

1 **Searching for a prodrome for rheumatoid arthritis in the primary care record: A case-control study**  
2 **in the Clinical Practice Research Datalink**

3

4 Sara Muller<sup>1</sup>, Samantha Hider<sup>1,2</sup>, Annabelle Machin<sup>1</sup>, Rebecca Stack<sup>3</sup>, Richard A Hayward<sup>1</sup>, Karim  
5 Raza<sup>4,5</sup>, Christian Mallen<sup>1</sup>

6

7 Corresponding author: Sara Muller. David Weatherall Building, Keele University, Keele, Staffordshire,  
8 ST5 5BG, UK. T: +44 (0)1782 734842

9

10 <sup>1</sup>Research Institute for Primary Care & Health Sciences, Keele University, Keele, UK

11 <sup>2</sup>Haywood Academic Rheumatology Centre, Haywood Hospital, Stoke-on-Trent, UK

12 <sup>3</sup>College of Business Law & Social Sciences, Nottingham Trent University, Nottingham, UK

13 <sup>4</sup>Institute of Inflammation and Ageing, Arthritis Research UK Rheumatoid Arthritis Pathogenesis

14 Centre of Excellence and MRC Arthritis Research UK Centre for Musculoskeletal Ageing Research,  
15 University of Birmingham, Birmingham, UK

16 <sup>5</sup> Department of Rheumatology, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham,  
17 UK

18 **ABSTRACT**

19 **Background**

20 Rheumatoid arthritis (RA) has articular and non-articular manifestations. Early, intensive treatment  
21 has substantial benefit for both. This requires patients be identified as soon as symptoms develop.

22 **Objectives**

23 To determine whether selected signs and symptoms can be identified in the primary care records of  
24 patients prior to a formal diagnosis of RA being made and, if so, how early they can be identified.

25 **Methods**

26 A case-control study was constructed within the UK Clinical Practice Research Datalink (CPRD). 3577  
27 individuals with 'definite' RA, were matched to 14287 individuals without inflammatory arthritis. An  
28 index date was established (i.e. date general practitioner (GP) first appeared to suspect RA). Rates of  
29 consultation and consultations for suspected early RA symptoms were compared in cases and  
30 controls in the two years prior to the index date using conditional logistic regression, adjusted for  
31 number of consultations.

32 **Results**

33 The mean (standard deviation) age of participants was 58.8 (14.5) years and 66.8% were female.  
34 Rates of any consultation were significantly higher in RA cases than in controls for at least two years  
35 prior to the index date. Cases were more likely to have a pre-diagnosis coded consultation for joint,  
36 and particularly hand symptoms (aOR 11.44 (9.60, 13.63)), morning stiffness (8.10 (3.54, 18.5)),  
37 carpal tunnel syndrome (4.57 (3.54, 5.88)) and other non-articular features.

38 **Conclusions**

39 In patients who develop RA, GP consultation rates are higher for at least two years prior to the first  
40 recorded suspicion of RA. This study highlights symptoms that should raise a GP's index of suspicion  
41 for RA.

42

43 **KEY WORDS:** Rheumatoid arthritis; Diagnosis; Primary care

44 **INTRODUCTION**

45 Rheumatoid arthritis (RA) causes joint pain, stiffness and damage and can lead to excess morbidity  
46 and mortality. It has a prevalence in the UK of around 0.67% [1]. It is known that early, intensive  
47 treatment can increase the likelihood of remission and reduce long term joint damage and  
48 comorbidities [2,3].

49 Delay in making a diagnosis of RA, and therefore in treating it, can occur at a number of points in the  
50 patient journey [3-6]; first in the patient recognising their symptoms and seeking help from primary  
51 care, second in the primary care physician recognising the potential for a diagnosis of RA and making  
52 a referral to a rheumatology specialist, and third in seeing a rheumatologist and starting appropriate  
53 treatment. Work has been ongoing to understand the causes of patient delay [7-12], which has  
54 comprised in-depth studies of the symptoms reported by patients prior to their diagnosis of RA [13-  
55 16]. These symptoms have included problems with joints, fatigue, weakness [13], muscle cramps,  
56 psychological distress, and loss of motor control [14].

