely to have a history of MI, stroke or heart failure (table 1), although for MI the difference was not significant.

Individuals with RA had higher rates of some adverse CV risk factors when compared with those without RA, including history of active smoking and diabetes, but there was no difference in systolic blood pressure (table 1). Paradoxically, a higher likelihood of favourable total cholesterol and LDL cholesterol profiles was seen in people with RA (table 1). Medication use over time in the two groups is detailed in online supplementary appendix 3.

CVD incidence before diagnosis

A greater number of cardiovascular events occurred within the 5 years preceding a diagnosis of RA compared with controls. CVD incidence within 5–10 years preceding a diagnosis and more than 10 years preceding a diagnosis were not significantly different (online supplementary appendix 4).

Incidence of CVD in early RA

In the matched cohort component, during a total follow-up duration of 71 067 person years (median 5.4 years in each group), 728 incident CVD events were recorded (table 2). Event rates for the composite CVD endpoint and all individual components were greater among individuals with RA with the exception of stroke (figures 1–4).

Risk factors for CVD events after diagnosis

In separate Cox proportional hazards models for individuals with and without RA, increasing age, male sex and prevalent CKD were significant risk factors for the composite CVD endpoint after the index date (figure 5 and online supplementary appendix 5). Among individuals with RA, current smoking, BMI and diabetes were also associated with a higher rate of the composite endpoint. Lastly, seropositivity was associated with greater cardiovascular events in the non-RA population in a fully adjusted model but not in the RA population. In the post hoc model, with interaction terms between RA and variables that were significantly associated with CVD in either of the separate models described above, we found no evidence of risk factor heterogeneity (no interaction term was significant).

Sensitivity analyses with the inclusion of pre-existing CVD in the regression model for the RA group made little difference to the associations above (online supplementary appendix 6). Repeating the Cox proportional hazards models for individuals with and without RA, using the extended composite endpoint, similar associations were observed except that BMI was imputation with five iterations to generate missing values. HRs presented are those from pooled models. All statistical analyses were performed in R statistical package software version 3.2.1. Multiple imputation was performed using the MICE package.

Patient and public involvement

There was no direct patient involvement in study design, analysis, interpretation or writing.
**Table 2** Event rates of cardiovascular outcomes among individuals with and without RA

<table>
<thead>
<tr>
<th>Outcome</th>
<th>RA (n=6591)</th>
<th>Events per 1000 person-years (95% CI)</th>
<th>No RA (n=6591)</th>
<th>Events per 1000 person-years (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite endpoint</td>
<td>410</td>
<td>10.62 (9.61 to 11.69)</td>
<td>318</td>
<td>8.13 (7.26 to 9.07)</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>123</td>
<td>3.10 (2.57 to 3.69)</td>
<td>78</td>
<td>1.95 (1.54 to 2.43)</td>
</tr>
<tr>
<td>Stroke</td>
<td>170</td>
<td>4.29 (3.67 to 4.99)</td>
<td>152</td>
<td>3.82 (3.24 to 4.48)</td>
</tr>
<tr>
<td>Congestive cardiac failure*</td>
<td>158</td>
<td>3.98 (3.39 to 4.65)</td>
<td>109</td>
<td>2.72 (2.23 to 3.29)</td>
</tr>
</tbody>
</table>

*People with pre-existing heart failure were excluded from this calculation (RA: 108; no RA: 65).

RA, rheumatoid arthritis; MI, myocardial infarction.

significantly associated with the extended CVD endpoint in both groups, rather than RA only (online supplementary appendix 7).

In the overall cohort, the unadjusted HR for the association of RA with the composite endpoint was 1.40 (95% CI 1.19 to 1.64, p<0.001). After adjustment for potential confounders including traditional and RA-related factors (online supplementary appendix 8), the HR for the composite endpoint was not substantially attenuated (HR 1.33, 95% CI 1.07 to 1.65, p=0.010).

**DISCUSSION**

People with newly diagnosed RA have an excess of stroke and heart failure that occurs predominantly in the 5 years preceding their diagnosis. The profile of individuals with newly diagnosed RA is not typical of a high-risk cardiovascular cohort, with comparable levels of systolic blood pressure, with paradoxically lower Total Cholesterol and LDL cholesterol levels, compared with individuals without RA. Despite this, in the immediate period following diagnosis, we observed a near 1.4-fold higher risk of CVD.

