The use of digital data to investigate the management of rheumatoid arthritis

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medicine, Biology and Health.

2021

Ruth E Costello

School of Biological Sciences

Contents

Word count: 15,054 words.

List of Tables

List of Figures

Abstract

The University of Manchester

Candidate name: Ruth E Costello

Degree title: PhD by published works

Title of thesis: The use of digital data to investigate the management of rheumatoid arthritis

Year: 2021

Increasingly there are large amounts of digital data that can be harnessed for epidemiological research. The theme of this thesis is to describe how digital data, in particular electronic health records (EHR) and data collected through an online health community, can be used to answer questions about the management of rheumatoid arthritis (RA). These two data sources have different strengths - EHRs produce large longitudinal datasets representing the whole of the UK and can be linked to other administrative datasets. For bespoke surveys, collecting data through an online health community is a quick and efficient method to identify issues important to patients. Both have limitations that need addressing through careful use of epidemiological techniques.

This thesis presents seven studies that highlight these methodological challenges focused on two interventions used to manage RA: vaccinations and glucocorticoids. Primary care EHR data was used to estimate influenza and pneumococcal vaccination uptake the UK (publication 1) and timing of pneumococcal vaccination in relation to starting disease-modifying anti-rheumatic drugs (publication 2), where the main challenges addressed were related to misclassification. EHR data was used to answer questions related to adverse outcomes associated with glucocorticoids: mortality, overall (publication 3) and in people with comorbid type 2 diabetes mellitus (publication 4), and hypertension (publication 5). These studies required careful preparation of drug data to correctly attribute risk and consideration of biases such as peri-mortal bias and surveillance bias. Data collected through an online health community measured patient perspectives of glucocorticoid side effects (publication 6) and the challenge of representativeness of responders was directly addressed through capturing the characteristics of patients with RA and comparing them to patients with RA identified using EHR data to understand the representativeness of online health community responders (publication 7).

The studies highlighted important clinical issues: 1) vaccination uptake was not occurring as guidelines recommend, 2) glucocorticoid use was associated with increased risk of mortality and hypertension, 3) the glucocorticoid side effects of importance to patients are not frequently researched and 4) patients recruited through online health communities represent a younger and more diverse RA population. The thesis shows how digital data can be used, to successfully address a variety of previously unanswered questions, related to different aspects of RA management, using methods applicable to the investigation of other chronic diseases.

Candidate Declaration

The University of Manchester

PhD by published work candidate declaration

Candidate name: Ruth E Costello

Faculty: Faculty of Biology, Medicine and Health

Thesis title: The use of digital data to investigate the management of rheumatoid arthritis

1) the nature and extent of your own contribution and the contribution of co-authors and other collaborators to each of the publications presented

Publication 1: Professor William Dixon (WGD) and Professor Kevin Winthrop (KW) conceived the study. WGD, KW and I designed the study. I was responsible for obtaining CPRD approval for the study, preparing the data, conducting the analysis, preparing the manuscript and submitting it for publication. All authors interpreted the results and critically revised the manuscript. In addition, I prepared abstracts for international conferences and gave an oral presentation and poster presentations.

Publication 2: WGD conceived the idea. All authors designed the study. I was responsible for preparing the data, conducting the analysis, preparing the manuscript and submitting it for publication. All authors interpreted the results and critically revised the manuscript. In addition, I prepared the abstract for an international conference and gave a poster presentation.

Publication 3: Dr Mohammad Movahedi was the lead author for this publication, he conducted the initial analysis, which I contributed to, and drafted the manuscript. Dr Movahedi left his post before the manuscript was submitted and I took on responsibility for the project. Subsequently I made significant contributions to preparing the manuscript. All authors interpreted the results and critically revised the manuscript. I submitted the

manuscript for publication and was responsible for reviewer comments and additional analysis suggested by the reviewers.

Publication 4: WGD conceived the idea. WGD, Professor Richard Emsley, Dr Antonia Marsden and I designed the study. I prepared the data, conducted the analysis, prepared the manuscript and submitted the manuscript for publication. All authors interpreted the results and critically revised the manuscript. In addition, I prepared an abstract for an international conference and gave an oral presentation.

Publication 5: WGD conceived the study. WGD, Dr Meghna Jani, Dr Belay Birlie Yimer and I designed the study. I prepared the data, conducted the analysis, prepared the manuscript and submitted the manuscript for publication. All authors interpreted the results and critically revised the manuscript. In addition, I prepared an abstract for international conferences and gave poster presentations.

Publication 6: WGD conceived the study and designed the survey. I contributed comments on the survey design. WGD, Professor John McBeth (JMcB), Dr Jenny Humphreys and I designed the study. I was responsible for submitting the study for ethics approval. I conducted the analysis and prepared the manuscript. All authors interpreted the results and critically revised the manuscript. In addition, I prepared an abstract for an international conference and gave an oral presentation.

Publication 7: WGD conceived the study. WGD, JMcB and I were responsible for the study design. I was responsible for obtaining CPRD approval, ethics approval, preparing the data, conducting the analysis and preparing the manuscript. All authors interpreted the results and critically revised the manuscript. In addition, I prepared an abstract for an international conference and gave a poster presentation.

2) All work presented was completed whilst a member of staff at the University of Manchester.

3) None of the work presented has been submitted in support of any other degree or qualification at this, or any other University, or of any professional or learned body.

I confirm that this is a true statement and that, subject to any comments above, the submission is my own original work.

Signed: Ruth Costello…….. Date: 17/05/2021…..……..……

Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses

Acknowledgements

I would like to thank everyone who has supported me throughout this work, my colleagues, past and present with special thanks to Therese and Mike with whom I have shared offices with at different points and been able to have a chat about my work, and life in general! Thanks also to my supervisors Will and Jenny for their advice, support and encouragement, particularly over the past year.

Thanks to my friends and family, with a special thanks to my parents, who are no longer with us, but I hope this would make them proud. Finally, I want to thank my husband, Kev, whom I sadly lost while writing this thesis and will always miss, he gave me lots of support and encouragement, and I will always be thankful for that.

Statement

i) particulars of the candidate's degrees, other qualifications and research experience, including all particulars required to establish eligibility under the University's regulations;

Employment history

June 2014 – Present	Research assistant, The University of Manchester
April 2013-June 2014	Clinical Trial Coordinator, University of Liverpool
February 2012-April 2013	Clinical Trials Support Officer, University of Liverpool
January 2010-February 2012	Postgraduate Medicine Administrator, Imperial College

(Part-time while studying)

June 2007-January 2010	Trial Dementia Coordinator, Imperial College
January 2005-June 2007	Clinical Trials Officer, Imperial College

Qualifications

2008 – 2012	London School of Hygiene and Tropical Medicine (distance learning)
	MSc Epidemiology
1998 - 2001	Southbank University
	BSc (Hons) 2.2 Human Biology and Psychology

Courses attended

July 2014: CPRD GOLD training day, London

April 2015: Getting the most out of research conferences, University of Manchester

April 2015: Bordeaux Pharmacoepidemiology festival, Université de Bordeaux.

February 2016: Turbocharge your writing, University of Manchester

July 2016: Advanced Epidemiology summer school (5 days), Manchester

November 2016: Introduction to version control using git, University of Manchester

November 2017: Data analysis using R, University of Manchester

May 2018: Software carpentry, University of Manchester

October-December 2018: EULAR peer review training, online.

December 2018: Observational Health Data Sciences and Informatics (OHDSI) study-athon, University of Oxford

December 2019: 4th International Conference on Administrative Data Research, Cardiff.

May 2020: Why successful people often feel like frauds - The Imposter Syndrome, online.

Research experience

Led on 13 abstracts for national and international rheumatology conferences, with 6 oral presentations and 7 poster presentations.

Peer reviewed 7 articles for a variety of international journals.

Presented at seminars within the Centre for Epidemiology.

Presented at a public engagement event at the University of Manchester.

Teaching experience

Co-supervised an APEP student project.

ii) a complete and numbered list of the publications submitted (grouped according to subject and type);

1. Costello R, Winthrop KL, Pye SR, Brown B, Dixon WG (2016) Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink. PLoS ONE 11(4): e0153848. https://doi.org/10.1371/journal.pone.0153848
2. Costello RE, Humphreys JH, Winthrop KL, Dixon WG, Pneumonia vaccination timing in relation to starting conventional synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis, Annals of the Rheumatic Diseases 2020;79:1665-1666
3. Movahedi M, Costello R., Lunt M, Pye SR, Sergeant JC, Dixon WG, Oral glucocorticoid therapy and all-cause and cause-specific mortality in patients with rheumatoid arthritis: a retrospective cohort study. Eur J Epidemiol 31, 1045–1055 (2016). https://doi.org/10.1007/s10654-016-0167-1
4. Costello RE, Marsden A., Movahedi M, Lunt M, Humphreys JH, Emsley R, Dixon WG. The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study. BMC Rheumatol 4, 4 (2020). https://doi.org/10.1186/s41927-019-0105-4
5. Costello RE, Yimer BB, Roads P, Jani M, Dixon WG, Glucocorticoid use is associated with an increased risk of hypertension, Rheumatology, , keaa209, https://doi.org/10.1093/rheumatology/keaa209
6. Costello R, Patel R, Humphreys J, McBeth J, Dixon WG, Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community, BMJ Open 2017;7:e014603. doi: 10.1136/bmjopen-2016-014603

7. Costello R, Jacklin C, Jameson Evans M, McBeth J, Dixon WG, Representativeness of a digitally engaged population and a patient organisation population with rheumatoid arthritis and their willingness to participate in research: a cross-sectional study, RMD Open 2018;4:e000664. doi: 10.1136/rmdopen-2018-000664

iii) an overall summary of the aims and achievement of the work, for which the publications submitted give evidence.

Chapter 1: Introduction

1.1 Thesis overview

This thesis will present seven studies with the theme of using digital data to answer questions around the management of rheumatoid arthritis (RA). Five of the studies presented use a research database derived from electronic health records (EHR) and the final two use an online health community to collect survey data. Within the thesis, I will discuss where my work fits in the current literature, the methodological challenges of using these types of data and how I overcame some of these challenges within my studies. The work presented in this thesis was completed while working as a research assistant in the Centre for Epidemiology so has also been shaped by the needs of the Centre. While the research questions may be less explicitly connected compared to a thesis planned from the start, there are consistent and recurring themes of digital data methods and RA treatments.

1.2 Chapter overview

This chapter will introduce digital data, its opportunity and use as a data source within epidemiology and methodological challenges of using these types of data. I will then introduce the subject area of RA, describing the disease itself and treatments used. I will finish with the overall aim of the thesis.

1.3 The use of digital data in epidemiological studies

Epidemiology is defined as “the study of the distribution and determinants of health-related states or events in specified populations and the application of this study to control of health problems” [1]. Studies within this field are broadly described as either interventional or observational. Interventional studies assign patients to a specific intervention such as a

medication or a non-pharmacological therapy, and outcomes are measured. Observational studies observe people without intervening. There are benefits and drawbacks to both types of study.

Interventional studies, such as randomised controlled trials (RCT) are essential for determining medication efficacy. Efficacy refers to how well an intervention works in ideal conditions. In RCTs, patients who meet study inclusion criteria are randomised to receive either the treatment under study or comparator, and outcomes are compared. Randomisation results in the balancing of characteristics, measured and unmeasured, allowing causal conclusions. The results of these studies provide evidence for licensing drugs. The major limitations of RCTs are: 1) the studies often use a specific population, although they will all have the disease of interest there are often restrictions on age and comorbidities, this means results may not generalise to all people with the disease of interest, 2) the studies usually have a short duration, which is not long enough to identify longer-term side effects and 3) the studies usually lack the power to study rare adverse events.

Observational epidemiology studies are essential to help answer questions arising due to these limitations, and determining medication effectiveness. Effectiveness refers to how well an intervention works in real-world conditions. The study designs frequently used are cross-sectional studies, case-control studies and cohort studies. Cross-sectional studies observe people at one point in time. Cohort and case-control studies observe people over time. In case-control studies, people with the outcome of interest are identified as the cases and then matched to controls (i.e. those without the outcome), on factors such as age and gender. Exposures are usually identified retrospectively, and associations with the outcome of interest are measured. This is not the case for nested case-control studies where a case-control study is nested within a cohort study, therefore exposures are measured prospectively. Cohort studies identify patients who meet eligibility criteria related to the exposure, for example being prescribed a specific treatment of interest or diagnosed with a

specific disease, and they are followed up for outcomes [1]. Cohort studies may be prospective, where people with particular exposures are followed up over time. These types of studies have the advantage of not relying on recall. However, they can take a long time to complete as follow-up can be many years, resulting in high costs. Cohort studies can also be retrospective, where historical data is used and people meeting eligibility criteria in the past are identified and then followed up to the present time. This reduces the length of time to complete the study and associated costs.

1.3.1 Pharmacoepidemiology

This thesis will mainly focus on pharmacoepidemiology. This is defined as “the study of the use and effects of drugs used in large numbers of people” [2]. Studies within this field are focused specifically on medications, for example they may be interested in how frequently specific medications are used, or the side effects, or long-term effects, of different medications. Cohort studies and registries are particularly useful in this context, however it can be difficult to recruit large numbers of people to prospective studies, particularly for rare diseases.

1.3.2 Where does digital data fit?

Various forms of digital data can be used in epidemiology. They include digital records of our interactions with the health services, data from digital devices where we track our lifestyle, and social media and online forums where we discuss our health and concerns. Data from these sources may already be collected and then repurposed for research, or the data source may enable specific data to be collected for a study in a more efficient manner. I will focus on the data sources used in the studies in this thesis, firstly digital data already collected in the form of databases of routinely collected health data and then using a digital data source to enable specific data collection.

1.4 Databases of routinely collected health data

Databases of routinely collected health data fall into two main types: 1) insurance claims and administrative data, and 2) databases derived from EHR data. Claims data refers to insurance claims databases such as Medicare [3]. These contain details of interactions with the health services and details of therapies, including the prescribed medications, for the purposes of billing. This type of data comes from countries, such as the United States, where the health system is funded through insurance. Administrative data refers to data recording interactions with public services, for examples educational attainment in schools and notifications of births and deaths. EHRs are routinely collected data recording patient's interactions with health services. These usually replace paper medical records and may contain information from consultations, such as symptoms and diagnoses, results of tests and prescriptions of medications. Research databases are available that are derived from this data. These types of database can be linked to administrative databases, such as notifications of deaths, to enhance the data.

The next sections will focus on research databases derived from EHRs in the UK. I will first set the scene by describing the healthcare service in the UK.

1.4.1 The healthcare system in the UK

The UK has a government funded National Health Service (NHS) which covers all healthcare and can be considered a "cradle to grave" service. It is structured with a primary care service provided by general practitioners (GPs), who act as gatekeepers for the service and are the first point of contact. They provide general medical care and make referrals to secondary care. Secondary care provides specialist care, usually in hospitals in either an outpatient or inpatient setting. If a condition is very complex or rare, a patient may be referred from secondary care to tertiary care. Tertiary care usually takes place in specialist hospitals.

1.4.2 Research databases derived from electronic health records in the UK

Research databases in the UK are usually derived from primary care EHR data, often with linkage to secondary care and administrative data and are pseudo-anonymised. There are different databases dependent on the software used by the practice, at the time of writing this thesis the biggest research database is the Clinical Practice Research Datalink (CPRD). There are two parts to this database, CPRD GOLD contains data derived from EHRs, for GP practices using specific software called Vision [4]. CPRD Aurum is a newer database that contains data derived from EHR for practices using a software package called EMIS web [5]. These both contain data on all interactions with the GP including family history, diagnoses, symptoms, prescriptions, immunisations and referrals. Other UK research databases based on primary care EHRs include THIN and QRESEARCH, which are very similar to CPRD, where the databases are based on data GP practices using EMIS and Vision software (THIN) [6] and EMIS alone (QRESEARCH) [7]. The Phoenix Partnership (TPP) are another UK primary care data provider, where data comes from GP practices using Systmone software [8]. TPP has the advantage of linkage to UK Biobank, a prospective cohort study of over 500,000 participants, that contains a wide range of data, often not available within primary care records [9]. For each of the databases, after study approval, the researcher downloads a subset of the data, based on their study population, for analysis within their own data environment. Another UK database that has been recently been developed in collaboration with TPP is OpenSAFELY [10]. This uses a different model, where the data remains within the secure OpenSAFELY data centre and is analysed there. This reduces the risks associated with data leaving the secure data environment and allows analysis to be performed on up to date records in a timely manner. The database was developed during the COVID-19 pandemic to allow rapid analysis of data.

1.4.3 Data recording

Most clinical details, such as diagnoses and symptoms are recorded as Read codes. Read codes are a coding system used in primary care, developed by Dr James Read in the

1980's. These codes enable detailed coding of clinical observations and symptoms, and other aspects such as occupation, lifestyle indicators such as smoking, social circumstances and ethnicity. When certain codes are entered into the system that indicate a measurement has been taken, for example blood pressure, the user will be prompted to enter details of the measurement. There is also the facility to make detailed notes in the form of free text, however this is not available within research databases due to confidentiality, as the text may contain, for example, the patient's name. Prescribed medications are recorded using a Gemscript Drug Dictionary and described as product codes, where generic and branded products at different doses can be prescribed. De-identified data from the EHR are collected and processed, forming the research database.