57 Primary care delay continues to be a significant contributor to overall diagnostic delay for people  
58 with RA [17]. This could be because GPs are not aware of the need to refer quickly [18], or because  
59 they find it difficult to identify 'red flag' symptoms, for example because of co-existing  
60 musculoskeletal conditions [19,20]. We hypothesised that we would be able to identify signs and  
61 symptoms in the coded part of the medical record that increase the likelihood of a future RA  
62 diagnosis, increasing GPs awareness and facilitating a more rapid referral.

63 We present a case-control study to assess the association of clinical features reported by patients in  
64 the earliest phases of RA with future diagnosis of the condition using the UK Clinical Practice  
65 Research Datalink (CPRD).

66

67 **MATERIALS AND METHODS**

68 **Data source: the Clinical Practice Research Datalink**

69 The CPRD is an anonymised source of routinely collected primary care health records covering  
70 approximately 6.9% of the UK's population. It is broadly representative of this population in terms of  
71 age, sex and ethnicity [21]. The data exist in coded form and include details of symptoms, diagnoses  
72 and prescriptions. Clinical data are assigned Read codes (the hierarchical clinical coding system  
73 currently used in UK primary care) by the GP. Data are collected for clinical purposes and so it can be  
74 assumed that anything the GP considers relevant might be coded, regardless of whether any clinical  
75 action was required as a result. The coding of data has been shown to be accurate for a range of  
76 conditions [22]. The CPRD assigns an 'up-to-standard' date for when a practice has a high enough  
77 quality of coding to be used for research. We use only up-to-standard data in this study.

78

#### 79 **Definition of rheumatoid arthritis**

80 Previous work in the General Practice Research Database (predecessor to CPRD) developed an  
81 algorithm to identify individuals with RA [23]. The algorithm combines Read codes for RA with  
82 prescription information and potential alternative diagnoses to create a case definition that is  
83 specific, but not overly sensitive. It has been updated to allow for the inclusion of new Read codes  
84 and the use of biological disease modifying anti-rheumatic drugs (DMARD) [24].

85

#### 86 **Analysis sample**

87 All individuals with a first Read code for RA in the CPRD between 2007 and 2012 were identified and  
88 the algorithm to define RA was applied [24]. Those who met this specific definition of RA were then  
89 matched to four individuals of the same sex and from the same practice who were born in the same  
90 3-year time interval who did not have a Read code for any inflammatory arthropathy up until the  
91 time of the case's first RA code, in order to form a case-control study.

92

#### 93 **The index date: first indication of RA in the records**

94 The date at which the first RA Read code was recorded may not accurately reflect the date at which  
95 the GP first suspected RA in the patient [25,26], as he/she may wait until diagnosis is confirmed by a  
96 specialist before entering this diagnostic code. As a proxy for the date of first clinical suspicion of RA,  
97 an index date was defined. Based on previous work in the CPRD [27,28], this index date was taken to  
98 be the earliest of the first RA code, or other code from the Read code subchapter N04 (rheumatoid  
99 arthritis and other inflammatory arthropathy), the date of the first prescription of a DMARD or first  
100 referral to rheumatology in the three years preceding the first RA Read code (Figure 1).

101

### 102 **Early signs and symptoms**

103 Signs and symptoms that may precede a diagnosis of RA were identified from the literature [13-16]  
104 and consultation with experts. Whilst some of the signs/symptoms described in the literature were  
105 not possible to define within the medical record (e.g. muscle burning as one would feel after  
106 exercise), other signs/symptoms were clearly defined syndromes and conditions that could be  
107 specified and studied in more detail (e.g. carpal tunnel syndrome, shoulder pain).

108 The final set of the signs/symptoms included is given in Tables 2 (articular) and 3 (non-articular). Lists  
109 of Read codes to define each sign/symptom, the concepts they represent and the process to achieve  
110 the lists are available at [29].

111 Where a sign/symptom was not recorded, it was assumed that the individual did not experience it,  
112 rather than data being missing.

113

### 114 **Statistical analyses**

#### 115 Rates of consultation

116 A consultation was defined as a day on which a Read coded contact with the practice was made.

117 Where there were multiple contacts/Read codes on the same day, only one was included in the

118 consultation rate analysis. Monthly rates of consultations in the 2-year period before the index date

119 were estimated and compared in cases and controls using incidence rate ratios with 95% confidence  
120 intervals (CI).

121

122 Signs and symptoms associated with RA

123 We investigated signs/symptoms in the 2-year period prior to the index date. Cumulative time  
124 periods were defined at 1, 6, 12 and 24 months prior to the index date. Each period included  
125 previous periods (Figure 1). Additionally, we considered the period 12 to 24 months before the index  
126 date to allow comparison with the period 0 to 12 months before the index date. In these analyses all  
127 Read coded contacts were considered, even when multiple codes were entered on the same date.