**Higher CVD risk precedes diagnosis**

We observed more prevalent CVD among individuals at diagnosis of RA, as compared with non-RA controls. A history of stroke and heart failure was significantly more common among those with RA, and a history of MI was more common, although not significantly. Our findings suggest a significantly higher event rates for CVD in the 5-year period immediately preceding a diagnosis. These data support previous analyses using the Finnish nationwide disease register, which suggests elevated cardiovascular risk precedes the clinical onset of RA, and the Rochester Epidemiology Project, where patients were more likely to be hospitalised for MI than age-matched and sex-matched controls in the 2 years preceding a diagnosis. A large Swedish population-based study found no increased prevalence of MI, angina or ischaemic heart disease in early RA, but they did not explore stroke or heart failure.

**Subsequent CVD risk is not explained by traditional risk factors or RA characteristics at diagnosis**

People with RA had a greater CVD risk after diagnosis than those without both before (HR 1.40, 95% CI 1.19 to 1.64, p<0.001) and after adjustment (HR 1.33, 95% CI 1.07 to 1.65, p=0.010) for baseline differences. This adjustment included RA-related factors such as inflammatory markers, seropositivity and use of glucocorticoids at diagnosis. Our findings are consistent with...
another cohort study within the Rochester Epidemiology Project, which suggested higher CVD risk in early RA in people without baseline CVD after adjusting for baseline CVD risk factors.24

Implications of our findings
The higher burden of CVD prior to RA diagnosis suggests that CVD is not just a late complication of disease. From a clinical standpoint, it is important to identify and risk stratify patients accurately. Based on the higher prevalence of CVD at diagnosis of RA, and the greater risk cardiovascular events thereafter, newly diagnosed patients should undergo thorough assessment for CVD and risk factors. While it is unlikely that modification of traditional risk factors will completely mitigate the higher CVD risk in people with RA, it is nevertheless important to address known CVD risk factors in this high-risk group. Care should be taken in considering CVD risk given the lipid paradox, whereby

at a population level, those with higher LDL levels are associated with reduced cardiovascular events, thought to be explained by confounding inflammation, which follows an inverse relationship with lipids.25

Strengths and limitations
Key strengths of our study include the large population-based sample size, RA identification using a validated algorithm and identification of controls from the same GP practice as cases. While the demographics of the RCGP RSC sample are nationally representative,17 there may be some differences between those registered with a GP and those unregistered. The unregistered population may be healthier than the registered population, which may have diminished the differences we observed between RA cases and controls. The converse may also be true if the unregistered population includes a higher proportion of people with poorer health such as the homeless.

The data are observational in nature and the possibility of residual confounding cannot be excluded. In common with other studies using administrative data, our study relied on the use of comprehensive code lists. This limitation was mitigated against by using a validated approach for variable definition.26 The data analysed in this study were not linked to secondary care data and therefore relied on characteristics and outcomes from primary care records. Family history of CVD was not available. It is likely that prescription data, particularly for bDMARDs, which tend to be prescribed in secondary care, were not all captured in primary care records. Our analyses were also limited by the lack of data on cause-specific mortality. For the post hoc analysis of heterogeneity in risk factors between RA and non-RA, we may have been underpowered to detect small differences. Lastly, the reason for blood tests prior to diagnosis was not available; some investigations such as CRP will have been performed for other reasons.

Figure 3  Cumulative incidence of stroke in people with rheumatoid arthritis (RA) and matched controls without RA.

Figure 4  Cumulative incidence of heart failure in people with rheumatoid arthritis (RA) and matched controls without RA.

Figure 5  Association of cardiovascular risk factors and the composite endpoint among individuals with and without RA. BMI, body mass index; BP, blood pressure; CKD, chronic kidney disease; CRP, C reactive protein; csDMARD, conventional synthetic disease-modifying anti-rheumatic drug; LDL, low-density lipoprotein.

such as concurrent infection, and therefore, not all inflammation captured by these investigations will be attributable to RA.