1.4.3 Linking databases

These research databases can be linked to a variety of other datasets, for more comprehensive information that may not be available in primary care records. Commonly seen linkages are linkage to various Hospital Episode Statistics (HES) data providing information about secondary care such as hospital admissions and outpatient appointments [11]. Within HES, datasets diagnoses are recorded as ICD-10 codes, an internationally recognised disease classification system. Linkage to Office for National Statistics (ONS) death registration data provides more comprehensive mortality data than is otherwise available, including date and cause of death. The linkages available are dependent on the research database provider.

1.5 Using a digital data source for data collection

Although research databases contain large amounts of data, there are questions that cannot be answered using this data source. To study aspects of healthcare that are not captured in health records, such as patient perspectives, it is possible to identify patients that appear in a research database, and ask them to answer a bespoke questionnaire. However, as researchers do not have direct access to patients, the study must be run through the database provider and relies on practices agreeing to take part and uptake has been lower

than expected in previous studies [12]. Alternatively, researchers could go to a hospital or GP clinics to directly recruit patients, or ask clinicians to recruit patients, but both options are again time-consuming and resource-intensive. Furthermore, recruitment from a small number of sites can introduce selection bias resulting in an unrepresentative sample. If the survey needed completing at more than one time point, this would be even more difficult. Online recruitment is one way to have quick, inexpensive access to a large number of patients for surveys that a person can self-complete, and using social media is a way to target a population. Studies have evaluated the use of generic social media such as Facebook and Twitter to recruit participants [13,14] and found them effective at recruiting a large number of patients and inexpensive. A less frequently used resource are online health communities.

Online health communities are websites where patients with the same condition can discuss issues and experiences related to their health condition; this is termed peer-to-peer support [15]. In 2013, it was estimated 1 in 4 people specifically seek health information from other people with the same condition [16]. In 2018, it was estimated that 54% of the UK population looked online for health-related information [17]. These platforms provide quick and inexpensive access to a large number of patients for surveys and recruitment can target people living with a particular disease more easily than generic social media platforms.

1.6 Pharmacoepidemiology using digital data

Studying pharmacoepidemiology involves studying 1) the use and 2) the effects of drugs.

1.6.1 Drug utilisation

When studying drug use, researchers may want to understand drug prescribing across a population, this type of study would fall under the category of a drug utilisation study. To study drug utilisation researchers may be reliant on surveys or audits within a single centre or across a number of centres. These types of studies are time-consuming, they may not give a full picture across the country and it may be difficult to measure trends over time.

Drug utilisation studies using research databases allow measurement of drug utilisation across a country, rather than in a specific centre. As data are usually available over a long period, it allows the measurement of trends over time.

1.6.2 Drug safety

As described earlier clinical trials are not usually long enough to identify all side effects of medications. Prospective studies such as registries can provide insights into drug safety, however they can be time-consuming and expensive. Research databases are well suited to drug safety studies as they contain prescription information as well as diagnoses and symptoms information, with dates, allowing identification of new diagnoses after the commencement of treatment. Alongside the disease group and outcome of interest, the data contains information on potential confounders, for example age, gender, body mass index (BMI), smoking status and other diseases or treatments. Linkage to datasets, such as HES data and ONS death registration data, amongst others, allows the study of more serious outcomes, such as admission for myocardial infarction and death. As these outcomes are relatively rare, the large numbers of patients available for analysis in these databases gives the power to detect differences between two groups that may not be possible to detect otherwise. For example, the association between NSAIDs and myocardial infarction was not clear from RCTs as the event is rare, and people with cardiovascular disease (CVD) were excluded. Observational data provided the power to study this association. One such study used data from QRESEARCH. A nested case-control study design was used, where 9,218 first ever myocardial infarction were identified, with an incident rate of 1.71 per 1000 person years. These patients were matched to 86,349 controls. The size of the study allowed the authors to compare different types of NSAIDs and provide evidence that in both COX-2 inhibitors and non-selective NSAIDs were associated with increased risk of MI. For example, ibuprofen was associated with a 24% increased odds of MI (odds ratio (1.24 (95% confidence interval 1.11 to 1.39) [18]. Without a database of this size it would have been very difficult to observe this: given the low incidence rate (1.17 per 1000 person-years) and

assuming the odds ratio is equivalent to the hazard ratio a sample size of 24.5 million would be needed to detect an effect of this magnitude.

1.7 Methodological challenges when using digital data

Although there are many benefits to using digital data within pharmacoepidemiology, there are some methodological challenges. When using EHR data, there are a number of steps when preparing the data: identifying the population, identifying exposures, outcomes and covariates and defining these in a suitable way for analysis. I will summarise some of the main challenges related to this data preparation and then challenges related to bias and confounding. I will then summarise the challenges related to data collection using digital data.

1.7.1 EHR challenges - data preparation – defining disease diagnosis

As described in section 1.4.3 diseases are recorded using Read codes. GP's can record different levels of detail about the diagnosis, therefore there is not just one Read code for a disease. To identify people with a particular disease, a code list is developed to include all possible codes. To avoid misclassification, the code list should be validated, either externally or internally. External validation would be either confirmation of the diagnosis by a GP, or a record of the diagnosis in linked HES data or linked national registry data. Internal validation would be other indicators of the disease such as confirmatory tests, medications prescribed or symptoms recorded in the EHR [19]. This is a time-consuming process, but for many diseases it has been completed already [20–22]. For some populations and diagnoses there may not be an algorithm or validation may not be possible. The researcher will need to consider whether an un-validated code list will suffice for the purposes of analysis.

1.7.2 EHR challenges - data preparation – defining drug exposures

As described in section 1.4.3, product codes are used to record medications prescribed. There are different product codes for different doses and brands of the same drug, therefore, code lists need to be generated to ensure all drug exposure is identified. Unlike disease

codes, codes for medications are only in one section of the record and are essential for dispensing, therefore are more reliable than disease codes. However, the data does not provide exact durations of the prescriptions, therefore data preparation is required to estimate stop dates of drug exposure. This will be discussed further in section 3.5.1.

1.7.3 EHR – missing data

Lifestyle factors are often important for studies, however these are often not systematically collected for all patients. For example, body weight or smoking information is usually collected opportunistically, leading to missing data. If not accounted for appropriately, missing data can lead to biased results, if those without missing data do not represent those with missing data. For variables with missing data, a researcher will need a strategy dependent on the nature of the variable, i.e. exposure, confounder or outcome, and the amount of missing data. The type of missingness will need to be determined: missing completely at random (missing for reasons unrelated to anything observed or unobserved), missing at random (missingness is related to observed data) or missing not at random (missingness is related to unobserved data) [23]. The decision on the type of missing data is based on judgement. If data are considered to be missing at random then multiple imputation can be used [24]. Multiple imputation involves predicting the missing value based on available data, including all variables to be included in the analysis and other variables that help predict missingness. If the amount of missing data is small or there are not many extra variables that provide information about missingness, or if the data is not missing at random then complete case analysis may be preferable [25].

1.7.4 EHR challenges – bias and confounding

Observational research challenges commonly relate to potential bias and confounding, and need consideration in the context of research databases [26].

1.7.4.1 Bias

Bias is described as “any systematic error in the design, conduct or analysis of a study that results in a mistaken estimate of an exposure’s effect on the risk of disease” [1]. As described in section 1.4.1, when using EHR data code lists and algorithms are used to define a disease population, these need to be comprehensive to capture the disease population fully. Systematic differences in those selected to take part in the study may result in bias. If the study population is defined further by having a specific characteristic measured, e.g. body weight, there may be selection bias if those who are heavier are more likely to have their weight measured and therefore are more likely to be included in the study. Other forms of bias may arise because data are not captured systematically. For example, surveillance bias occurs when the exposed group has more of an opportunity for the outcome to be measured due to greater surveillance compared to the unexposed group [27] and may be more of an issue when using EHR data, this will be discussed further in section 3.8.1. Further examples will be described throughout the thesis.

1.7.4.2 Confounding

A confounder is a variable that is associated with the exposure, a risk factor for the outcome and not on the causal pathway. If a confounder is not accounted for, it can result in a distorted association, where either a true association is masked or an association is seen where one does not exist [1]. A particular concern when using EHR data is confounding by indication. This occurs when people who are treated with a specific drug may differ from those who are not treated, due to other characteristics that influence a clinician’s decision to treat [28]. For example, a person with severe disease is more likely to receive therapies that are more intensive compared to those with less severe disease. This can result in an imbalance in baseline covariates and biased results if disease severity is associated with the outcome. This can be taken into account in the design or in the analysis, through adjustment, and will be discussed further in section 3.5.2.

1.7.5 Challenges in data collection using digital data

One big challenge of data collection through an online platform is recruitment: the people who use this platform, and agree to take part in studies, may not be representative of the underlying patient population, this will be discussed further in chapter 5. Another challenge is that the data relies on self-report, resulting in possible recall bias. This could be differential if recall is different between cases and controls, for example, cases with lung cancer may recall their smoking history in more detail than controls who do not have lung cancer. It may be that there is recall error and this is non-differential which will bias towards the null.

Having described the use of digital data in epidemiology, I will now give an overview of the subject area of this thesis - RA and its treatments.

1.8 Rheumatoid arthritis - overview

RA is a chronic autoimmune disease characterised by inflammation in the joints. This inflammation causes pain, stiffness and swelling in the joints and can lead to irreversible damage [29,30].

1.8.1 Diagnosis

In the UK, the National Institute for Health and Care Excellence (NICE) guidelines [31] recommend that people who attend primary care with symptoms suggestive of inflammatory arthritis are referred to a rheumatologist in secondary care, for further assessment and to determine if the patient has RA. Although classification criteria exist for the purposes of research [32], diagnosis of RA is based on clinical signs and symptoms, blood tests and imaging. Severity of disease is measured using composite scores such as the 28 joint disease activity score (DAS28) [33].

1.8.2 Prognosis

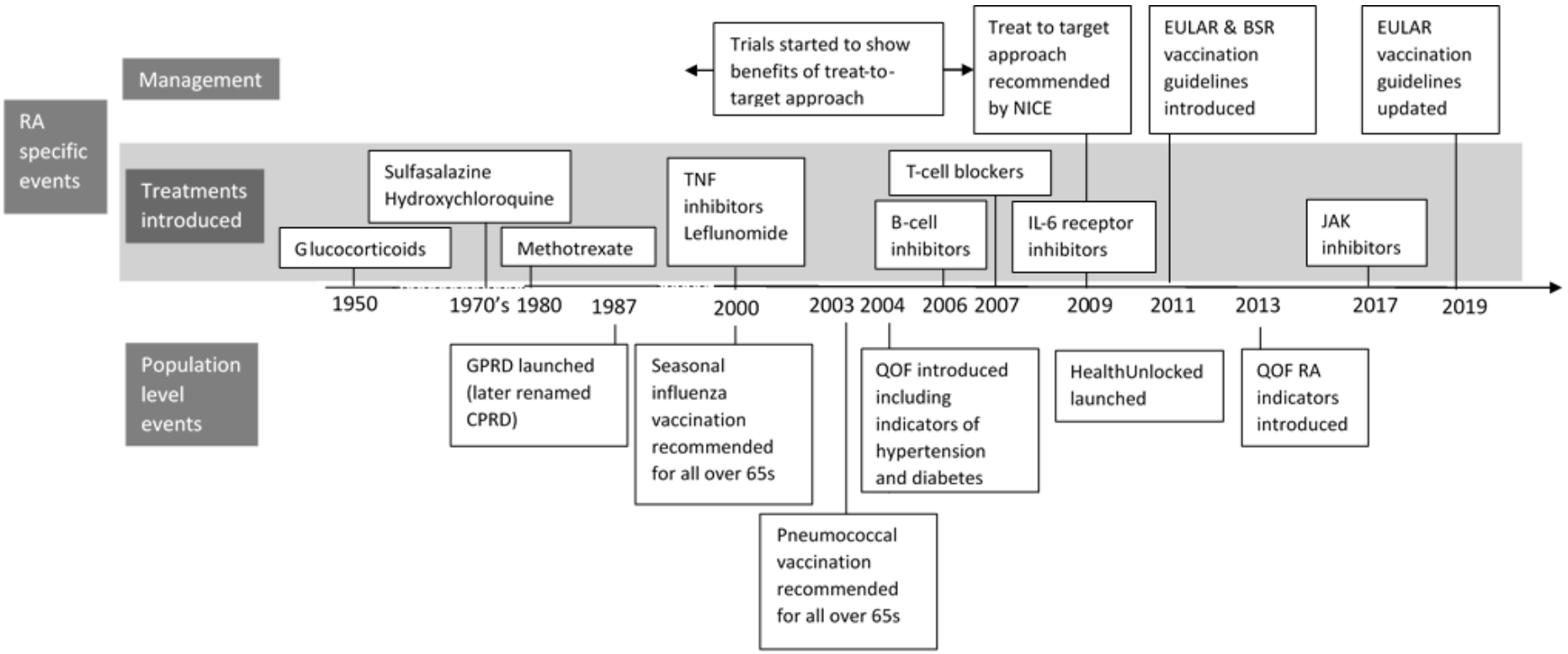
In the past, severe disease would result in major bone deformity and immobility, particularly in the hands, and this would result in disability. Nowadays patients are treated with disease-modifying anti-rheumatic drugs (DMARDs) (Figure 1), and severe deformity is seen less

often, though disability is still present [34–36]. RA is a systemic disease, this means that not only the joints are affected. Extra-articular features can affect the cutaneous, respiratory and cardiovascular systems therefore the frequency of comorbidities is high [37], with the most frequent being CVD, cancers, infection and osteoporosis [37–43]. Mortality rates are also increased in patients with RA are increased compared to the general population [44]. CVD is the most frequent cause of death in patients with RA [45,46].

1.8.2 Pharmacological therapies

RA is primarily treated with DMARDs, a class of drug that interfere with the disease process thereby reducing joint pain and inflammation, reducing consequent damage and improving long-term outcomes of the disease [47]. There are now several classes of DMARDs: conventional synthetic DMARDs (csDMARDs), biological DMARDs (bDMARDs) and targeted synthetic DMARDs (tsDMARDs), and prescribing is guided by NICE, European League against Rheumatism (EULAR) and British Society for Rheumatology (BSR) guidelines [31,48,49]. csDMARDs comprise of drugs including methotrexate, sulfasalazine, leflunomide and hydroxychloroquine, prescribed alone or in combination depending on disease severity. These are prescribed initially in secondary care, with ongoing prescribing in primary care once treatment is stable. bDMARDs and tsDMARDs were developed to target specific immune pathways. bDMARDs are large proteins, whereas tsDMARDs are smaller synthetic compounds. NICE guidelines indicate that bDMARDs and tsDMARDs are only used in severe disease, if csDMARDs have not bought the disease under control. They are exclusively prescribed in secondary care, usually in combination with methotrexate, unless methotrexate is not tolerated. For further symptom control, non-steroidal anti-inflammatory drugs (NSAIDs) may be recommended, depending on a patient's age and risk of side effects associated with NSAIDs. Glucocorticoid (GC) treatment may also be needed, both at the start of treatment, and to control disease after treatment has started. GCs will be discussed in more detail in chapters 3 and 4. GCs and NSAIDs are prescribed in primary care.

Figure 1: Timeline of events in history of rheumatoid arthritis management



1.8.3 Non-pharmacological therapies and preventative interventions

In addition to these pharmacological therapies, non-pharmacological therapies may also be of benefit to patients, such as physiotherapy, occupational therapy and exercise programmes. In addition, there are preventative interventions and assessments to prevent common comorbidities. It is recommended that patients with RA should have annual reviews where comorbidities, such as CVD and osteoporosis, are monitored, allowing early identification and treatment [31]. These take place in primary care and form part of the quality and outcomes framework (QOF), a programme to incentivise GPs described in section 2.1 [50]. It is also recommended that patients are vaccinated to reduce the chance of infections. Vaccinations take place in primary care and will be discussed further in chapter 2.

1.9 Chapter conclusions and thesis aims

This chapter has introduced digital data, focusing on research databases derived from EHR's and using digital data as a source for data collection, as well as some of the resultant methodological challenges. This thesis will show how digital data can be used to describe treatment uptake and understand outcomes in relation to treatments for people living with RA. The thesis will have a focus on the type of data used and the methodological considerations while using these types of data. My research has concentrated on two areas of RA management: vaccinations and GCs. The following chapters will provide more details of these two areas and my work. The specific objectives will be given throughout the chapters.

Chapter 2: Vaccinations

Vaccinations are an important part of the management of RA, due to a predisposition to infection [41,42,51]. The risk of infections is increased in patients with RA due to the disease itself [41] and the immunosuppressive drugs these patients are prescribed to treat the condition, including GCs, csDMARDs and bDMARDs [52,53]. Further to this, as mentioned in section 1.8.2, patients with RA have an increased risk of mortality, and serious infections are one of the leading causes of death [54,55]. Vaccinations are one way to reduce the risk of common community-acquired infections [56,57] , and influenza and pneumococcal vaccinations are recommended for patients with RA [58,59].