128

129 Conditional logistic regression was used to assess the association between signs/symptoms and case  
130 control (RA) status, allowing for the matched design in each of the time periods described above.

131 Results are presented as unadjusted odds ratios and then adjusted for the rate of consultations (as  
132 defined above) in the period in question. All values are presented as odds ratios (OR) with 95% CI.

133

134 Data management and analyses were conducted in Stata 14.2 [30].

135

136 Approval for this study was granted by the Independent Scientific Advisory Committee of the CPRD  
137 (reference 13\_126).

138

## 139 **RESULTS**

### 140 **Sample**

141 Between 2007 and 2012, 4161 people were identified with a first Read code for RA. Of these, 3577  
142 met the criteria for RA using the algorithm described above [24] and were matched to 14287  
143 controls (Table 1). The mean (standard deviation) age of the both cases and controls was 58.8 (14.5)

144 years and 66.8% were female (Table 1). Current and previous smoking were more common in cases  
145 than controls.

146

#### 147 **Numbers of consultations**

148 For two years prior to the index date, the overall consultation rate was significantly higher in each  
149 month in cases than in controls (Figure 2) (incidence rate ratio 1.22 (1.21, 1.22)). This increase  
150 became more pronounced in the 6 months before the index date and in the final month before,  
151 cases consulted at 2.68 (95% CI 2.61, 2.76) times the rate of controls (mean: 2.39 consultations per  
152 person).

153

#### 154 **Signs and symptoms preceding a diagnosis of RA**

##### 155 Articular symptoms

156 All articular signs/symptoms were associated with a future diagnosis of RA in all time periods in both  
157 unadjusted and adjusted analyses, with the exception of jaw pain which was not significantly  
158 associated with RA following adjustment for the number of consultations in the 0 to 1-month period  
159 (Table 2). In the 0 to 1-month period, joint symptoms (adjusted OR (95% CI), 14.82 (12.48, 17.60)),  
160 and specifically hand problems (61.07 (31.58, 118.10)), were strongly associated with the  
161 development of RA. Palindromic rheumatism occurred only in cases in the 12 months preceding the  
162 index date. The strength of all associations was lessened by adjustment for the number of  
163 consultations (except for jaw pain in the 0 to 6-month period) and associations were generally  
164 stronger for consultations closer to the index date.

165

##### 166 Non-articular symptoms

167 In the 0 to 1-month period there were four non-articular signs/symptoms that had large (odds  
168 ratio $\geq$ 6) unadjusted associations with development of RA (morning stiffness, muscle pain: 13.83  
169 (5.11, 37.42) and carpal tunnel syndrome: 2.96 (1.38, 6.34)). All remained strongly and significantly

170 associated after adjustment (Table 3). Morning stiffness was recorded in 14 cases (0.39%), but not in  
171 any controls, hence an estimate of the strength of association could not be made. Unintentional  
172 weight loss was not significantly associated with RA in the month before the index date, but was in  
173 all other time periods (except adjusted analysis in 12 to 24-month period (Supplementary tables)). In  
174 all time periods, there was a significant unadjusted association between fatigue and development of  
175 RA, but this association was attenuated and not significant after adjustment for number of  
176 consultations. No association was seen with sleep problems or flu-like illness (Table 3).  
177 Psychological problems were significantly associated with a higher odds of RA in all unadjusted  
178 analyses (except 0 to 24 months before the index date, where association was positive, but not  
179 significant), but significantly associated with a decreased odds of RA after adjustment.

180

181 Comparison of consultation 0-12 and 12-24 months before the index date

182 Comparison of the associations of signs/symptoms during the periods 0 to 12-months and 12 to 24-  
183 months before the index date suggested that signs/symptoms grouped together (Supplementary  
184 tables). A similar pattern of association was seen in both time periods for fatigue, altered sensations,  
185 postnatal occurrence of RA, weakness, psychological problems and carpal tunnel syndrome. There  
186 was no unadjusted association between sleep and RA in either time period and only after  
187 adjustment in the 12 to 24-month period for falling. Flu-like symptoms were associated with RA in  
188 the 12 to 24-month period (adjusted analyses only (1.46 (1.04, 2.07))), but not closer to the index  
189 date. All articular signs/symptoms and the remaining non-articular signs/symptoms were more  
190 strongly associated with RA in the 0 to 12-month period than in the 12 to 24-month period. This was  
191 particularly noticeable for hand symptoms (aOR 0 to 12-months: 23.75 (18.49, 30.51); 12 to 24-  
192 months: 2.70 (2.05, 3.56)), morning stiffness (0 to 12-months: 9.72 (3.84, 24.60); 12 to 24-months:  
193 0.93 (0.08, 10.64)) and muscle pain (0 to 12-months: 3.15 (2.22, 4.47); 12 to 24-months: 1.22 (0.75,  
194 1.97)).