CONCLUSIONS
Individuals with RA have an excess of MI, stroke and heart failure at diagnosis, although the differences with MI were not significant. This excess appears to accumulate within the 5 years prior to diagnosis. People with RA also have an excess risk of cardiovascular events (predominantly MI and heart failure) compared with age-matched and sex-matched controls without RA from diagnosis onwards. The excess CVD risk is not explained by traditional risk factors or disease biomarkers present at diagnosis.

Key messages
► Cardiovascular disease (CVD) is the major cause of premature death in people with rheumatoid arthritis, but the timing of onset and causes of this excess risk remain unclear.

What might this study add?
► We identify that excess CVD is present at, and prior to, the time of diagnosis and continues in early rheumatoid arthritis. This excess risk is not explained by differences in traditional cardiovascular risk factors in people with rheumatoid arthritis or key rheumatoid arthritis-related factors at diagnosis (C reactive protein, rheumatoid arthritis-related autoantibodies or medication use).

How might this impact on clinical practice?
► Our findings underscore the importance of performing a thorough CVD risk assessment at diagnosis and initiating early management of any identified cardiovascular risk factors.

Author affiliations
1Centre for Rheumatic Diseases, King’s College London, London, UK
2Department of Rheumatology, King’s College Hospital, London, UK
3Department of Clinical and Experimental Medicine, University of Surrey, Guildford, UK
4Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), Royal College of General Practitioners, London, UK
5Primary Care Centre Versus Arthritis, School of Primary, Community and Social Care, Keele University, Keele, UK
6Pfizer Medical Affairs, Inflammation & Immunology, International Developed Markets, Pfizer Innovative Health, Pfizer Ltd, Tadworth, UK
7Rheumatology Research Group, Institute of Inflammation and Ageing and Research into Inflammatory Arthritis Centre Versus Arthritis, University of Birmingham, College of Medical and Dental Sciences, Birmingham, UK
8University of Oxford, Kennedy Institute of Rheumatology, London, UK
9Department of Rheumatology, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK

Acknowledgements
We acknowledge additional medical writing and statistical support from Andrew McGovern and Jack Brownrigg (Momentum Data) and project management support from Filipa Ferreira (University of Surrey).

Contributors
All authors designed the study, supervised the data analysis, provided the interpretation of results and contributed to the drafting and critical review of the manuscript. All authors approved the final draft. EN is the guarantor.

Funding
Medical writing and statistical support were provided by John Dennis at Momentum Data and was funded by Pfizer. CDM is funded by the National Institute for Health Research (NIHR) Collaborations for Leadership in Applied Health Research and Care West Midlands, the NIHR School for Primary Care Research and an NIHR Research Professorship in General Practice (NIHR-RP-2014-04-026); CDB and KR are supported by the NIHR Birmingham Biomedical Research Centre.

Competing interests
EN has received speaker honoraria and has participated in advisory boards for Pfizer, Sanofi, 110 Gilead, Celltrion, AbbVie and Lilly. SDL declares no conflicts of interest. CMD declares no conflicts of interest. KK is an employee of Pfizer. GB is an employee of Pfizer. CDB declares no conflicts of interest. JG has received honoraria and/or sponsorships for conferences from AbbVie, Celgene, Janssen, Pfizer and UCB. KR has received research funding from AbbVie and Pfizer and honoraria/consultancy fees from AbbVie, Sanofi, Lilly, Bristol-Meyers Squibb, UCB, Pfizer, Janssen and Roche Chugai.

Patient and public involvement
Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication
Not required.

Provenance and peer review
Not commissioned; externally peer reviewed.

Data availability statement
Data may be obtained from a third party and are not publicly available. The Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) dataset can be accessed by bona fide researchers on a case-by-case basis. Ethical approval by the National Health Service Research Ethics Committee may be required before data requests to be considered. Researchers wishing to directly analyse the patient-level anonymised data will be required to complete information governance training and work on the data from the secure servers at the University of Surrey or University of Oxford. Further information, data request forms and contact details can be found on the RCGP RSC website: https://www.rcgp.org.uk/clinical-and-research/our-programmes/research-and-surveillance-centre.aspx.

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/0000-0002-8553-2641

REFERENCES
9 Symmons DPM, Gabriel SE. Epidemiology of CVD in rheumatoid arthritis, with a focus on RA and SLE. Nat Rev Rheumatol 2011;7:399–408.
Special populations