2.1 Vaccinations in rheumatoid arthritis

Both EULAR and BSR have recommended influenza and pneumococcal vaccinations in patients with autoimmune inflammatory rheumatic diseases, including RA, since 2011 [49,60–62]. National guidelines in the UK recommend influenza and pneumococcal vaccinations for all people ≥65 years and for patients in a clinical risk group, such as patients with RA taking immunosuppressive medications [58,59]. Influenza vaccinations in those aged ≥65 years and in clinical risk groups are incentivised in primary care through the Quality and Outcomes Framework (QOF) [63]. This is an annual reward and incentive programme where GP practices achieve points by hitting specific targets, and are paid based on the points they achieve and what they aspired to achieve. Pneumococcal vaccinations in the same groups are incentivised through the enhanced services programme where practices are paid based on the number of vaccinations recorded. Vaccinations eligible for incentives are identified through the Read codes.

Understanding whether these recommendations are being followed is important from a public health perspective. If vaccinations are not taking place, there is the opportunity to highlight this and encourage vaccinations. The UK national figures of vaccination uptake describe uptake for patients who are immunosuppressed, but not disease groups, such as

RA, within this category [64]. On reviewing the literature, studies describing pneumococcal and/or influenza vaccination uptake in the UK in patients with RA did not provide a clear picture of vaccination status of patient with RA across the UK. Studies were limited to small, single centre, cross-sectional studies (N<200) therefore may not be representative of the whole UK. The studies also relied on self-reported vaccination status means there is the potential for selection bias and recall error in these studies [65–70]. Sowdon et al did additionally reviewed hospital records and information from a local immunisation centre [69] and reported the lowest uptake: 53% had received an influenza vaccination and 28% pneumococcal vaccination, compared to 77-59% and 40-43% for influenza and pneumococcal vaccinations, respectively, in studies relying on recall only [65–68,70]. This is consistent with other studies that found self-reported vaccination status tended towards false positives in the unvaccinated [71].

More recently, international guidelines have specified that the timing of vaccination in relation to starting immunosuppression may be important and made recommendations related to this. An RCT in 2018 found that a temporary 2 week discontinuation of methotrexate after influenza vaccination significantly improved immunogenicity, [72] suggesting timing of vaccination may be important in relation to methotrexate. Meta-analyses have also found that methotrexate and bDMARDs in particular reduced the immunogenicity of both influenza and pneumococcal vaccinations [73,74]. EULAR guidance only recommended vaccination prior to B-cell depleting therapy in the past, but the 2019 update did expand this to all DMARD types [62]. The BSR guidelines have recommended vaccination prior to immunosuppression since 2011 [49,61]. Only one previous study, conducted at a single centre in the US, has described timing of vaccinations [75]. This study used EHR and claims data and found 37% patients with RA, received a pneumococcal vaccination prior to starting immunosuppressive medication, but only represents people in the US.

The first objective of this thesis was:

- To describe vaccination uptake in RA and its timing in relation to starting disease modifying anti-rheumatic drugs using data from primary care EHRs.

I conducted two studies to investigate this objective. The first examined vaccine uptake, while the second looked at the timing of its administration with respect to DMARD use. The studies used data from CPRD GOLD, as described in section 1.4.2, (referred to as CPRD throughout the rest of thesis, as Aurum did not exist at the time of the studies described).

2.2 Publication 1: Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink

2.2.1 Methodology

Population: This study used a cohort of patients with incident RA identified during the study period 1st January 2000 to 31st December 2013, followed from RA diagnosis until death, leaving their practice, when their practice stopped contributing to CPRD or the end of the study period. People with RA were identified using a previously validated algorithm [76]. The algorithm uses multiple indicators for RA, with either >1 Read code for RA, or an RA code and DMARD medication codes and no alternative indications for DMARDs in the previous 5 years. External validation through GP confirmation of diagnosis showed these criteria had sensitivity and specificity of >80%. The algorithm was slightly updated in 2015 [77] but it was after the majority of the studies described in this thesis were started therefore the older algorithm is used throughout this thesis.

Exposure: Vaccinations during follow-up were identified using Read and product codes. As influenza is an annual vaccine, vaccination status each year was determined, with the year start corresponding with the influenza season (1st September) and vaccinations until 31st March identified. For pneumococcal vaccination, only the first vaccination was counted. For

both types of vaccination, it was determined whether the first vaccination was prior to starting csDMARDs.

Analyses: For both vaccination types, the proportion with at least one vaccination and the proportion vaccinated prior to starting csDMARDs was tabulated. To understand if influenza vaccinations were taking place annually, the proportion of vaccinations received was determined, with the denominator being the expected number of vaccinations to the number of vaccinations based on the years of follow-up available. All analyses were stratified by age (<65 years compared to ≥65 years).

2.2.2 Results

This study found that vaccination uptake was suboptimal for both influenza and pneumococcal vaccination. Although 80% of patients received at least one influenza vaccination, in the first 5 years after RA diagnosis, up to two-thirds were not vaccinated annually. Only 50% patients received their pneumococcal vaccination. Those aged under 65 years were less frequently vaccinated. Overall, just under half (49%) of patients received an influenza vaccination prior to starting csDMARDs and 42% received a pneumococcal vaccination prior to starting csDMARDs.

2.2.3 Contribution to the literature and impact

This was the first study to report on influenza and pneumococcal vaccination in patients with RA in the UK as a whole, using nationally representative data. The results show that vaccinations are not taking place as recommended, particularly for those under 65 years. This suggests that people may be vaccinated due to meeting the age threshold, where there are incentives to vaccinate, rather than because of their RA. The incentive programme for pneumococcal vaccination provides fewer incentives for practices to improve compared to the QOF, which may explain the lower rates of pneumococcal vaccination. Therefore, clinicians should be particular vigilant to ensure those under 65 years are vaccinated.

The yearly uptake of influenza vaccination was lower than in previous studies that relied on self-report [65–68,70], and most similar to Sowden et al where medical records were reviewed [69]. This provides further evidence that self-reported vaccination status overestimates the proportion vaccinated, and shows the suitability of EHR data for studying vaccination uptake.

Overall, the results are important from a public health perspective, as they highlight that patients with RA are missing their vaccinations, and rheumatologists and GPs should be encouraged to offer vaccinations to reduce the chance of infection. This publication has been cited 26 times (as of April 2021). This includes a citation in the latest EULAR recommendations in adult patients with autoimmune inflammatory rheumatic diseases, where this work provided part of the evidence that vaccination uptake was suboptimal [62]. An abstract of this work was selected for oral presentation at the 2015 EULAR annual conference and for e-poster presentation at 2015 BSR annual conference.

2.3 Publication 2: Pneumonia vaccination timing in relation to starting conventional synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis.

Since the previous study in 2014, EULAR recommendations have been updated to recommend vaccination should be prior to all immunosuppression, rather than just b-cell depleting therapy [62]. Following these guideline updates, I wanted to explore the timing of vaccination in relation to starting csDMARDs in more detail. Changes over time may indicate the recommendations have changed practice and measuring how close vaccinations were to starting csDMARDs may indicate that vaccination was taking place due to starting immunosuppressive medication rather than another indication, such as age. I chose to focus on pneumococcal vaccinations for a two reasons. Firstly methotrexate in particular has been shown to reduce the immunogenicity of pneumococcal vaccinations [74] and secondly

influenza vaccination is annual, so it would not be possible to identify a particular temporal association with starting csDMARDs. There is NICE guidance that pneumococcal vaccination should be “repeated at 10-yearly intervals if given before starting the DMARD, or at 5-yearly intervals if given after starting the DMARD” [78] and this was taken into account when designing the study.

2.3.1 Methodology

Population: This study used the same data as publication 1, though the study period was extended to 31st December 2018 and there were some slight changes to the cohort definition. As starting csDMARDs was when follow-up started and therefore central to this study, there was the additional requirement of csDMARDs being prescribed up to a maximum of 3 months prior to, or any time after, the first RA code.

Exposure: To be sure that the pneumococcal vaccination was up to date, only pneumococcal vaccinations up to 5 years prior to starting csDMARDs and any time after were included.

Analyses: For each patient it was determined if the first vaccination was prior to starting csDMARDs. As with paper 1, I stratified by age. I compared the proportions vaccinated within the first year or within 3 years, prior to or after starting csDMARDs. I determined the proportion vaccinated prior to starting csDMARDs by year csDMARDs were started.

2.3.2 Results

This study showed that 36.5% were vaccinated prior to starting csDMARDs. Those aged 65 years or over were more frequently vaccinated prior to starting csDMARDs. Although the proportion vaccinated prior to starting csDMARDs was low, there were some more positive findings: there was evidence that RA diagnosis was triggering pneumococcal vaccination as the frequency of vaccination was highest around this point, with 22% vaccinated in the year after RA diagnosis. There were increases in the proportion vaccinated prior to starting csDMARDs over time, from 17% in 2000 to 55% in 2016.

2.3.3 Contribution to the literature and impact

Besides paper 1, only one other study has reported on timing of vaccination in relation to csDMARDs. Desai et al found that 37% of patients with RA were vaccinated prior to starting csDMARDs over 2 years (2008-2010), in a single centre in the USA [75], which was similar to these findings. However, the data I used enabled further investigation into when vaccination was occurring and whether there had been changes over time at a national level. This helps show clinicians where improvements can be made, and provides some encouragement that timing was improving over time, perhaps reflecting the increased awareness given the changes in recommendations. Again, the results highlight the disparities by age, indicating that those under 65 years are most frequently unvaccinated.

With this study, though misclassification around the date of vaccination was unlikely, there could be misclassification with the start date of csDMARDs, as csDMARDs may be started in secondary care, this means the proportions vaccinated prior to starting csDMARDs could be an overestimate.

An abstract of this work was presented as a poster at EULAR 2020. Since publication in June 2020, the letter has been downloaded 1491 times (as of April 2021).

2.4 Chapter conclusions

The objective of this chapter was to describe the uptake of influenza and pneumococcal vaccination in RA and its timing in relation to starting csDMARDs. The studies showed that vaccination uptake and timing was suboptimal, particularly in those under 65 years. This is important in this group at risk of infections. The studies highlights the need for clinicians to consider vaccinations, particularly when starting csDMARDs.

Chapter 3: Glucocorticoid associated outcome

As described in section 1.9 the other area of RA management I have been interested in is GCs. In this chapter, I will give a background to GCs as a therapy in RA, before describing the studies I have conducted using CPRD data to answer questions related to the risk of specific GC side effects in patients with RA.

3.1 History of glucocorticoids

The clinical benefits of GCs were discovered in the 1950's by Phillip Hench. He showed their anti-inflammatory properties and efficacy treating RA and other inflammatory conditions [79]. Hench won the Nobel Prize in 1950 for this work. At this time, RA was a debilitating disease with a paucity of treatments. Rheumatologists were invited to see the effects of GCs in patients with RA and reported that "during the course of two days we saw them miraculously improve" [80]. At this time, cortisone was given as intra-muscular injections and the steroid compounds themselves came from animals. Significant side effects, however, were noted from early on (see section 3.3). In the following decades synthetic steroids were developed, the most frequently used being prednisolone or prednisone [81], with various routes of administration: oral tablets, intra-muscular injections or inhalation. As well as treating rheumatic diseases, they have a range of applications in other inflammatory diseases involving all systems of the body. For example, they are frequently used in the treatment of respiratory diseases such as asthma, where typically inhaled GCs are used; skin conditions, such as eczema, where topical GCs are the most common form; and as part of active and palliative treatment of some cancers [81,82]. A recent example is the use of dexamethasone to treat COVID-19 [83,84].

3.2 Prevalence of glucocorticoid use

In 2011, Fardet et al conducted a study of the prevalence of long-term GC therapy over the previous 20 years. They showed that on average 0.75% of the population were prescribed oral GCs for at least 3 months and the most frequent indications were asthma, polymyalgia

rheumatic/GCA and chronic obstructive pulmonary disease [81]. Another study in 2000 had shown similarly that 1% of the UK population was prescribed GCs, with the most frequent indication being respiratory disease (40% patients) and 6% for musculoskeletal disorders [82].

3.3 Glucocorticoid side effects

As GCs are synthetic versions of naturally occurring hormones, they are able to bind to most cell types. This means they have effects throughout the body [85]. The anti-inflammatory effect is of benefit, but there are a number of adverse effects. At the time of discovery, there were few regulations around the introduction of medicines. Today, new medicines go through multiple stages of clinical trials testing safety and efficacy in order to gain approval by regulatory bodies (Medicines and Health Regulatory Authority in the UK) and come to market. When first studied in patients with RA, the doses of cortisone were large at 100mg. There were reports of facial rounding and weight gain with these large doses [86]. Nowadays, GCs are administered at lower doses where possible, for example maintenance therapy in rheumatic disease is often <7.5mg prednisolone equivalent dose (PED) per day [87]. A variety of GC side effects have been reported, from weight gain and insomnia to clinically serious side effects including type 2 diabetes mellitus (DM), osteoporosis, infections and cardiac disorders. However, the evidence is not clear-cut. The summary of product characteristics for prednisolone, for example, says the frequency for all undesirable effects is unknown [88]. Similarly EULAR recommendations for medium to high dose GCs state there is a paucity of data on “wanted and unwanted clinical effects of GCs”. This makes it difficult for clinicians and patients to weigh up the benefits and risks of GCs. Indeed, there are opposing views on how much GCs should be used by clinicians due to the side effects [89].

An understanding of the side effects is important particularly in patients with RA, given their frequent comorbidities and increased risk of mortality (section 1.8.2). Communicating risks and benefits to patients is particularly important in the course of shared decision making where clinicians and patients decide together the path of treatment [48,90,91], and is

described as comprising of four dimensions (nature of benefits/harms, probability they will occur, importance to the individual and maximising benefits/minimising harms) [92]. The work on GCs that I will describe relate to two of these dimensions, the first refers to the probability of side effects, based on the available evidence (this chapter) and the second refers to the importance of side effects to individuals (chapter 4).

3.4 Chapter objective

The second objective of this thesis, related to the probability of side effects is:

- To investigate the probability of GC related outcomes in RA: all-cause and cardiovascular mortality and hypertension using data derived from EHRs.

There are three studies related to this objective, the first describes the association between GC use and mortality, the second describes how comorbid type 2 DM affects the association between GC use and mortality and the third describes the association between GC use and hypertension. CPRD data is well placed to address these questions. In particular this is because oral GCs are typically prescribed in primary care, therefore are well captured in CPRD, allowing the exposure to be modelled in a time-varying manner. Further, CPRD can be linked to ONS death registration data (as I will go onto describe) to provide robust mortality data, and both hypertension and DM are routinely assessed and managed in primary care, so are also well-captured.

3.5 Methodological points for this chapter

Before describing the specific methodology for the publications, I will describe medication data preparation and confounding by indication more broadly, which applies to all three papers in this chapter.

3.5.1 Medication data preparation

Before analysis, medication data in CPRD requires preparation as the raw dataset only contains the start date of the prescription, and a number of different variables that can help

estimate the prescription end date: 1) quantity and numeric daily dose, 2) number of days of the prescription and 3) dose duration. Further, within the dataset, 2 and 3 are frequently missing. In pharmacoepidemiology studies, this step of data preparation is often not well described. Different researchers are likely to prepare the data slightly differently, which has been shown to impact results and result in misclassification [93,94]. An algorithm, known as DrugPrep [94], has been created to improve both the efficiency and transparency of drug exposure data preparation. It contains a sequence of steps for cleaning the data, deciding which data points to use to determine the length of the prescription, handling multiple and overlapping prescriptions and handling gaps between prescriptions. The transparency of data preparation has gained greater importance more recently through its inclusion in the reporting guidelines for pharmacoepidemiology studies using routinely collected data [95]. I will describe where this algorithm has been applied for each publication.

3.5.2 Confounding by indication

For studies throughout this chapter, a potential confounder by indication is disease severity, as people with high disease activity may be more likely to be prescribed GCs and this high disease activity may make patients more likely to have the outcomes of interest (mortality, type II DM and hypertension). This means that GCs may appear to be associated with worse outcomes but actually, it is disease activity that is associated with these poor outcomes. In CPRD, measures of disease activity are not routinely captured, as this is usually monitored in secondary care rather than primary care. However, disease activity can be inferred through proxy measures. Ward et al showed that frequency of rheumatology visits did correlate with functional disability [96]. In the following studies, I have used measures of healthcare utilisation: average number of rheumatology outpatient visits in those with linkage to HES outpatient data and the mean number of GP visits per year. This allowed me to address this problem, although residual confounding remains possible.

3.6 Publication 3: Oral glucocorticoid therapy and all-cause and cause-specific mortality in patients with rheumatoid arthritis: a retrospective cohort study

The association between GC use and mortality is an important question, as described in section 3.3 GCs have a range of side effects, some of these (such as CVD) are serious and could be potentially fatal. Therefore, it is conceivable that GCs could increase the risk of mortality. However, in RA, it is also possible that GCs may reduce mortality risk by lowering inflammation and disease activity. Previous studies examining GC use and mortality in patients with RA produced conflicting results. Many found that GC use was associated with increased mortality [45,46,97–100], but not all [101–103]. These studies had a variety of definitions of GC use, often summarising GC use, for example summarising dose category over the previous 12 months [98]. This risks not fully capturing the inconsistent nature of GC prescribing, which changes in response to disease activity. Further, some studies had small numbers and many were based on data from a limited number of centres [45,97,99,100,103].