195



196 **DISCUSSION**

197 The rate of consultations increases rapidly in the period before the index date in those with RA  
198 compared to controls, and key signs and symptoms are recorded at a higher rate before an RA  
199 diagnosis. In the final month before the index date, these include all joint symptoms, but particularly  
200 those involving the hand, and the non-articular symptoms morning stiffness, muscle pain and carpal  
201 tunnel syndrome. In longer periods before the index date, there is also an increase in the recording  
202 of other features such as unintentional weight loss. Other symptoms reported by patients in  
203 previous studies (e.g. fatigue, cramping, poor sleep) [13-16], showed less clear associations.

204

205 The strengths of our study include the large sample size and use of a validated definition of RA [24].  
206 Whilst it could be argued that the definition had good specificity at the expense of sensitivity, the  
207 exclusion of controls with any record of inflammatory arthropathy should reassure that there was no  
208 contamination of the control group with potential cases. The data for this study were taken from a  
209 high quality database containing a representative sample of individuals in UK primary care. As such,  
210 the results should be generalizable to other primary care settings. Despite its strengths, this study  
211 also has some weaknesses. First, multiple statistical testing, which could result in false positive  
212 associations. Second, we do not know the thought processes of the GPs who coded the  
213 consultations and how this might have affected our findings. This question cannot be answered in  
214 routinely collected data, but would require in-depth interviews with GPs as to their views and clinical  
215 practice. This in itself may prove difficult, as an individual GP will see a new case of RA only rarely  
216 and may not be able to report what action they would take [20].

217

218 We adjusted our findings for the total number of consultations (days with  $\geq 1$  Read coded  
219 consultation) in order to adjust for ascertainment/surveillance bias, whereby the presence of the  
220 patient in the surgery makes it easier for the GP to identify and code signs/symptoms. This  
221 adjustment for number of consultations attenuated the association of a number of signs/symptoms

222 (physical functioning, cramps, weakness and restless legs) with RA. This may suggest that these  
223 signs/symptoms are more common in those with RA, but are only recorded in those who attend  
224 more frequently. A similar process may explain the change in direction of association with  
225 psychological problems when adjusting for number of consultations: people with coded  
226 psychological problems consult more frequently and it may be that controls receive psychological  
227 codes, but RA cases receive codes for physical symptoms because these take priority for the GP.  
228 Due to small numbers, non-significant associations in the final month before the index date should  
229 be interpreted with caution, as they may well represent a type II statistical error, especially where  
230 the signs/symptom was associated with RA in longer time periods and the absolute estimate of the  
231 size of association is similar across time periods (e.g. unintentional weight loss). However, it could be  
232 that these symptoms are simply more common at an earlier stage in the pre-clinical picture of RA  
233 and become less commonly reported or over-shadowed by other symptoms in the final weeks  
234 before the GP suspects RA.

235

236 Previous literature has described the symptoms patients report before a diagnosis of RA [13-16], and  
237 there is a feeling among rheumatologists that they know what symptoms they expect to see in early  
238 RA. However, to our knowledge this is the first paper to consider whether these signs and symptoms  
239 occur in the primary care record, whether they are more common in those who later received a  
240 diagnosis of RA than in those who do not and how long before RA is suspected by the GP these signs  
241 and symptoms are present.

242

243 Within this study, classical features of RA such as hand pain and stiffness were more frequently  
244 coded within the primary care record, and were seen more frequently up to 2 years before the index  
245 date. However, musculoskeletal symptoms in regions not traditionally associated with early RA (e.g.  
246 neck and shoulder pain) were also reported more frequently by patients who eventually developed  
247 RA. Joint symptoms, particularly in the hands, and other well-recognised non-articular features

248 should raise the index of suspicion of RA in patient presenting in primary care, particularly when  
249 accompanied by a general increase in patient contact with the primary care. However, GPs should  
250 also be aware that these features have low specificity and only a small proportion of patients with  
251 these symptoms will go on to receive a diagnosis of RA. For example, whilst we have confirmed that  
252 people with RA are more likely to have hand symptoms, an RA outcome is seen in only a minority of  
253 patients that have hand symptoms recorded in primary care. Further studies will be needed to  
254 investigate what other symptoms/signs increase the likelihood of an RA diagnosis.

255 Other early symptoms reported by patients such as falls and sleep problems did not show any  
256 association with RA. This may represent a true lack of association, or it may be that either patients  
257 did not report these symptoms to the GP or that GPs did not code them, especially if they did not fit  
258 with the GP's concept of what is important.

259 The association of flu-like symptoms with RA only in the 12 to 24-month period may suggest that  
260 rather than being part of an RA prodrome, flu-like symptoms may be a marker of an insult on the  
261 immune system that reflects the phase of immune tolerance breakage [31].

262

263 The next steps should be to identify groups of symptoms that constitute a prodrome of RA and at  
264 the same time educate GPs as to the key symptoms that may indicate RA prior to the cardinal  
265 symptoms of morning stiffness and hand symptoms that they already appear to recognise, although  
266 further work is needed to refine the specificity of these common symptoms.

267 In the future, it may be possible to create automated electronic alerts for the GP within the records  
268 system that highlight the risk for an individual patient when certain codes are entered. This already  
269 happens for example to alert the GP to the possibility of sepsis.

270

271 What our study was not able to do was to identify new signs or symptoms from the record that may  
272 occur at a higher rate in those who go on to receive an RA diagnosis than in those who do not; to do

273 so would have required an alternative methodological approach to identify patterns in consultation  
274 that were not defined by code lists (e.g. 32]).

275

## 276 **CONCLUSION**

277 We have provided definitive evidence of the presence of some key features of early RA in the  
278 primary care medical record prior to the GP appearing to recognise the condition. Primary care  
279 professionals should be aware of the range of articular and non-articular features, specifically hand  
280 symptoms, muscle pain, carpal tunnel syndrome and unintentional weight loss, accompanied by an  
281 increased rate of consultation, as potentially forming a prodromal syndrome for RA. Increased  
282 awareness of these symptoms combined with education on the need for early referral could  
283 facilitate earlier treatment of RA, increasing the likelihood of remission and reducing long term joint  
284 damage and comorbidities.

285

## 286 **CONFLICT OF INTERESTS**

287 KR reports personal fees from BMS, personal fees from Abbvie, grants from Pfizer, personal fees  
288 from Pfizer, personal fees from UCB, outside the submitted work. The other authors report no  
289 conflicts of interest.

290

## 291 **ACKNOWLEDGEMENTS**

292 The authors are grateful to Dr Alyshah Abdul Sultan and Mrs Rebecca Whittle for their help in  
293 processing the CPRD data.

## 294 **Funding**

295 SM and AM are funded by the National Institute of Health Research School for Primary Care  
296 Research. CDM is funded by the National Institute for Health Research (NIHR) Collaborations for  
297 Leadership in Applied Health Research and Care West Midlands, the NIHR School for Primary Care  
298 Research and a NIHR Research Professorship in General Practice (NIHR-RP-2014-04-026). KR is

299 supported by the NIHR Birmingham Biomedical Research Centre. The views expressed are those of  
300 the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