ONS linkage was critical to this study. Death registration is a legal requirement in England and Wales, and therefore provides comprehensive coverage. Although date of death is available in CPRD data, an algorithm is used to derive this date, thus is considered less accurate [104]. In addition, ONS data includes cause of death (unavailable in CPRD), allowing investigation of cause-specific mortality.

3.6.1 Methods

Study population: A cohort of patients with RA and linkage to ONS death registration data during the study period 1st January 1998 to 1st October 2011 was identified. Patients entered the study at the latest of RA diagnosis date, date of ONS linkage or 1st January 1998, so were a mixture of prevalent and incident cases.

Exposures and outcomes: Date of death was identified from the linked data only. This study was conducted prior to the completion of the DrugPrep algorithm, therefore GC data

preparation did not follow all of the decisions of the algorithm, though they were similar. The steps used are described in the manuscript appendix. GC use was defined in six different ways: ever GC use, current GC use, current GC dose, current GC dose category, cumulative GC dose and cumulative GC dose category. This allowed an understanding of the impact of dose, and provided more certainty that any association seen was a true association, if seen across different definitions of GC use.

Analyses: In the main analyses, Cox proportional hazards (PH) regression models measured the association between the six definitions of GC exposure and all-cause and cause-specific mortality. The proportion of missing data for each variable was assessed. Variables with less than 5% proportion of missing data (smoking and deprivation) were included in the final model in a complete case analysis. BMI was the only variable with more than 5% missing data. The analysis plan *a priori* was to impute such a variable if it was associated with the outcome or altered the hazard ratio by more than 10%. However, this was not the case for BMI and it was therefore excluded from the final model.

Studying mortality associated with medication use can be difficult, as researchers need to disentangle mortality associated with medication use and medication use due to worse disease that results in increased mortality (i.e. not due to the medication itself). This is a form of protopathic bias – where an exposure is thought to be the associated with an outcome, when in fact the exposure is due to early signs or symptoms of the outcome. In this study this was termed “peri-mortal bias” where being in the later stages of life influenced GC prescribing, which in turn affects the association seen between GC use and mortality. This was examined in three ways: firstly, the overall and cause-specific mortality rate in the first six months after GC initiation was compared to the mortality rate more than six months after GC initiation. Secondly, making the same comparison but using the proportion of deaths as opposed to mortality rate. Both these measures provided an indication of GC prescribing in response to a terminal illness. The third approach was to use a lag-time window and excluding GC use in the six months prior to deaths [105]. Using these three approaches

allowed a greater understanding, and greater certainty, of whether peri-mortal bias was a problem.

3.6.2 Results

This study found that all definitions of GC use were associated with an increased risk of mortality. When dose was categorised, no association was seen at the lowest GC dose (<7.5mg) however above this, risk increased with increasing dose. The peri-mortal bias analyses indicated its presence, however it did not completely explain the association seen. As GC use was associated with death from other causes, unmeasured confounding was explored using the rule out approach [106], where the amount of confounding required to completely explain the confounding is determined. For this study, unmeasured confounding would need to have a 40% prevalence, increase the relative risk of mortality by a factor of 3 and increase the odds of GC exposure by 3.5 to completely explain the findings, which seems unlikely.

3.6.3 Contribution to the literature and impact

The findings of this study fit with previous literature [45,46,99,100], in particular with two studies where only doses over 5mg were associated with increased risk of all-cause mortality [98] and cardiovascular mortality [97]. Specifically, this study adds to the previous literature through use of a large national study population, and multiple models of time-varying GC exposure, providing more confidence in the results, as well as investigating “peri-mortal” bias and potential unmeasured confounding, which were not addressed in the previous studies. Clinically the implications remain that GCs should be used at the lowest effective dose to reduce the risk of mortality. It has been cited 25 times, and was referenced in a systemic review informing the 2019 update of EULAR recommendations for the management of RA [107].

3.7 Publication 4: The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study.

This study explores whether comorbid type 2 DM affected the association between GC use and mortality (as seen in publication 3). GC use is associated with an increased risk of DM [108,109] and DM itself is also associated with an increased risk of CVD and mortality [110][111]. Given the increased risk of CVD and mortality in RA [44][112–114], understanding possible interaction between GC use and DM, may be particularly important for patients with RA.

3.7.1 Methodology

Study population: In this study, I replicated the study period in publication 3, but included only patients with incident RA. Prevalent cases were not included to avoid potential survivor bias, where prevalent cases who survive to be included in a cohort may be systematically different to those who do not.

Exposures and outcomes: To ensure as complete case capture as possible, patients with DM were identified in three ways: 1) a Read code for type 2 DM, 2) at least 2 prescriptions for anti-diabetic medication, or 3) fasting blood sugar ≥ 7.0 mmol/litre, random glucose test ≥ 11.1 mmol/litre, glucose tolerance test ≥ 11.1mmol/litre or a glycosylated haemoglobin (HbA1C) ≥7%. As patients with polycystic ovary syndrome could be prescribed an anti-diabetic medication without having DM, patients with this diagnosis were excluded. For this study, I used the time-varying definition of current GC use as this most accurately reflected the intermittent nature of prescribing. GCs were prepared using the DrugPrep algorithm. As the end date of prescriptions is an estimate and to allow for potentially long-lasting effects of GCs, a six-month risk attribution window was used. This meant that events were attributed to GC use for six months after the estimated end date of the prescription. This definition will be described as recent GC use.

Analyses: In standard regression models, measures of risk including interaction are on the multiplicative scale, so risk is estimated relative to baseline risk. Measures on the additive scale, such as risk difference, provide an estimate of the additional cases. This is more useful to assess the public health impact, as it is not dependent on a baseline risk. If the baseline risk of mortality in people with DM is higher compared to those without DM, relative risk may not provide a true reflection of mortality associated with GC use in people with DM. To estimate the effect of GC use and comorbid DM on mortality, interaction was measured on both the additive and multiplicative scales. Multiplicative interaction was measured by including an interaction term in the Cox PH model. Additive interaction relates to risk difference and was measured by calculating the relative excess risk due to interaction [115] and the ratio of absolute effects (manuscript appendix). Missing BMI and smoking data were imputed using multiple imputation. For this study disease activity was not expected to confound differentially by DM status, i.e. patients with DM were not expected to have more or less active RA than patients without DM, so additional analysis accounting for this was not performed.

3.7.2 Results

This study found that when measuring the association between GC use and all-cause mortality, patients with DM had a lower risk ratio but higher risk difference compared to those without DM (Table 1). This emphasised the impact of higher baseline mortality in patients with DM and the importance of measuring additive interaction.

Table 1: Association between recent GC use and all-cause mortality stratified by diabetes mellitus status.

Diabetes status	**Risk ratio (95% confidence interval (CI))**	**Risk difference (95% CI)**
Diabetes mellitus	2.99 (2.32 to3.87)	44.9 (32.9 to 56.8)
No diabetes mellitus	4.37 (3.77 to 5.07)	34.4 (30.1 to 38.7)

In the Cox PH model for all-cause mortality, there was no multiplicative interaction seen (0.86 (95% CI: 0.64–1.15)), however additive interaction indicated increased risk of mortality, though it was not statistically significant (adjusted ratio of absolute effects: 1.22 (95% CI: 0.86 to 1.72)). This means that patients with RA and comorbid DM who are prescribed GCs have a 1.22 times increased absolute risk of mortality compared to those with RA but no comorbid DM who are prescribed GCs.

3.7.3 Contribution to the literature and impact

To my knowledge, this is the first study to estimate the mortality risks associated with GC use and comorbid type 2 DM in patients with RA. In the general population, Olivarius et al investigated mortality in those using GCs at DM diagnosis compared to those not using GCs, there were only 35 deaths and increased mortality in the GC group was explained by age [116]. Methodologically this study shows the importance of considering the baseline risks and measuring additive interaction, as recommended by Knol and Vanderweele [115]. Clinically this study highlights the importance of considering comorbid DM when prescribing GCs. This publication has one citation as of April 2021. An abstract of this work was selected for oral presentation at EULAR 2018.

3.8 Publication 5: Glucocorticoid use is associated with an increased risk of hypertension.

Hypertension associated with GC use had not been well studied in patients with RA. As patients with RA are at higher risk of CVD and hypertension can be easily measured and treated, it is important to understand if an association exists. Indeed, a study of patient and rheumatologist views on GC side effects found hypertension was the third most worrisome side effect for rheumatologists, though less so for patients [117]. Only a few studies had previously investigated GC-associated hypertension in patients with RA and have had conflicting results. A positive association was seen in two studies, however Huscher et al

used self-reported data, making it prone to recall bias [118], and the study by Panoulas et al was cross-sectional limiting interpretation over time [119]. Two studies did not find an association, Wilson et al used primary care data, but only used Read codes, therefore may not have captured hypertension completely [120], the other study by Jackson et al had small numbers [121].

There are a number of key advantages in using CPRD for this study. Firstly, blood pressure is frequently measured in primary care, and since 2004 GPs have been incentivised to measure blood pressure in those aged 40 years and over through QOF [50]. Further, blood pressure measurements are directly inputted into the EHR, allowing identification of patients with hypertension in addition to Read codes. Alongside this antihypertensive medication is prescribed in primary care, and can be used to support case definition.

3.8.1 Methodology

Study population: A cohort of patients with incident RA who did not have hypertension at RA diagnosis were identified, the study period was extended from the previous two studies to 1st January 1992 to 31st June 2019.

Exposures and outcomes: GC data were prepared using the DrugPrep algorithm (decisions described in supplemental data 1). There were three definitions of GC exposure: current use, current dose and cumulative dose. For each definition, a 3-month risk attribution window was used. A validated definition of hypertension was used where patients were required to have either two consecutive systolic blood pressure measurements over 140mmHg, two consecutive blood pressure measurements over 90mmHg or a hypertension Read code and therapy with anti-hypertensives [122].

Analyses: Adjusted Cox PH regression models measured the association between GC use and hypertension. Missing data for BMI and smoking was imputed using multiple imputation. A number of sensitivity analyses were conducted: the inclusion of healthcare utilisation indicators as proxies for disease activity, different length attribution windows (window

increased to six months and decreased to one month) and a stricter definition of hypertension.

A concern with this study was potential surveillance bias. People who were prescribed GCs may have had their blood pressure measured more frequently because of the risk of hypertension with GC use, therefore may have more opportunity for hypertension to be identified. To explore this, the frequency of blood pressure measures was compared between people at three levels of GC exposure (no use, intermittent use and continuous use) in the first two years since RA diagnosis. As people with hypertension have their blood pressure measured more frequently, if hypertension diagnosis occurred within these first two years, follow up was censored at diagnosis.

3.8.2 Results

This study found the incidence rate of hypertension was 64.1 per 1000 person-years. Hypertension was most frequently identified through consecutive high blood pressure measurements. Of those patients with consecutive high readings, only 60% were subsequently prescribed antihypertensive medication. The incidence rate of hypertension was higher in those prescribed GCs (GC use: 87.6 vs no GC use: 59.7 per 1000 person years). Current GC use was associated with a 17% increased risk of hypertension after adjustment for confounders. Only GC doses ≥7.5 mg were associated with increased risk of hypertension. In the first 2 years, the frequency of blood pressure measurements did not differ by level of GC exposure, indicating surveillance bias was not a problem.

3.8.3 Contribution to the literature and impact

This study adds to the literature by providing robust estimates of risk, where there was previous uncertainty. These results agree with the two studies that found an association [118,119]. Another study used CPRD data and did not find an association, but only used Read codes to identify hypertension resulting in a lower IR of 23 per 1000 person-years. This misclassification, if across both GC and non-GC users would bias results to the null and

may explain why an association was not seen [120]. I found that 40% of patients' hypertension was untreated, in keeping with another UK study in secondary care that also identified 40% of patients had untreated hypertension [123]. The study indicates clinicians need to be vigilantly monitoring blood pressure in those prescribed GCs, and act on these measurements. This is particularly important patients with RA, who are already at high risk of CVD. An abstract of this work was presented as a poster at BSR 2020 and EULAR 2020.

3.9 Chapter conclusions

The objective of this chapter was to investigate GC related outcomes (all-cause and cardiovascular mortality and hypertension) using data derived from primary care EHRs. All three studies indicated that GC use was associated with increased risk of their respective outcomes. In publications 3 and 5 which investigated dose, low dose GCs (<7.5mg) were not associated with increased risk of mortality and hypertension, supporting current recommendations to treat at the lowest dose possible [48]. The use of data derived from EHRs allowed me to investigate these clinically important outcomes in a large population, addressing issues including peri-mortal bias, surveillance bias, unmeasured confounding and measure interaction on both the multiplicative and additive scales. This meant I was able to confidently provide estimates of risk for each of the outcomes, and population groups, that clinicians can use to inform patients, and that also have wider impact such as informing clinical guidelines and wider public health.

Chapter 4: The importance of GC side effects to patients

This thesis has focused so far on what are considered "clinically important" GC side effects [117], but understanding patient perspectives of side effects is also an important aspect of drug safety [92]. As described in section 1.5, surveying patients identified in CPRD data is not an efficient method of collecting this data because many practices do not agree to take part in these types of study. To address this, I conducted a study collecting data on patient perspectives using an online health community.

4.1 Patient perspectives

GCs have a wide range of adverse effects, and those that are priorities for patients may differ from those considered important to clinicians and health services. Van Der Goes et al compared the most worrisome side effects of GCs of patients with rheumatic disease versus rheumatologists, using focus group discussions. The top 10 most worrisome side effects were ranked from a list of 37, and it was found that the ranking varied between patients and rheumatologists. Further, they identified that the side effects of concern to patients were often those more frequently occurring rather than those most clinically serious [117]. Focus groups provide rich qualitative data, however, the dynamics of group discussion means it is possible that some individuals may not express their true feelings, and the full range of perspectives may not be captured [124]. Surveys are another way to measure patient perspectives, usually as quantitative data. Although this can reduce some of the richness of the information gathered compared to qualitative research, they can be complimentary. Surveys are typically completed individually, and capture each participant's perspective, which can then be combined to understand perspectives at the population level. Importantly this means each person's perspective is not influenced by anyone else. Side effects of GCs has not previously been investigated through surveys, and as discussed in section 1.5 online platform provide an opportunity to conduct a survey over a short period and in a less

resource intense manner than traditional postal or interview surveys. Therefore the objective of this chapter was:

- To describe the side effects of glucocorticoids that are most important to patients as reported by patients using an online health community.

4.2 Publication 6: Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community.

For this study, participants were recruited from HealthUnlocked, which is the largest online health community in the UK, hosting hundreds of groups that relate to specific medical conditions or interventions. For example, there is a National Rheumatoid Arthritis Society (NRAS) group and a couch to 5km group. Within these groups people post questions, answer other people's posts, or just read the posts. HealthUnlocked therefore provided direct access to thousands of patients, with the ability to target a given diagnosis or medication through the groups or identifying posts with titles or tags with words related to the diagnosis or medication of interest. The use of keywords to target the population is relatively novel in the setting of epidemiology, online surveys usually post a link out to members of a patient organisation or online platform [125], or are advertised through social media [126] rather than targeting particular patients directly in this way.

4.2.1 Methodology

Survey: A short survey (see manuscript appendix) was designed that collected information about timing and beliefs about GCs. The survey popped up on the HealthUnlocked website when a user clicked on a post, within any group, with the title word 'steroid' or the tags 'glucocorticoid', 'prednisolone', 'prednisone', 'steroid' or 'dexamethasone'.

Population: Users who were currently taking GCs or had taken them in the last month were eligible for the survey. This was to limit recall bias, as people who had taken GCs some time ago may not remember all side effects equally and therefore their perspectives may not be truly representative.

Analysis: This study focused on the question where respondents rated how important particular side effects were to them on a scale from 1 to 10. The distribution of responses was plotted in histograms for each side effect. Scores were categorised and then stratified by community group to see if there were differences in perspectives between disease groups.

4.2.2 Results

In this study, 604 users completed the survey, from 17999 pop-ups. Most came from the NRAS group (n=229) and the polymyalgia rheumatica and giant cell arteritis UK (PMRGCAUK) group (n=221). When ranked the side effects most important to respondents were weight gain, insomnia and moon face. Weight gain was the top concern for all groups except the PMRGCAUK group where eye disease was of most concern.

4.2.3 Contribution to the literature and impact

These results were broadly similar to the small number of previous studies [117,127,128]. There were some differences compared to the study by Van der Goes et al, where osteoporosis ranked top and weight gain was 5th. The study populations may explain this, as Van der Goes et al comprised of patients with rheumatic diseases only, most frequently RA (61%). In my study osteoporosis ranked higher in the RA population, compared to the overall population, in keeping with Van der Goes et al [117]. This study adds to the literature by providing patient perspectives based on individual rankings rather than focus group combined rankings. It highlights that side effects not frequently studied are important to patients, this means clinicians cannot provide evidence-based information to patients on side effects that are important to them. These results should therefore drive the future research agenda. This study also had methodological impact, as it was a novel study design, in that the survey popped-up based on keywords, and shows how over 600 responses can be collected over only 3 months. An abstract of this work was selected for oral presentation at ACR 2018. The paper has been cited 21 times including more recent patient perspective

studies on GC side effects and a review of GCs and European guidelines for paediatric Crohns disease [129].

4.3 Chapter conclusions

Publication 6 showed that weight gain, insomnia and moon face were the side effects most important to patients. These side effects are less frequently, if ever, studied and highlights areas where research is needed. The use of an online health community to conduct a survey online allowed me to understand patient perspectives easily and relatively quickly, and provided me with data that was not available in CPRD data.

Chapter 5: Representativeness of patients using an online health community

While conducting publication 6 I was aware that it was not clear how representative the survey respondents were of the general UK population. There may be selection bias if, for example, people who use online communities are younger than the disease population as a whole, this has implications for any researchers using such methodology to understand the external validity of their results. Therefore, the objective of this chapter was:

- To understand how representative patients with RA from online health communities are compared to patients with RA from the general population.

At the time, there were no available data on the characteristics of online health communities or survey respondents from those communities that would allow me to better understand this. I was able to use my knowledge of both HealthUnlocked and CPRD data as a comparator to help address this data gap.

5.1 Publication 7: Representativeness of a digitally engaged population and a patient organisation population with rheumatoid arthritis and their willingness to participate in research: a cross-sectional study.

This study was focussed on patients with RA, as this disease has been a topic of the majority of my work so far. To understand how people from different populations may differ, I explored the characteristics of members of a patient organisation (NRAS), as well as people from an online health community and compared them to the well characterised and representative population within CPRD. Although NRAS has been used to recruit patients with RA in research previously [130–132], the representativeness of the group has not been described.

5.1.1 Methods

This was a cross-sectional study comparing three groups of people with RA: 1) those who completed a survey on HealthUnlocked, 2) those who are members of NRAS and 3) those identified in CPRD. For group 1 (HealthUnlocked), I designed a survey (see manuscript appendix) to capture characteristics and willingness to take part in research, with input from the rest of the study team including other epidemiologists and clinical rheumatologists. A combined patient and public involvement group and NRAS then reviewed the survey to ensure it did not burden respondents. Once finalised, it was embedded on HealthUnlocked within all posts in the NRAS community and popped up for completion when a user viewed an NRAS post. For group 2 (NRAS), NRAS provided an anonymised database of current members that contained information about their characteristics. For group 3, people with RA were identified from CPRD using the validated algorithm used in previous studies throughout this thesis. The characteristics of a prevalent cohort of patients with RA were chosen to represent the general RA population. The groups were compared by testing the difference in proportions for each characteristic.

5.1.2 Results

There were differences in the age distributions of both NRAS and HealthUnlocked populations compared to CPRD. NRAS had an over-representation of people aged 55-75 and an under-representation of those aged 75 years and over. HealthUnlocked had a significantly younger population with fewer respondents aged 65 years and over. Both NRAS and HealthUnlocked had an over-representation of females. Disease duration also differed between data sources, HealthUnlocked respondents were more recently diagnosed, whereas NRAS members were more likely to have had the disease for longer with fewer people diagnosed between 2010 and 2016. Respondents from HealthUnlocked were more likely to be from deprived areas compared to CPRD.

5.1.3 Contribution to the literature and impact

Similar to publication 6, this study showed that a reasonably large sample of participants can be recruited over a short space of time though HealthUnlocked, with 615 responses from patients with RA available for this study. The results indicate that researchers need to consider where participants are recruited, as it may affect the study population characteristics. The greater representation of people from more deprived areas was an interesting finding, as this differs from traditional survey respondents who more frequently have higher socioeconomic status [133]. The use of online health community may therefore allow improved representation of this “harder to reach” group, and increase diversity amongst research participation [134–136]. The results have specific implications for the interpretation of publication 6, as they suggest that those study results may not fully reflect the perspectives of older people and those with longer disease duration. This does not mean those results should be disregarded, but when interpreting the results for people with RA, we need to be aware of the groups that survey results may be less applicable to. It may also explain the differences between the publication 6 results and those of van der Goes et al where study participants had a mean disease duration of 14 years [117]. An abstract of this work was selected for poster presentation at EULAR 2018, and the publication has been cited twice.

5.2 Chapter conclusions

This chapter has shown that people with RA from HealthUnlocked are younger and more recently diagnosed compared to the general RA population, and this approach may encourage inclusion of people from more diverse backgrounds than traditional survey methodology. I was able to use various digital data sources to understand representativeness. These findings are important for researchers when considering which study population to use.

Chapter 6: Discussion

6.1 Summary of work presented

The overall aim of this thesis was to show how digital data can be used answer questions around the management of RA. I used CPRD data to show how influenza and pneumococcal vaccination uptake in the UK is suboptimal, and frequently not taking place prior to starting csDMARDs, as recommended. Vaccinations had not been studied at this scale in the UK previously, and would have been difficult to study without routinely collected digital data. I also used CPRD data to study GC drug safety, and showed that GC use is associated with increased risk of mortality, particularly in those with comorbid DM, and that GC use is associated with increased risk of hypertension. For both mortality and hypertension, lower doses were associated with less risk. CPRD data allowed me to answer these previously unanswered questions related to GC drug safety, in a large study population using robust methods. I used an online health community to collect survey data to enable me to describe patient perspectives of GC side effects. I showed that weight gain, insomnia and moon face were most important to patients. I also used an online health community to understand the representativeness of survey respondents with RA. I found that survey respondents from this platform were younger, more recently diagnosed and from more deprived areas compared to the general population, represented by patients with RA identified using CPRD data. This had implications for the interpretation of the previous study, as the results may reflect the perspectives of patients with these characteristics rather than the general RA population. We do not know if the perspectives of those less well represented would be different, and the results are still useful but it is important to have an awareness of this issue. Using an online health community to collect survey data was a novel method. The platform brings together people with specific diagnoses allowing the targeting of specific patient groups, in both studies this resulted in data being collected for over 600 people in a short period.

6.2 Benefits and challenges of the data

There are benefits and challenges common to both data sources. For example, both require no or minimal data collection resources but challenges include representativeness, where those who complete surveys and those who attend GPs, may not truly represent the disease population and issues around exposure measurement quality, as survey data relies on self-reported data and EHR data does not contain the prescription end date. These issues may, or may not be a problem, but need consideration when designing the study. Alongside these, each data source has specific benefits and challenges.

6.2.1 EHR data

The major advantage with EHR data is its size, a vast amount of data is captured, often with more detail than other data sources. For example, all prescribing by GPs is captured and is reliable as the prescribing is required for dispensing by pharmacy. This makes the data particularly suitable for pharmacoepidemiology studies as I have shown. Data are routinely captured, as a record of clinical care, and are not affected by information bias when entered by the GP. However, this routine capture presents different challenges when using the data.

6.2.2 EHR Challenges - misclassification

One aim of data preparation (section 1.7.1) is to correctly attribute diseases, exposures and events to patients, thereby avoiding misclassification. An example of this, used throughout the thesis, is the validated algorithm used to identify patients with RA [76]. This algorithm has sensitivity and specificity of >80%, thereby reducing misclassification, however it should be acknowledged that misclassification cannot be removed completely. For example, using the RA algorithm, a person is considered a case from the first Read code for RA, however it is possible that they have a diagnosis and are prescribed csDMARDs, but the code has not yet made it onto the EHR from a hospital letter. This means there is time prior to the first Read code that could be misclassified, which may, or may not, affect study results. In publication 2, starting csDMARDs was central to the study so the first RA Read code had to be within 3 months of starting csDMARDs. For the other studies, this potential

misclassification was less of a problem. Another example where misclassification was examined was for the outcome of hypertension (publication 5). Though identified using a validated algorithm [122], with multiple indicators of hypertension, sensitivity analyses were conducted with a stricter definition of hypertension using Read codes and anti-hypertensive medication to explore potential misclassification. With this definition, there were fewer cases, though many did meet the blood pressure criteria prior to meeting this strict criteria. The association between GCs and hypertension was still seen with this definition suggesting misclassification was not a major problem and we could be confident in the main results. EHR data will never be perfect, but I was able to understand and identify where there could be potential misclassification and minimise important misclassification.

6.2.3 EHR Challenges – data availability

As certain data is not captured systematically for all patients, this can result in bias from missing data. For example, BMI (a relevant covariate in publications 3, 4 and 5) was frequently missing, as it is typically measured opportunistically. Different methods were used to address this, in publication 3, the impact of the variable was assessed and it was not included in the final model and in publications 4 and 5, BMI was imputed as I judged it to be missing at random. The aim of these different methods were to reduce the impact of this missing data, and I have continued to learn about how best to deal with missing data [25]. After publication 3, I learned that a complete case analysis in those with BMI could be biased if missingness was associated with the outcome. As it is possible those who had BMI measured did so because they were obese, which is associated with mortality, there could be bias. To avoid this potential bias I decided to impute in later publications. In publication 3, BMI would be an unmeasured confounder, but we showed that an unmeasured confounder would need to be highly prevalent and highly associated with the outcome and exposure. This seemed unlikely to be the case for BMI.

Alongside biases due to non-systematic data capture, there are also challenges related to the absence of potentially important information in EHR datasets. Prescription indication is

not provided, this would be useful, particularly when drugs have multiple indications. A specific example of this is in publication 3 where potential peri-mortal bias was examined in different ways, as we were not able to measure it directly. If we could have known what the GCs were prescribed for, people with GC prescriptions for terminal illnesses could have been excluded. Another challenge, caused by not having disease specific information is confounding by indication (section 3.5.2). As RA is managed in secondary care, there are not measures of disease severity in the primary care record, therefore we cannot account for disease severity directly in analyses. I used the proxy of healthcare utilisation, this did not impact on any of the analyses so it is hard to know how well this measured disease severity. I did consider using the number of csDMARDs prescribed, as increased numbers of csDMARDs may indicate higher disease severity. However, there are other reasons for a person switching csDMARDs such as adverse reactions to a particular csDMARD, so this may not represent disease severity well and this was shown in a previous study [137]. CRP is measured as a marker of inflammation, which may indicate disease severity, however this is measured too infrequently in primary care to be of use. As healthcare utilisation has some evidence behind it [96], this seemed to be the best proxy to use.

Another challenge related to unavailable data is unmeasured confounding [106]. I had planned to study whether influenza and pneumococcal vaccinations prevent infection in patients with RA, as most studies only measure short-term immunogenicity. Unfortunately, despite trying many approaches, we could not deal with the problem of unmeasured confounding. We did not feel confident in publishing the initial results from our Cox PH modelling as influenza vaccination was associated with increased flu symptoms and pneumococcal vaccination was associated with increased risk of pneumonia infection, which was not biologically plausible. I adjusted for additional confounders such as frailty but this did not change the results, I then used a negative control (cellulitis) and saw that vaccinations were associated with increased risk of cellulitis, further supporting unmeasured confounding. I considered a self-controlled study as this eliminates confounding, as each person is their

own control. To use this method the requirements are: 1) recurrent events need to be independent, 2) the event should not influence future exposure and 3) the event must not censor the observation period [138]. This study did not meet the requirements for this type of study as previous pneumonia or influenza infection could influence whether a person received a vaccination. We spent some time working on using an instrumental variable (IV). An IV is a variable that is associated with the exposure, but not the confounders, and is not directly associated with the outcome. Using an IV can reduce unmeasured confounding [139]. The predicted probability of vaccination at the practice and region level was used as an IV, however this was found to be a weak instrument, resulting in wide confidence intervals, therefore we did not continue with this analysis. This was one of the first studies I led, it showed me the limitations of the data and the importance of considering all possible confounders. In this situation the limiting factor was that there was no good measurement of a person's propensity to receive a vaccination. Ultimately, we did not feel sufficiently confident in the results to publish.

In general, these challenges occur because of the way the data is, or is not, captured. After identifying potential problems through insights from clinicians and comparison to previous studies, various methods can be applied to identify whether bias is present, to minimise the chance of bias and, if we cannot minimise its occurrence, to measure the impact of bias. If we produce results that we are confident in then the benefits of the data outweigh the challenges posed.

6.2.4 Data collection using online health communities

The major benefit of using online health communities to collect data is that data can be collected efficiently, over a short period of time, directly from patients. I have shown that this method of data collection can represent some of the traditionally harder to reach groups, such as those who are more deprived. However, there are limitations on the data that can be collected. The studies presented were cross-sectional studies, and descriptive. These types of analyses provide value and important insights for researchers of certain questions,

however this type of data is unlikely to be suitable for studies of complex associations. Surveys can be conducted at more than one time point, and I was involved with a longitudinal study interested in the impact of using HealthUnlocked on health outcomes. Another group designed the study and I was involved with the analysis only. The data was collected at two time points, and only a third of people completed the second survey, those who did complete the second survey were those who engaged with the site frequently and may not represent all HealthUnlocked users, therefore attrition is likely to be a problem in longitudinal studies collecting data in this method. There were interesting findings, but as the data were collected from HealthUnlocked users only, it was difficult to make causal conclusions as there was no control group, so we do not know whether changes seen would have happened regardless of HealthUnlocked use [140]. Although longitudinal studies can be conducted through this platform, careful study design is important.

Publication 7 addressed the issue of representativeness when using online health communities. The results are useful for future studies, if the less well represented groups were particularly important for a study, it may be that data collection could be supplemented by using another data source, such as NRAS itself. Overall, despite the challenges, online health communities are a valuable resource for studies requiring information directly from patients that is difficult to capture otherwise.

6.3: Future research

The studies presented would benefit from further work to translate the results to patients. I will do this through a blog post on the department website and creating a summary video of my findings, to share on social media. The understanding gained from the work in this thesis has led me to number of future research questions, some of which are already in progress. I have used my knowledge of developing code lists and using linked data in another study using CPRD GOLD and Aurum data, investigating the occurrence and outcomes of juvenile idiopathic arthritis (JIA). I am using my experience of modelling GC exposure and risk attribution to develop a protocol that expands the adult GC safety analyses to a wider range

of inflammatory diseases, and extends beyond what I have done to date, in terms of the population studied and the methodology. For example, we plan to use weighted cumulative dose, where exposure is attributed to the event based on the GC dose, duration and recency [141]. Using these new methods will continue my development as an epidemiologist. This work has also led me to be interested in the impact of GC use in children and young people with JIA, partly as most GC research focuses on adults, but also because it brings additional methodological challenges. These drugs are used differently in children and young people compared to adults, with more intra-articular injections in children and young people, for example.

6.4: Conclusions

These studies show how different forms of digital data can be used to study the management of RA. I have shown how EHR data can be used to study both treatment uptake, and complex associations between treatment and side effects. Through these studies, I have learnt about the importance of understanding the data source and how data are (or are not) collected which then informs study design. The importance of exploring the impact of any issues identified, this allows us to be confident in the results. Collecting data through an online health community allowed me to answer questions related to patient perspectives and representativeness that could not have been answered using EHR data. Learning about this efficient method of data collection complements my learning around EHR data. Through these studies, I have gained knowledge of methods that are applicable to other chronic disease areas, and successfully provided answers to clinically important questions related to the management of RA.

References

1 Gordis L. *Epidemiology*. Elsevier/Saunders 2009. https://www.elsevier.com/books/epidemiology/gordis/978-1-4160-4002-6 (accessed 14 Aug2019).

2 Strom B, Kimmel SE, Hennessy Sean. *Pharmacoepidemiology* . 6th ed. 2012. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119413431 (accessed 14 Apr2020).

3 Mues KE, Liede A, Liu J, *et al.* Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. *Clin Epidemiol* 2017;9:267–77.

4 Herrett E, Gallagher AM, Bhaskaran K, *et al.* Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.

5 Wolf A, Dedman D, Campbell J, *et al.* Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol* 2019;48:1740-1740G.

6 Lewis JD, Schinnar R, Bilker WB, *et al.* Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 2007;16:393–401.

7 Home - QResearch. https://www.qresearch.org/ (accessed 9 Nov2020).

8 TPP | SystmOne. https://www.tpp-uk.com/products/systmone (accessed 4 Nov2020).

9 Sudlow C, Gallacher J, Allen N, *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med* 2015;12:1001779.

10 Williamson EJ, Walker AJ, Bhaskaran K, *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430–6.

11 Herbert A, Wijlaars L, Zylbersztejn A, *et al.* Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol* 2017;46:1093-1093i.

12 Joseph RM, Soames J, Wright M, *et al.* Supplementing electronic health records through sample collection and patient diaries: A study set within a primary care research database. *Pharmacoepidemiol Drug Saf* 2018;27:239–42.

13 Khatri C, Chapman SJ, Glasbey J, *et al.* Social Media and Internet Driven Study Recruitment: Evaluating a New Model for Promoting Collaborator Engagement and Participation. *PLoS One* 2015;10:e0118899.