301

## 302 **References**

- 303 1. Abhishek A, Doherty M, Kuo CF, Mallen CD, Zhang W, Grainge MJ. Rheumatoid arthritis is  
304 getting less frequent-results of a nationwide population-based cohort study. *Rheumatology*  
305 *(Oxford)* 2017; **56**: 736-44.
- 306 2. Moura CS, Abrahamowicz M, Beauchamp ME, Lacaille D, Wang Y, Boire G, et al. Early  
307 medication use in new-onset rheumatoid arthritis may delay joint replacement: results of a  
308 large population-based study. *Arthritis Res Ther* 2015; **17**: 197.
- 309 3. van der Linden MP, le Cessie S, Raza K, van der Woude D, Knevel R, Huizinga TW, et al. Long-  
310 term impact of delay in assessment of patients with early arthritis. *Arthritis Rheum* 2010; **62**:  
311 3537-46.
- 312 4. Barhamain AS, Magliah RF, Shaheen MH, Munassar SF, Falemban AM, Alshareef MM, et al.  
313 The journey of rheumatoid arthritis patients: a review of reported lag times from the onset  
314 of symptoms. *Open Access Rheumatol* 2017; **9**: 139-50.
- 315 5. Kumar K, Daley E, Carruthers DM, Situnayake D, Gordon C, Grindulis K, et al. Delay in  
316 presentation to primary care physicians is the main reason why patients with rheumatoid  
317 arthritis are seen late by rheumatologists. *Rheumatology (Oxford)* 2007; **46**: 1438–40.
- 318 6. Feldman DE, Bernatsky S, Houde M, Beauchamp ME, Abrahamowicz M. Early consultation  
319 with a rheumatologist for RA: does it reduce subsequent orthopaedic use? *Rheumatology*  
320 *(Oxford)* 2013; **52**: 452–9.
- 321 7. Simons G, Lumley S, Falahee M, Kumar K, Mallen CD, Stack RJ, et al. The pathway to  
322 consultation for rheumatoid arthritis: exploring anticipated actions between the onset of  
323 symptoms and face-to-face encounter with a healthcare professional. *BMC Musculoskelet*  
324 *Disord* 2017; **18**: 258.

- 325 8. Simons G, Mason A, Falahee M, Kumar K, Mallen CD, Raza K, et al. Qualitative Exploration of  
326 Illness Perceptions of Rheumatoid Arthritis in the General Public. *Musculoskeletal Care* 2017;  
327 **15**: 13-22.
- 328 9. Mølbaek K, Hørslev-Petersen K, Primdahl J. Diagnostic Delay in Rheumatoid Arthritis: A  
329 Qualitative Study of Symptom Interpretation Before the First Visit to the Doctor.  
330 *Musculoskeletal Care* 2016; **14**: 26-36.
- 331 10. Van der Elst K, De Cock D, Vecoven E, Arat S, Meyfroidt S, Joly J, et al. Are illness perception  
332 and coping style associated with the delay between symptom onset and the first general  
333 practitioner consultation in early rheumatoid arthritis management? An exploratory study  
334 within the CareRA trial. *Scand J Rheumatol* 2016; **45**: 171-8.
- 335 11. Molina E, Del Rincon I, Restrepo JF, Battafarano DF, Escalante A. Association of  
336 socioeconomic status with treatment delays, disease activity, joint damage, and disability in  
337 rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2015; **67**: 940-6.
- 338 12. Simons G, Mallen CD, Kumar K, Stack RJ, Raza K. A qualitative investigation of the barriers to  
339 help-seeking among members of the public presented with symptoms of new-onset  
340 rheumatoid arthritis. *J Rheumatol* 2015; **42**: 585-92.
- 341 13. Stack RJ, Sahni M, Mallen CD, Raza K. Symptom complexes at the earliest phases of  
342 rheumatoid arthritis: a synthesis of the qualitative literature. *Arthritis Care Res (Hoboken)*  
343 2013; **65**: 1916-26.
- 344 14. Stack RJ, van Tuyl LH, Sloots M, van de Stadt LA, Hoogland W, Maat B, et al. Symptom  
345 complexes in patients with seropositive arthralgia and in patients newly diagnosed with  
346 rheumatoid arthritis: a qualitative exploration of symptom development. *Rheumatology*  
347 *(Oxford)* 2014; **53**: 1646-53.
- 348 15. van Tuyl LH, Stack RJ, Sloots M, van de Stadt LA, Hoogland W, Maat B, et al. Impact of  
349 Symptoms on Daily Life in People at Risk of Rheumatoid Arthritis. *Musculoskeletal Care*  
350 2016; **14**: 169-73.