14 Whitaker C, Stevelink S, Fear N. The use of Facebook in recruiting participants for health research purposes: A systematic review. J. Med. Internet Res. 2017;19:e290.

15 Eysenbach G, Powell J, Englesakis M, *et al.* Health related virtual communities and electronic support groups: Systematic review of the effects of online peer to peer interactions. Br. Med. J. 2004;328:1166–70.

16 Fox S, Duggan M. Health Online 2013. 2013. https://www.pewinternet.org/2013/01/15/health-online-2013/ (accessed 19 Jul2019).

17 Internet access – households and individuals, Great Britain - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2018 (accessed 19 Jul2019).

18 Hippisley-Cox J, Coupland C. Risk of myocardial infarction in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis. *BMJ* 2005;330:1366.

19 Herrett E, Thomas SL, Schoonen WM, *et al.* Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010;69:4–14.

20 Margulis A V., Fortuny J, Kaye JA, *et al.* Validation of Cancer Cases Using Primary Care, Cancer Registry, and Hospitalization Data in the United Kingdom. *Epidemiology* 2018;29:308–13.

21 Quint JK, Müllerova H, DiSantostefano RL, *et al.* Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *BMJ Open* 2014;4:5540.

22 Nissen F, Morales DR, Mullerova H, *et al.* Validation of asthma recording in the Clinical Practice Research Datalink (CPRD). *BMJ Open* 2017;7:e017474.

23 Rubin DB. Inference and missing data. *Biometrika* 1976;63:581–92.

24 Rubin DB, Wiley J, York N, *et al.* Multiple Imputation for Nonresponse in Surveys. 1987.

25 Hughes RA, Heron J, Sterne JAC, *et al.* Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *Int J Epidemiol* 2019;48:1294–304.

26 Prada-Ramallal G, Takkouche B, Figueiras A. Bias in pharmacoepidemiologic studies using secondary health care databases: A scoping review. BMC Med. Res. Methodol. 2019;19:1–14.

27 Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. JAMA - J. Am. Med. Assoc. 2011;305:2462–3.

28 McMahon AD. Approaches to combat with confounding by indication in observational studies of intended drug effects. *Pharmacoepidemiol Drug Saf* 2003;12:551–8.

29 Smolen JS, Aletaha D, Barton A, *et al.* Rheumatoid arthritis. *Nat Rev Dis Prim* 2018;4:18001.

30 van der Woude D, van der Helm-van Mil AHM. Update on the epidemiology, risk factors, and disease outcomes of rheumatoid arthritis. Best Pract. Res. Clin. Rheumatol. 2018;32:174–87.

31 Rheumatoid arthritis in adults: Rheumatoid arthritis in adults: management management NICE guideline. 2018. www.nice.org.uk/guidance/ng100 (accessed 20 Mar2020).

32 Aletaha D, Neogi T, Silman AJ, *et al.* 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against

Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.

33 Prevoo MLL, Van'T Hof MA, Kuper HH, *et al.* Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.

34 Welsing PMJ, Fransen J, van Riel PLCM. Is the disease course of rheumatoid arthritis becoming milder?: Time trends since 1985 in an inception cohort of early rheumatoid arthritis. *Arthritis Rheum* 2005;52:2616–24.

35 Diffin JG, Lunt M, Marshall T, *et al.* Has the severity of rheumatoid arthritis at presentation diminished over time? *J Rheumatol* 2014;41:1590–9.

36 Gwinnutt JM, Symmons DPM, MacGregor AJ, *et al.* Have the 10-year outcomes of patients with early inflammatory arthritis improved in the new millennium compared with the decade before? Results from the Norfolk Arthritis Register. *Ann Rheum Dis* 2018;77:848–54.

37 Dougados M, Soubrier M, Antunez A, *et al.* Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: Results of an international, cross-sectional study (COMORA). *Ann Rheum Dis* 2014;73:62–8.

38 Van Halm VP, Peters MJL, Voskuyl AE, *et al.* Rheumatoid arthritis versus diabetes as a risk factor for cardiovascular disease: A cross-sectional study, the CARRÉ Investigation. *Ann Rheum Dis* 2009;68:1395–400.

39 Naranjo A, Sokka T, Descalzo MA, *et al.* Cardiovascular disease in patients with rheumatoid arthritis: results from the QUEST-RA study. *Arthritis Res Ther*;10:R30.

40 Simon TA, Thompson A, Gandhi KK, *et al.* Incidence of malignancy in adult patients with rheumatoid arthritis: A meta-analysis. *Arthritis Res Ther* 2015;17:1–10.

41 Doran MF, Crowson CS, Pond GR, *et al.* Frequency of infection in patients with rheumatoid arthritis compared with controls: a population-based study. *Arthritis Rheum* 2002;46:2287–93.

42 Furer V, Rondaan C, Heijstek M, *et al.* Incidence and prevalence of vaccine preventable infections in adult patients with autoimmune inflammatory rheumatic diseases (AIIRD): A systemic literature review informing the 2019 update of the EULAR recommendations for vaccination in adult patients with AIIRD. *RMD Open* 2019;5:e001041.

43 Haugeberg G, Uhlig T, Falch JA, *et al.* Bone mineral density and frequency of osteoporosis in female patients with rheumatoid arthritis: Results from 394 patients in the Oslo County rheumatoid arthritis register: Results from 394 Patients in the Oslo County Rheumatoid Arthritis Register. 2000.

44 Dadoun S, Zeboulon-Ktorza N, Combescure C, *et al.* Mortality in rheumatoid arthritis over the last fifty years: Systematic review and meta-analysis. *Jt Bone Spine* 2013;80:29–33.

45 Wolfe F, Mitchell DM, Sibley JT, *et al.* The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;37:481–94.

46 Sihvonen S, Korpela M, Mustonen J, *et al.* Mortality in patients with rheumatoid arthritis treated with low-dose oral glucocorticoids. A population-based cohort study. *J Rheumatol* 2006;33:1740–6.

47 Smolen JS, Van Der Heijde D, MacHold KP, *et al.* Proposal for a new nomenclature of disease-modifying antirheumatic drugs. Ann. Rheum. Dis. 2014;73:3–5.

48 Smolen JS, Landewé R, Bijlsma J, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.

49 Ledingham J, Gullick N, Irving K, *et al.* BSR and BHPR guideline for the prescription and monitoring of non-biologic disease-modifying anti-rheumatic drugs. Rheumatol. (United Kingdom). 2017;56:865–8.

50 Quality and Outcomes Framework, 2019-20 - NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2019-20 (accessed 23 Sep2020).

51 Testa A, Arnold Chan Allison Smitten KL, Choi HK, *et al.* The risk of hospitalized infection in patients with rheumatoid arthritis. 2008.

52 Dixon WG, Abrahamowicz M, Beauchamp M-E, *et al.* Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case–control analysis. *Ann Rheum Dis* 2012;71:1128–33.

53 Singh JA, Cameron C, Noorbaloochi S, *et al.* Risk of serious infection in biological treatment of patients with rheumatoid arthritis: a systematic review and meta-analysis. *Lancet* 2015;386:258–65.

54 Kapetanovic M, Lindqvist E, Geborek P, *et al.* Long-term mortality rate in rheumatoid arthritis patients with disease onset in the 1980s. *Scand J Rheumatol* 2011;40:433–8.

55 Sokka T, Abelson B, Pincus T. Mortality in rheumatoid arthritis: 2008 update. *undefined* 2008.

56 Dear KB G, Andrews R, Holden J, *et al.* Vaccines for preventing pneumococcal infection in adults. In: *The Cochrane Database of Systematic Reviews (Complete Reviews)*. John Wiley & Sons, Ltd 2003.

57 Demicheli V, Jefferson T, Ferroni E, *et al.* Vaccines for preventing influenza in healthy adults. Cochrane Database Syst. Rev. 2018;2018.

58 Ramsay M. Green Book chapter Pneumococcal disease. 2018.

59 England PH, editor. Influenza: the green book, chapter 19. 2014.

60 van Assen S, Agmon-Levin N, Elkayam O, *et al.* EULAR recommendations for vaccination in adult patients with autoimmune inflammatory rheumatic diseases. *Ann Rheum Dis* 2011;70:414–22.

61 BSR Statement on Vaccination in Adult Patients with Rheumatic Diseases. 2011.

62 Furer V, Rondaan C, Heijstek MW, *et al.* 2019 update of EULAR recommendations for vaccination in adult patients with autoimmune inflammatory rheumatic diseases. *Ann Rheum Dis* 2019;:annrheumdis-2019-

215882.

63 Lester H, Bch MB, Frcgp MA, *et al.* Developing Quality and Outcomes Framework (QOF) indicators and the concept of ‘QOFability’. iMedPub

64 Centre H and SCI, editor. NHS immunisation statistics 2013-14. 2014.

65 Doe S, Pathare S, Kelly CA, *et al.* Uptake of influenza vaccination in patients on immunosuppressant agents for rheumatological diseases: a follow-up audit of the influence of secondary care. *Rheumatol* 2007;46:715–6.

66 Saravana S. Uptake of influenza vaccination in rheumatology patients. *Rheumatol* 2004;43:1055.

67 Bridges MJ, Coady D, Kelly CA, *et al.* Factors influencing uptake of influenza vaccination in patients with rheumatoid arthritis. *Ann Rheum Dis* 2003;62:685.

68 Clarke AJ, Gulati P, Abraham SM. A cross-sectional audit of the uptake of seasonal and H1N1 influenza vaccination amongst patients with rheumatoid arthritis in a London hospital. *Clin Exp Rheumatol* 2011;29:596.http://www.ncbi.nlm.nih.gov/pubmed/21722505

69 Sowden E, Mitchell WS. An audit of influenza and pneumococcal vaccination in rheumatology outpatients. *BMC Musculoskelet Disord* 2007;8:58.

70 Pradeep J, Watts R, Clunie G. Audit on the uptake of influenza and pneumococcal vaccination in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;66:837–8.

71 Jackson ML. Use of self-reported vaccination status can bias vaccine effectiveness estimates from test-negative studies. *Vaccine X* 2019;1:100003.

72 Park JK, Lee YJ, Shin K, *et al.* Impact of temporary methotrexate discontinuation for 2 weeks on immunogenicity of seasonal influenza vaccination in patients with rheumatoid arthritis: A randomised clinical trial. *Ann Rheum Dis* 2018;77:898–904.

73 Subesinghe S, Bechman K, Rutherford AI, *et al.* A Systematic Review and Metaanalysis of Antirheumatic Drugs and Vaccine Immunogenicity in Rheumatoid Arthritis. *J Rheumatol* 2018;45:733–44.

74 Hua C, Barnetche T, Combe B, *et al.* Effect of methotrexate, anti-tumor necrosis factor α, and rituximab on the immune response to influenza and pneumococcal vaccines in patients with rheumatoid arthritis: A systematic review and meta-analysis. Arthritis Care Res. 2014;66:1016–26.

75 Desai SP, Turchin A, Szent-Gyorgyi LE, *et al.* Routinely measuring and reporting pneumococcal vaccination among immunosuppressed rheumatology outpatients: the first step in improving quality. *Rheumatol* 2011;50:366–72.

76 Thomas SL, Edwards CJ, Smeeth L, *et al.* How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis Rheum* 2008;59:1314–21.

77 Muller S, Hider SL, Raza K, *et al.* An algorithm to identify rheumatoid arthritis in primary care: a Clinical Practice Research Datalink study. *BMJ Open* 2015;5:e009309.

78 Scenario: General principles of managing DMARDs | Management | DMARDs | CKS | NICE. https://cks.nice.org.uk/topics/dmards/management/general-principles-of-managing-dmards/ (accessed 16 Mar2021).

79 Hench PS, Kendall EC, Slocumb CH, *et al.* Effects of Cortisone Acetate and Pituitary Acth on Rheumatoid Arthritis, Rheumatic Fever and Certain Other Conditions - A Study in Clinical Physiology. *Arch Intern Med* 1950;85:545–666.isi:A1950UF22100001

80 Benedek TG. History of the development of corticosteroid therapy. *Clin Exp Rheumatol* 2011;29.

81 Fardet L, Petersen I, Nazareth I. Prevalence of long-term oral glucocorticoid prescriptions in the UK over the past 20 years.

82 van Staa TP, Leufkens HGM, Abenhaim L, *et al.* Use of oral corticosteroids in the United Kingdom. *QJM An Int J Med* 2000;93:105–11.

83 WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group, Sterne JAC, Murthy S, *et al.* Association Between Administration of Systemic Corticosteroids and Mortality Among Critically Ill Patients With COVID-19: A Meta-analysis. *JAMA* Published Online First: 2 September 2020.

84 Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report. *N Engl J Med* Published Online First: 17 July 2020.

85 Hardy RS, Raza K, Cooper MS. Therapeutic glucocorticoids: mechanisms of actions in rheumatic diseases. Nat. Rev. Rheumatol. 2020;16:133–44.

86 Hench PS, Kendall EC, Slocumb CH, *et al.* Adrenocortical Hormone in Arthritis : Preliminary Report. *Ann Rheum Dis* 1949;8:97–104.

87 Buttgereit F, Da Silva JPA, Boers M, *et al.* Standardised nomenclature for glucocorticoid dosages and glucocorticoid treatment regimens: Current questions and tentative answers in rheumatology. *Ann Rheum Dis* 2002;61:718–22.

88 Pevanti 2.5mg Tablets - Summary of Product Characteristics (SmPC) - (emc). https://www.medicines.org.uk/emc/product/1741/smpc#UNDESIRABLE_EFFECTS (accessed 13 May2020).

89 Strehl C, Bijlsma JWJ, De Wit M, *et al.* Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: Viewpoints from an EULAR task force. Ann. Rheum. Dis. 2016;75:952–7.

90 Shared decision making. 2019.www.nice.org.uk/guidance/ktt23 (accessed 7 Sep2020).

91 Stoffer MA, Smolen JS, Woolf A, *et al.* Development of patient-centred standards of care for rheumatoid arthritis in Europe: The eumusc.net project. *Ann Rheum Dis* 2014;73:902–5.

92 Herxheimer A. Communicating with Patients about Harms and Risks. *PLoS Med* 2005;2:e42.

93 van Staa TP, Abenhaim L, Leufkens H. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol* 1994;47:183–9.

94 Pye SR, Sheppard T, Joseph RM, *et al.* Assumptions made when preparing drug exposure data for analysis have an impact on results: An unreported step in pharmacoepidemiology studies. *Pharmacoepidemiol Drug Saf* 2018;27:781–

8.

95 Langan SM, Schmidt SA, Wing K, *et al.* The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363:3532.

96 Ward MM. Rheumatology visit frequency and changes in functional disability and pain in patients with rheumatoid arthritis. *J Rheumatol* 1997;24:35–42.

97 del Rincon I, Battafarano DF, Restrepo JF, *et al.* Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. *Arthritis Rheumatol* 2014;66:264–72.

98 Listing J, Kekow J, Manger B, *et al.* Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNFalpha inhibitors and rituximab. *Ann Rheum Dis* 2015;74:415–21.

99 Leigh JP, Fries JF. Mortality predictors among 263 patients with rheumatoid arthritis. *J Rheumatol* 1991;18:1307–12.http://www.ncbi.nlm.nih.gov/pubmed/1757929

100 Corbett M, Dalton S, Young A, *et al.* Factors predicting death, survival and functional outcome in a prospective study of early rheumatoid disease over fifteen years. *Br J Rheumatol* 1993;32:717–23.http://www.ncbi.nlm.nih.gov/pubmed/8348275

101 Wallberg-Jonsson S, Ohman ML, Dahlqvist SR. Cardiovascular morbidity and mortality in patients with seropositive rheumatoid arthritis in Northern Sweden. *J Rheumatol* 1997;24:445–51.http://www.ncbi.nlm.nih.gov/pubmed/9058647

102 Mitchell DM, Spitz PW, Young DY, *et al.* Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29:706–14.http://www.ncbi.nlm.nih.gov/pubmed/3718563

103 Allebeck P, Rodvall Y, Allander E. Mortality in rheumatoid arthritis, particularly as regards drug use. *Scand J Rheumatol* 1985;14:102–8.http://www.ncbi.nlm.nih.gov/pubmed/4001886

104 Gallagher AM, Dedman D, Padmanabhan S, *et al.* The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in

England compared with the Office for National Statistics death registrations. *Pharmacoepidemiol Drug Saf* 2019;28:563–9.

105 Tamim H, Tahami Monfared AA, LeLorier J. Application of lag-time into exposure definitions to control for protopathic bias. *Pharmacoepidemiol Drug Saf* 2007;16:250–8.

106 Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf* 2006;15:291–303.

107 Sepriano A, Kerschbaumer A, Smolen JS, *et al.* Safety of synthetic and biological DMARDs: A systematic literature review informing the 2019 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2020;79:S760–70.

108 Movahedi M, Beauchamp ME, Abrahamowicz M, *et al.* Risk of Incident Diabetes Mellitus Associated With the Dosage and Duration of Oral Glucocorticoid Therapy in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol* 2016;68:1089–98.

109 Solomon DH, Love TJ, Canning C, *et al.* Risk of diabetes among patients with rheumatoid arthritis, psoriatic arthritis and psoriasis. *Ann Rheum Dis* 2010;69:2114–7.

110 Fox CS, Golden SH, Anderson C, *et al.* Update on prevention of cardiovascular disease in adults with type 2 diabetes mellitus in light of recent evidence: A scientific statement from the American Heart Association and the American diabetes association. *Diabetes Care* 2015;38:1777–803.

111 Sarwar N, Gao P, Kondapally Seshasai SR, *et al.* Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *Lancet* 2010;375:2215–22.

112 Solomon DH, Karlson EW, Rimm EB, *et al.* Cardiovascular morbidity and mortality in women diagnosed with rheumatoid arthritis. *Circulation* 2003;107:1303–7.

113 Turesson C, Jarenros A, Jacobsson L. Increased incidence of cardiovascular

disease in patients with rheumatoid arthritis: Results from a community based study. *Ann Rheum Dis* 2004;63:952–5.

114 Han C, Robinson DW, Hackett M V, *et al.* Cardiovascular disease and risk factors in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. *J Rheumatol* 2006;33.

115 Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol* 2012;41:514–20.

116 Olivarius Nde F, Siersma V, Dyring-Andersen B, *et al.* Patients newly diagnosed with clinical type 2 diabetes during oral glucocorticoid treatment and observed for 14 years: all-cause mortality and clinical developments. *Basic Clin Pharmacol Toxicol* 2011;108:285–8.

117 Van Der Goes MC, Jacobs JWG, Boers M, *et al.* Patient and rheumatologist perspectives on glucocorticoids: An exercise to improve the implementation of the European League Against Rheumatism (EULAR) recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Ann Rheum Dis* 2010;69:1015–21.

118 Huscher D, Thiele K, Gromnica-Ihle E, *et al.* Dose-related patterns of glucocorticoid-induced side effects. *Ann Rheum Dis* 2009;68:1119–24.

119 Panoulas VF, Douglas KM, Stavropoulos-Kalinoglou A, *et al.* Long-term exposure to medium-dose glucocorticoid therapy associates with hypertension in patients with rheumatoid arthritis. *Rheumatology*;47:72–5.

120 Wilson JC, Sarsour K, Gale S, *et al.* Incidence and risk of glucocorticoid-associated adverse effects in patients with rheumatoid arthritis. *Arthritis Care Res* 2019;71:498–511.

121 Jackson SH, Beevers DG, Myers K. Does long-term low-dose corticosteroid therapy cause hypertension? *Clin Sci*;61 Suppl 7:381s-383s.

122 de Burgos-Lunar C, Salinero-Fort MA, Cárdenas-Valladolid J, *et al.* Validation of diabetes mellitus and hypertension diagnosis in computerized medical records in primary health care. *BMC Med Res Methodol* 2011;11:146.

123 Panoulas VF, Douglas KM, Milionis HJ, *et al.* Prevalence and associations of

hypertension and its control in patients with rheumatoid arthritis. *Rheumatology*;46:1477–82.

124 Smithson J. Using and analysing focus groups: limitations and possibilities. 2000. http://www.tandf.co.uk/journals (accessed 4 Dec2020).

125 Wicks P, Thorley EM, Simacek K, *et al.* Scaling PatientsLikeMe via a "generalized platform" for members with chronic illness: Web-based survey study of benefits arising. J. Med. Internet Res. 2018;20:e9909.

126 Murray K, Quinn S, Turk M, *et al.* COVID-19 and rheumatic musculoskeletal disease patients: infection rates, attitudes and medication adherence in an Irish population. *Rheumatology* 2021;60:902–6.

127 Guidry JA, George JN, Vesely SK, *et al.* Corticosteroid side-effects and risk for bleeding in immune thrombocytopenic purpura: patient and hematologist perspectives. *Eur J Haematol* 2009;83:175–82.

128 Forss M, Batcheller G, Skrtic S, *et al.* Current practice of glucocorticoid replacement therapy and patient-perceived health outcomes in adrenal insufficiency - a worldwide patient survey. *BMC Endocr Disord* 2012;12:8.

129 van Rheenen PF, Aloi M, Assa A, *et al.* The Medical Management of Paediatric Crohn's Disease: an ECCO-ESPGHAN Guideline Update. *J Crohn's Colitis* 2020;2020:1–24.

130 Rouse PC, Veldhuijzen Van Zanten JJJCS, Ntoumanis N, *et al.* Measuring the positive psychological well-being of people with rheumatoid arthritis: A cross-sectional validation of the subjective vitality scale. *Arthritis Res Ther* 2015;17:312.

131 Aladul MI, Fitzpatrick RW, Chapman SR. Patients' Understanding and Attitudes Towards Infliximab and Etanercept Biosimilars: Result of a UK Web-Based Survey. *BioDrugs* 2017;31:439–46.

132 McBain H, Shipley M, Newman S. Clinician and Patient Views About Self-Management Support in Arthritis: A Cross-Sectional UK Survey. *Arthritis Care Res* 2018;70:1607–13.

133 Galea S, Tracy M. Participation Rates in Epidemiologic Studies. *Ann Epidemiol*

2007;17:643–53.

134 Chambers D, Preston L, Topakas A, *et al.* Review of diversity and inclusion literature and an evaluation of methodologies and metrics relating to health research.

135 Improving inclusion of under-served groups in clinical research: Guidance from INCLUDE project. https://www.nihr.ac.uk/documents/improving-inclusion-of-under-served-groups-in-clinical-research-guidance-from-include-project/25435 (accessed 26 Apr2021).

136 Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res* 2012;14:e46.

137 Vinet E, Kuriya B, Widdifield J, *et al.* Rheumatoid arthritis disease severity indices in administrative databases: a systematic review. *J Rheumatol* 2011;38:2318–25.

138 Whitaker HJ, Hocine MN, Farrington CP. The methodology of self-controlled case series studies. *Stat Methods Med Res* 2009;18:7–26.

139 MacKenzie TA, Tosteson TD, Morden NE, *et al.* Using instrumental variables to estimate a Cox's proportional hazards regression subject to additive confounding. *Heal Serv Outcomes Res Methodol* 2014;14:54–68.

140 Costello RE, Anand A, Evans MJ, *et al.* Associations between engagement with an online health community and changes in patient activation and health care utilization: Longitudinal web-based survey. *J Med Internet Res* 2019;21:e13477.

141 Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat Med* 2009;28:3437–53.

Copies of each of the publications in their published form.

Publication 1: Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink.

PLOS | ONE

RESEARCH ARTICLE

Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink

Ruth Costello[1]*, Kevin L. Winthrop[2], Stephen R. Pye[1], Benjamin Brown[3], William G. Dixon[1,3]

1 Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute for Inflammation and Repair, Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom, **2** Division of Infectious Diseases, Oregon Health and Science University, Portland, Oregon, United States of America, **3** Health eResearch Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom

* ruth.costello@manchester.ac.uk

OPEN ACCESS

Citation: Costello R, Winthrop KL, Pye SR, Brown B, Dixon WG (2016) Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink. PLoS ONE 11(4): e0153848. doi:10.1371/journal.pone.0153848

Editor: Michael Nurmohamed, VU University Medical Center, NETHERLANDS

Received: November 30, 2015

Accepted: April 5, 2016

Published: April 20, 2016

Data Availability Statement: CPRD data from patients with codes for rheumatoid arthritis (codelists included in S1 Table) were used. CPRD data can be accessed with an appropriate license from the CPRD and with approval from the Independent Scientific Advisory Committee. Licenses are available from CPRD: The Clinical Practice Research Datalink Group, The Medicines and Healthcare products Regulatory Agency, 5th Floor, 151 Buckingham

Abstract

Introduction

Guidelines for the management of rheumatoid arthritis (RA) recommend using influenza and pneumococcal vaccinations to mitigate infection risk. The level of adherence to these guidelines is not well known in the UK. The aims of this study were to describe the uptake of influenza and pneumococcal vaccinations in patients with RA in the UK, to compare the characteristics of those vaccinated to those not vaccinated and to compare vaccination rates across regions of the UK.

Methods

A retrospective cohort study of adults diagnosed with incident RA and treated with non-biologic immunosuppressive therapy, using data from a large primary care database. For the influenza vaccination, patients were considered unvaccinated on 1st September each year and upon vaccination their status changed to vaccinated. For pneumococcal vaccination, patients were considered vaccinated after their first vaccination until the end of follow-up. Patients were stratified by age 65 at the start of follow-up, given differences in vaccination guidelines for the general population.

Palace Road, Victoria, London SW1W 9SZ, England or (http://www.cprd.com).

Funding: This work was supported by the Arthritis Research UK Centre for Epidemiology (20380). WD was supported by an MRC clinican scientist fellowship (G0902272). BB was supported by a Wellcome Trust Research Training fellowship (104438/Z/14/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Results

Overall (N = 15,724), 80% patients received at least one influenza vaccination, and 50% patients received a pneumococcal vaccination, during follow-up (mean 5.3 years). Of those aged below 65 years (N = 9,969), 73% patients had received at least one influenza vaccination, and 43% patients received at least one pneumococcal vaccination. Of those aged over 65 years (N = 5,755), 91% patients received at least one influenza vaccination, and 61% patients had received at least one pneumococcal vaccination. Those vaccinated were older, had more comorbidity and visited the GP more often. Regional differences in vaccination rates were seen with the highest rates in Northern Ireland, and the lowest rates in London.

Conclusions

One in five patients received no influenza vaccinations and one in two patients received no pneumonia vaccine over five years of follow-up. There remains significant scope to improve uptake of vaccinations in patients with RA.

Introduction

Patients with rheumatoid arthritis (RA) are known to have a two-fold increased risk of infections compared to the general population [1]. This is thought to be due to the disease itself, shared risk factors such as smoking, comorbidities and immunosuppressive treatment [2]. For certain infections such as influenza and pneumonia, vaccinations may confer protection [3, 4]. Studies have shown that influenza and pneumococcal vaccinations are safe in patients with RA and produce an antibody response despite immunosuppressive medication [5].

The European League Against Rheumatism (EULAR) recommendations for vaccination in patients with rheumatic diseases recommend vaccination during stable disease, and ideally prior to starting disease-modifying anti-rheumatic drug (DMARD) therapy. A pneumococcal vaccination and annual influenza vaccinations are recommended for patients with RA being treated with immunosuppressive medication. It is unknown whether patients should have booster pneumococcal vaccinations [5]. The UK guidelines recommend a single vaccination for immunocompromised patients [6] and the US guidelines recommend revaccination 5 years after first vaccinations for those below 65 years of age and at age 65 years, or later if at least 5 years have elapsed since the previous dose [7].

Previous studies have investigated the uptake of both influenza and pneumococcal vaccines in patients with rheumatic diseases [8–20]. These have typically been small single centre studies mostly conducted in the US and Europe, often reliant on self-report of vaccinations. They have shown suboptimal uptake of vaccinations, especially pneumococcal vaccinations [9–13, 18–20]. At a national level in the UK, the Department of Health publishes data on vaccination uptake [21], but the figures are not broken down by indication. Only one study in the US has looked at whether patients with rheumatic diseases are being vaccinated prior to starting immunosuppressive therapy [15].

Patients with RA are a high risk group with specific guidelines about vaccination and infection prevention, with little data about vaccination uptake at a national level in the UK. Given this, the aims of this study were to describe the influenza and pneumococcal vaccination uptake

in patients with incident RA in the UK, to compare the characteristics of those vaccinated to those who were not vaccinated and to compare vaccination coverage across regions of the UK.

Materials and Methods

This study used data from the Clinical Practice Research Datalink (CPRD)—a large database of anonymised primary care electronic medical records from general practitioners in the UK. As of March 2011, it contains data for over 12 million patients from the mid 1980's onwards [22]. The records provide a rich source of information on clinical diagnoses and symptoms, immunisations, prescriptions, referrals and tests. The CPRD use data quality metrics to ensure the quality of the data at the individual level, by indicating poor data recording or non-continuous follow-up with an acceptability flag, and at the practice level, by indicating when a practice's data is up to research standard.

Definition of patients with incident RA

This was a retrospective cohort study, with the study period 1st January 2000 to 31st December 2013. To be included in the cohort patients needed to be i) diagnosed with RA for the first time within the study window, identified using Read codes (see S1 Table for codelist) according to a validated definition [23], ii) treated with immunosuppressive therapy at some point during follow-up (identified through product codes (medication codes) for methotrexate, sulfasalazine, leflunomide, hydroxychloroquine and other non-biologic DMARDs) (see S2 Table for codelist). iii) aged 18 years or over and iv) have at least 12 months electronic medical record data prior to entry to the study, to allow determination of baseline characteristics. To ensure data quality, data from patients deemed unacceptable and data prior to the practice being up to standard was not used. Patients entered the study on the date of their RA diagnosis, defined as the first Read code for RA. Follow-up time was censored at either death, transfer out of the practice, when the GP practice stopped contributing CPRD data, or the time of data extraction from CPRD, whichever came sooner.

Vaccination status

Influenza vaccine is an annual vaccine that is adjusted each year depending on the strains of influenza predicted for the influenza season, estimated to be 1st September to 31st March. Exposure to influenza vaccination was therefore time dependent. On 1st September all individuals had an unvaccinated status and, upon vaccination their status changed to vaccinated. On 31st August their status returned to unvaccinated and this was repeated each year. Vaccination for influenza was identified through Read and product codes (see S3 Table for codelists). If there was more than one entry for immunisation during the season the first date was used. Pneumococcal vaccination (PPV23) was identified using Read and product codes (see S4 Table for codelists). Following pneumococcal vaccination patients were considered vaccinated for the rest of follow-up during the study time-period. Any repeat vaccinations remained in the database but did not alter vaccination status. For both vaccines, patients were classified by whether they had any vaccinations during follow-up and the number of vaccinations during follow-up. The expected number of vaccinations during follow-up was calculated for influenza vaccinations.

Covariates

Age at baseline (date of RA diagnosis) was calculated by subtracting year of RA diagnosis from year of birth. Patients were then divided into those above and below 65 years of age, due to

differences in vaccination guidelines in the general population. Baseline covariates were determined using data from at least 12 months prior to cohort entry. Baseline smoking status (ever smoker vs never smoker) was identified using the latest Read code or product code indicating smoking status prior to baseline (see S5 Table for codelists). Median height (calculated using height measurements prior to baseline and during follow-up) and the nearest weight measurement within 5 years prior to RA diagnosis were used to calculate body mass index (BMI) at baseline. Other disease groups for which influenza and pneumonia vaccinations are recommended were identified using Read codes: these included chronic respiratory disease, chronic heart disease, chronic kidney disease, chronic liver disease, chronic neurological disease, diabetes, asplenia or dysfunction of the spleen and other immunosuppression (see S6 Table for codelists). Patients were classified as meeting another clinical risk category at baseline if they had a Read code for at least one of these diseases prior to baseline. DMARD therapy was identified using product codes. The date of first DMARD prescription was identified for each patient, which may have been prior to the date of RA diagnosis. Using this information patients were classified by whether they were vaccinated prior to starting DMARD therapy, using data prior to RA diagnosis where necessary. The number of face-to-face GP consultations during the year prior to baseline was identified from the database, and the mean number of consultations was calculated.

Analysis

For both vaccines, the number and percentage of patients who had at least one vaccination during follow-up, and who had their first vaccination prior to starting immunosuppressive therapy, stratified by age at baseline, was tabulated. The number of expected vaccinations was compared to the number of vaccinations received during follow-up, stratified by age at baseline. For the pneumococcal vaccination, this was described for the whole follow-up period. For the influenza vaccination only the first 5 years are reported due to the high number of vaccinations in some patients. The proportion of those vaccinated for each characteristic was calculated. Proportions vaccinated were compared by calculating the difference in the proportion vaccinated, and 95% confidence intervals (CI), between strata for each characteristic. The region of the practices was identified, and the percentage of patients who received at least one vaccination during follow-up was calculated for each practice, and displayed as a box and whisker plot by region.

The protocol for this study has been approved by Independent Scientific Advisory Committee for Medicines and Healthcare Regulatory Agency database research (Protocol number: 14_173). As this study used routinely collected anonymised electronic health records consent was not required.