- 351 16. Newsum EC, de Waal MW, van Steenberg HW, Gussekloo J, van der Helm-van Mil AH.  
352 How do general practitioners identify inflammatory arthritis? A cohort analysis of Dutch  
353 general practitioner electronic medical records. *Rheumatology (Oxford)* 2016; **55**: 848-53.
- 354 17. Healthcare Quality Improvement Partnership. A patient and public guide to the National  
355 Clinical Audit for Rheumatoid and Early Inflammatory Arthritis 1st Annual Report 2015 (Data  
356 collection: 1 February 2014 – 30 April 2015). Available at:  
357 [https://www.hqip.org.uk/public/cms/253/625/19/456/WEB-](https://www.hqip.org.uk/public/cms/253/625/19/456/WEB-BSR%20Patient%20Report%202015.pdf?realName=SMFSaA.pdf&v=0)  
358 [BSR%20Patient%20Report%202015.pdf?realName=SMFSaA.pdf&v=0](https://www.hqip.org.uk/public/cms/253/625/19/456/WEB-BSR%20Patient%20Report%202015.pdf?realName=SMFSaA.pdf&v=0) Accessed December  
359 2017
- 360 18. Patient view on behalf of the National Audit Office. People with Rheumatoid Arthritis, their  
361 carers, and the NHS. A national survey of patients with RA and their carers. 2009. Available  
362 at: [https://www.nao.org.uk/wp-content/uploads/2009/07/0809823\\_Patient\\_survey.pdf](https://www.nao.org.uk/wp-content/uploads/2009/07/0809823_Patient_survey.pdf)  
363 Accessed December 2017
- 364 19. Lard LR, Huizinga TW, Hazes JM, Vliet Vlieland TP. Delayed referral of female patients with  
365 rheumatoid arthritis. *J Rheumatol* 2001; **28**: 2190-2.
- 366 20. Meyfroidt S, Stevens J, De Lepeleire J, Westhovens R, De Cock D, Van der Elst K, et al. A  
367 general practice perspective on early rheumatoid arthritis management: A qualitative study  
368 from Flanders. *Eur J Gen Pract* 2015; **21**: 231-7.
- 369 21. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource  
370 Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015; **44**: 827-36.
- 371 22. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses  
372 in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010;  
373 **69**: 4-14.
- 374 23. Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for  
375 rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research  
376 database? *Arthritis Rheum* 2008; **59**: 1314-21.

- 377 24. Muller S, Hider SL, Raza K, Stack RJ, Hayward RA, Mallen CD. An algorithm to identify  
378 rheumatoid arthritis in primary care: a Clinical Practice Research Datalink study. *BMJ Open*  
379 2015; **5**: e009309.
- 380 25. Ford E, Carroll J, Smith H, et al. What evidence is there for a delay in diagnostic coding of RA  
381 in UK general practice records? An observational study of free text. *BMJ Open* 2016; **6**:  
382 e010393.
- 383 26. Ford E, Nicholson A, Koeling R, Davies K, Koeling R, Petersen I, et al. Optimising the use of  
384 electronic health records to estimate the incidence of rheumatoid arthritis in primary care:  
385 what information is hidden in free text? *BMC Med Res Methodol* 2013; **13**: 105.
- 386 27. Nicholson A, Ford E, Davies KA, Smith HE, Rait G, Tate AR, et al. Optimising use of electronic  
387 health records to describe the presentation of rheumatoid arthritis in primary care: a  
388 strategy for developing code lists. *PLoS One* 2013; **8**: e54878.
- 389 28. Tate AR, Martin AG, Murray-Thomas T, Anderson SR, Cassell JA. Determining the date of  
390 diagnosis--is it a simple matter? The impact of different approaches to dating diagnosis on  
391 estimates of delayed care for ovarian cancer in UK primary care. *BMC Med Res Methodol*  
392 2009; **9**: 42.
- 393 29. Medical Record Data Research. 2018. keele.ac.uk/mrr. Accessed 23 Mar 2018.
- 394 30. StataCorp. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP, 2015.
- 395 31. Tracy A, Buckley CD, Raza K. Pre-symptomatic autoimmunity in rheumatoid arthritis: when  
396 does the disease start? *Semin Immunopathol* 2017; **39**: 423-35.
- 397 32. Zhou SM, Fernandez-Gutierrez F, Kennedy J, Cooksey R, Atkinson M, Denaxas S, et al.  
398 Defining Disease Phenotypes in Primary Care Electronic Health Records by a Machine  
399 Learning Approach: A Case Study in Identifying Rheumatoid Arthritis. *PLoS One* 2016; **11**:  
400 e0154515.  
401



402 Corresponding author: Sara Muller, Research Institute for Primary Care & Health Sciences, David  
403 Weatherall Building, Keele University, Keele, UK. s.muller@keele.ac.uk  
404 Second corresponding author: Samantha Hider, Research Institute for Primary Care & Health  
405 Sciences, David Weatherall Building, Keele University, Keele, UK. s.hider@keele.ac.uk

406 **Figure Legends**

407 Figure 1 Schematic representation of data set

408 Figure 2 Consultation rates (per person year) in cases and controls prior to index date