Results

As shown in Fig 1 there were 15,724 patients with RA who met the criteria for inclusion in the study. These patients had a mean follow-up of 5.3 years (range: 0.003–14.0 years). Thirty-seven percent of the cohort were age 65 years and over at baseline, and 69% were female. Just over half (54%) of the cohort were smokers or ex-smokers at baseline, 20% had normal BMI, 26% met another clinical risk category at baseline and 22% visited their GP for a face-to-face consultation more than 5 times in the year prior to baseline. The DMARD most frequently prescribed during follow-up was methotrexate, and 50% of patients were prescribed oral glucocorticoids during follow-up (Table 1). 21% had a DMARD prescription prior to baseline, with patients prescribed DMARDs for a mean of 20.6 days (standard deviation: 17.7) prior to RA diagnosis. Overall, 12,492 (80%) patients had received at least one influenza vaccination

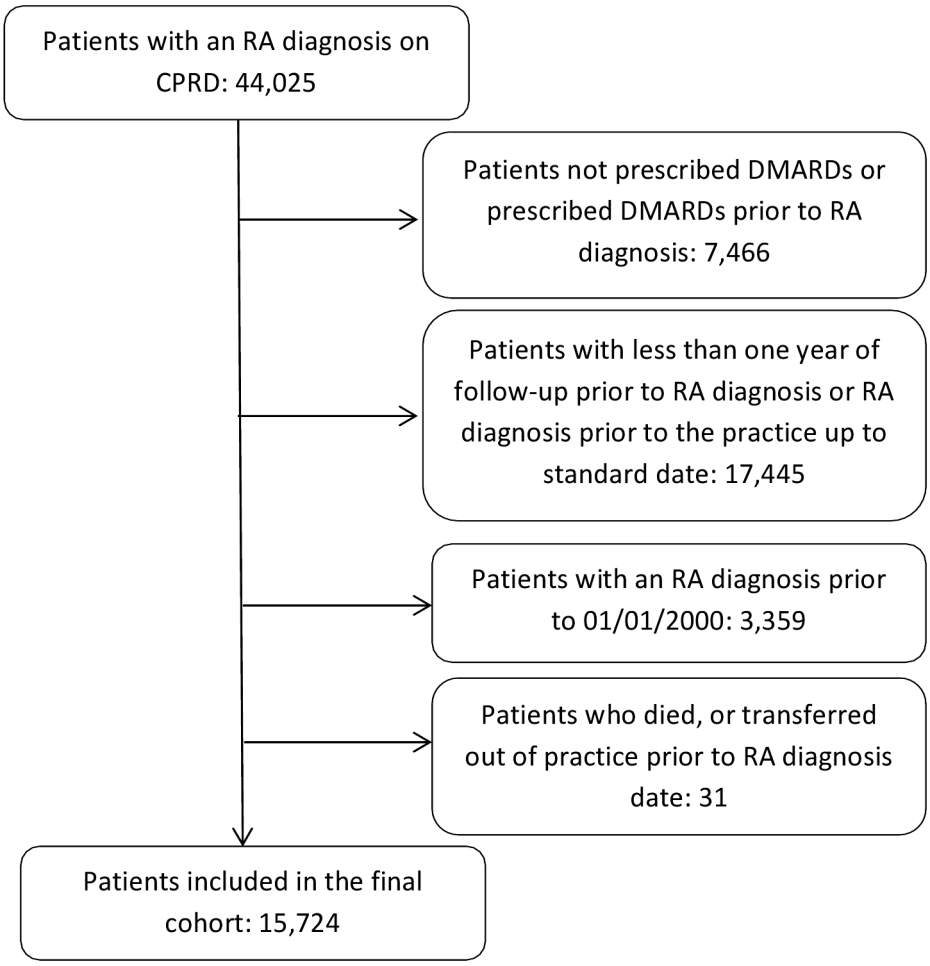


Fig 1. Flowchart of patients eligible for the cohort.

doi:10.1371/journal.pone.0153848.g001

during follow-up and 7,780 (50%) had received at least one pneumococcal vaccination during follow-up (Table 2).

Of those aged below 65 years at baseline (N = 9,969), 7282 (73%) patients had received at least one influenza vaccination. Of those whose first DMARD was during the influenza season (N = 4,092) 1,415 (35%) were vaccinated prior to starting DMARD therapy. Of those expected to have up to 5 vaccinations (N = 4,309), 21%—31% received all expected vaccinations. There were 4,278 (43%) patients who received at least one pneumococcal vaccination, of whom 1,059 (25%) were vaccinated prior to starting DMARD therapy and 175 (1.8%) were revaccinated at least once (Tables 2 and 3).

Of those aged 65 years and above at baseline (N = 5,755), 5210 (91%) patients received at least one influenza vaccination. Of those whose first DMARD was during the influenza season (N = 2,991), 2,220 (74%) were vaccinated prior to starting DMARD therapy. Of those expected to have up to 5 vaccinations (N = 2,961), 55%-76% received all expected vaccinations. There were 3,502 (61%) patients who received at least one pneumococcal vaccination, of whom 2,199 (63%) received a vaccination prior to starting DMARD therapy and 181 (3.1%) were revaccinated at least once (Tables 2 and 3).

Table 1. Characteristics of RA cohort, by vaccination status (N = 15724).

Characteristic	All	Vaccinated for influenza			Vaccinated for pneumonia		
	N (% of cohort)	N	Proportion vaccinated	Difference in proportion vaccinated (95% Confidence Interval)	N	Proportion vaccinated	Difference in proportion vaccinated (95% Confidence Interval)
Baseline age							
18–44 years	2910 (18.5)	1835	63.1	Reference	887	30.5	Reference
45–54 years	3005 (19.1)	2145	71.4	8.3 (5.9, 10.7)	1121	37.3	6.8 (4.4, 9.2)
55–64 years	4054 (25.8)	3302	81.5	18.4 (16.3, 20.5)	2270	56.0	25.5 (23.2, 27.8)
65–74 years	3502 (22.3)	3186	91.0	27.9 (25.9, 29.9)	2318	66.2	35.7 (33.4, 38.0)
Over 75 years	2253 (14.3)	2024	89.8	26.8 (24.6, 28.9)	1184	52.6	22.1 (19.4, 24.7)
Gender							
Female	10781 (68.6)	8525	79.1	Reference	5318	49.3	Reference
Male	4943 (31.4)	3967	80.3	1.2 (-0.2, 2.5)	2462	49.8	0.5 (-1.2, 2.2)
Baseline smoking							
Never smoker	6250 (39.8)	4919	78.7	Reference	3027	48.4	Reference
Ever smoker	8505 (54.1)	6877	80.9	2.2 (0.8, 3.4)	4293	50.5	2.0 (0.4, 3.7)
Missing	969 (6.2)	696	71.8	-6.9 (-9.9, -3.9)	460	47.5	-1.0 (-4.3, 2.4)
Baseline BMI							
Underweight	166 (1.1)	124	74.7	-5.3 (-12.1, 1.4)	69	41.6	-7.6 (-15.3, 0.1)
Underweight (<18.5)							
Normal	3102 (19.7)	2483	80.0	Reference	1526	49.2	Reference
Normal (18.5–24.9)							
Overweight	3502 (22.3)	2949	84.2	4.2 (2.3, 6.0)	1818	51.9	2.7 (0.3, 5.1)
Overweight (25–29.9)							
Obese	2607 (16.6)	2184	83.8	3.7 (1.7, 5.7)	1359	52.1	2.9 (0.3, 5.5)
Obese (30–39.9)							
Morbidly obese	372 (2.4)	297	79.8	-0.2 (-4.5, 4.1)	188	50.5	1.3 (-4.0, 6.7)
Morbidly obese (> = 40)							
Missing	5975 (38.0)	4455	74.6	-5.5 (-7.3, -3.7)	2820	47.2	-2.0 (-4.2, 0.2)
Met at least one other clinical risk category at baseline							
Yes	4034 (25.7)	3652	90.5	14.9 (16.1, 13.7)	2127	52.7	4.4 (6.2, 2.6)
No	11690 (74.3)	8840	75.6	Reference	5653	48.4	Reference
Number of face-to-face GP visits in year prior to baseline							
<5 visits	12329 (78.4)	9466	76.8	Reference	5916	48.0	Reference

(*Continued*)

Table 1. (*Continued*)

Characteristic	All	Vaccinated for influenza			Vaccinated for pneumonia		
	N (% of cohort)	N	Proportion vaccinated	Difference in proportion vaccinated (95% Confidence Interval)	N	Proportion vaccinated	Difference in proportion vaccinated (95% Confidence Interval)
5–9.9 visits	2684 (17.1)	2371	88.3	11.6 (10.1, 13.0)	1429	53.2	5.3 (3.2, 7.3)
10–14.9 visits	556 (3.5)	510	91.7	14.9 (12.5, 17.4)	335	60.3	12.3 (8.1, 16.4)
≥15 visits	155 (1.0)	145	93.5	16.7 (12.8, 20.7)	100	64.5	16.5 (8.9, 24.1)
Prescribed oral glucocorticoids during follow-up							
Yes	7792 (49.5)	6735	84.9	11.0 (12.3, 9.8)	4347	54.8	10.7 (12.3, 9.2)
No	7932 (50.5)	5757	73.9	Reference	3433	44.1	Reference
Prescribed methotrexate during follow-up							
Yes	11453 (72.8)	9517	83.1	13.4 (15.0, 11.9)	6006	52.4	10.9 (12.6, 9.2)
No	4271 (27.2)	2975	69.7	Reference	1774	41.5	Reference
Prescribed hydroxchloroquine during follow-up							
Yes	5593 (35.6)	4433	79.3	-0.2 (1.0, -1.6)	2749	49.2	-0.5 (1.1, -2.1)
No	10131 (64.4)	8059	79.5	Reference	5031	49.7	Reference
Prescribed sulfasalazine during follow-up							
Yes	7344 (46.7)	5869	79.9	0.9 (2.1, -0.4)	3773	51.4	3.6 (5.1, 2.0)
No	8380 (53.3)	6623	79.0	Reference	4007	47.8	Reference
Prescribed leflunomide during follow-up							
Yes	1838 (11.7)	1594	86.7	8.2 (9.9, 6.5)	1093	59.5	11.3 (13.7, 8.9)
No	13886 (88.3)	10898	78.5	Reference	6687	48.2	Reference
Prescribed other non-biologic DMARDs during follow-up							
Yes	1205 (7.7)	1047	86.9	8.1 (10.1, 6.0)	786	65.2	17.1 (19.9, 14.2)
No	14519 (92.3)	11445	78.8	Reference	6994	48.2	Reference

doi:10.1371/journal.pone.0153848.t001

Of those aged 65 years and above at baseline (N = 5,755), 5210 (91%) patients received at least one influenza vaccination. Of those whose first DMARD was during the influenza season (N = 2,991), 2,220 (74%) were vaccinated prior to starting DMARD therapy. Of those expected to have up to 5 vaccinations (N = 2,961), 55%-76% received all expected vaccinations. There were 3,502 (61%) patients who received at least one pneumococcal vaccination, of whom 2,199

Table 2. Influenza and pneumonia vaccination uptake, and timing of vaccinations in relation to starting DMARD therapy (N = 15724).

	Influenza vaccination N (%)			Pneumonia vaccination N (%)		
	<65 years	≥65 years	Total	<65 years	≥65 years	Total
Ever had a vaccination						
Yes	7282 (73.0)	5210 (90.5)	12492 (79.5)	4278 (42.9)	3502 (60.9)	7780 (49.5)
No	2687 (27.0)	545 (9.5)	3232 (20.5)	5691 (57.1)	2253 (39.2)	7944 (50.5)
First vaccination prior to starting DMARDs[1]						
Yes	1415 (34.6)	2220 (74.2)	3635 (51.3)	1059 (24.7)	2199 (62.8)	3258 (41.9)
No	2677 (65.4)	771 (25.8)	3448 (48.7)	3219 (75.3)	1303 (37.2)	4522 (58.1)

[1] For influenza vaccination: Includes patients whose first DMARD was during influenza season (September-March) as patients would not be vaccinated between April and August (N = 7083).

doi:10.1371/journal.pone.0153848.t002

(63%) received a vaccination prior to starting DMARD therapy and 181 (3.1%) were revaccinated at least once (Tables 2 and 3).

The characteristics of the study cohort are shown in Table 1. Those who were vaccinated for influenza were older, 91% of those aged 65–74 years at baseline and 90% of those aged 75 years or older at baseline were vaccinated compared to 63% of those aged 18–44 years at baseline, giving differences of 28 percentage points (95% CI: 26, 30 percentage points) and 27 percentage

Table 3. The expected versus received influenza and pneumonia vaccinations of the RA cohort, by age group.

Vaccinations expected	Vaccinations received	Influenza vaccination[1] N (%)			Pneumonia vaccination[2] N (%)		
		<65 years	≥65 years	Total	<65 years	≥65 years	Total
1	0/1	140 (68.6)	33 (24.3)	173 (50.9)	5691 (57.1)	2253 (39.2)	7944 (50.5)
	1/1	64 (31.4)	103 (75.7)	167 (49.1)	4103 (41.2)	3321 (57.7)	7424 (47.2)
	2+/1	-	-	-	175 (1.8)	181 (3.1)	356 (2.3)
2	0/2	470 (46.8)	106 (14.6)	576 (33.3)			
	1/2	263 (26.2)	133 (18.4)	396 (22.9)	-	-	-
	2/2	272 (27.1)	485 (67.0)	757 (43.8)			
3	0/3	395 (35.8)	92 (11.8)	487 (25.8)			
	1/3	195 (17.7)	65 (8.3)	260 (13.8)	-	-	-
	2/3	253 (22.9)	160 (20.4)	413 (21.9)			
	3/3	262 (23.7)	466 (59.5)	728 (38.6)			
4	0/4	297 (29.2)	60 (8.9)	357 (21.1)			
	1/4	133 (13.1)	28 (4.1)	161 (9.5)			
	2/4	128 (12.6)	47 (7.0)	175 (10.3)	-	-	-
	3/4	251 (24.7)	146 (21.6)	397 (23.5)			
	4/4	208 (20.5)	395 (58.4)	603 (35.6)			
5	0/5	269 (27.5)	62 (9.7)	331 (20.4)			
	1/5	108 (11.0)	20 (3.1)	128 (7.9)			
	2/5	93 (9.5)	21 (3.3)	114 (7.0)	-	-	-
	3/5	123 (12.6)	49 (7.6)	172 (10.6)			
	4/5	172 (17.6)	135 (21.0)	307 (19.0)			
	5/5	213 (21.8)	355 (55.3)	568 (35.1)			

[1] Includes those who were expected to receive up to 5 vaccinations only (N = 7270).

[2] Whole cohort (N = 15724)

doi:10.1371/journal.pone.0153848.t003

points (95% CI: 25, 29 percentage points), respectively. This was similar for pneumococcal vaccinations, though those aged over 75 years at baseline had a lower percentage of coverage (53%) than those aged 65–74 years at baseline (66%). There were small differences in influenza vaccination coverage between baseline BMI categories, with higher coverage observed in the overweight and obese categories (84% in both categories) compared to the normal category (80%), a difference of only 4 percentage points. A similar pattern was observed for pneumococcal vaccination, though again the coverage was lower than for influenza vaccination. A greater proportion of those who met at least one other clinical risk category had been vaccinated, compared to those who did not meet another clinical risk category. This was true for both influenza and pneumococcal vaccinations, though there was a greater difference for those vaccinated for influenza, with a difference 15 percentage points (95% CI: 16, 14 percentage points) compared to a difference of 4 percentage points (95% CI 6, 3 percentage points) for pneumococcal vaccinations. Those who visited their GP more often were more likely to be vaccinated with either vaccine. A greater proportion of those prescribed oral glucocorticoids, methotrexate, leflunomide and other DMARDs were vaccinated, for both types of vaccine. A greater proportion of those prescribed sulfasalazine had a pneumococcal vaccination, though this was not true for influenza vaccination. There were no differences in the proportion vaccinated in those prescribed hydroxychloroquine. There were small differences in the proportion vaccinated by smoking status. Ever smokers were vaccinated with either vaccine slightly more than never smokers. The proportion vaccinated with either vaccine did not differ by gender.

Regional differences were observed in both influenza and pneumococcal vaccination coverage. The highest coverage of influenza and pneumococcal vaccination was in Northern Ireland (86% and 61% respectively). The lowest coverage was in London (72% and 43% respectively) (Fig 2). The differences in regional variation were unchanged when stratified by age.

Discussion

This large cohort study, using electronic health records from GPs in the UK, has shown that one in five immunosuppressed patients with RA did not receive any influenza vaccinations during follow-up, and up to two thirds were not vaccinated annually. Half of those vaccinated received their first vaccination prior to starting DMARDs. Uptake of pneumococcal vaccinations was much lower, less than half were vaccinated, and less than 3% of patients received booster vaccinations. Of those vaccinated, 50% were vaccinated prior to starting DMARDs. Those who were younger, who did not meet another clinical risk category, and who visited their GP less often were least likely to be vaccinated. There was some variation in the proportion who received at least one vaccination by region with Northern Ireland having the highest coverage, and London having the lowest coverage.

The EULAR recommendations for vaccination in patients with rheumatic diseases [5], and the UK vaccination guidelines for patients who are immunosuppressed [6, 24], recommend an annual influenza vaccination and a pneumococcal vaccination. To our knowledge, this is the first large study to describe the uptake of influenza and pneumococcal vaccinations, and timing of vaccinations in relation to starting DMARD therapy, in patients with incident RA in the UK. There have been a small number of single centre audits of vaccination uptake in patients with rheumatic diseases in the UK [9–13, 19, 20] (Table 4), only 3 of these were investigating patients with RA specifically [10–12]. The audits had between 64 and 169 patients and found influenza vaccination uptake to be between 56% and 79% [9–13, 19, 20]. Pneumococcal vaccination rates were lower at between 33% and 43% [11, 13, 20]. Though in one study influenza vaccination uptake varied between 54% and 93%, and pneumococcal vaccination varied between 38% and 64% depending on the type of DMARD the patient was taking and whether

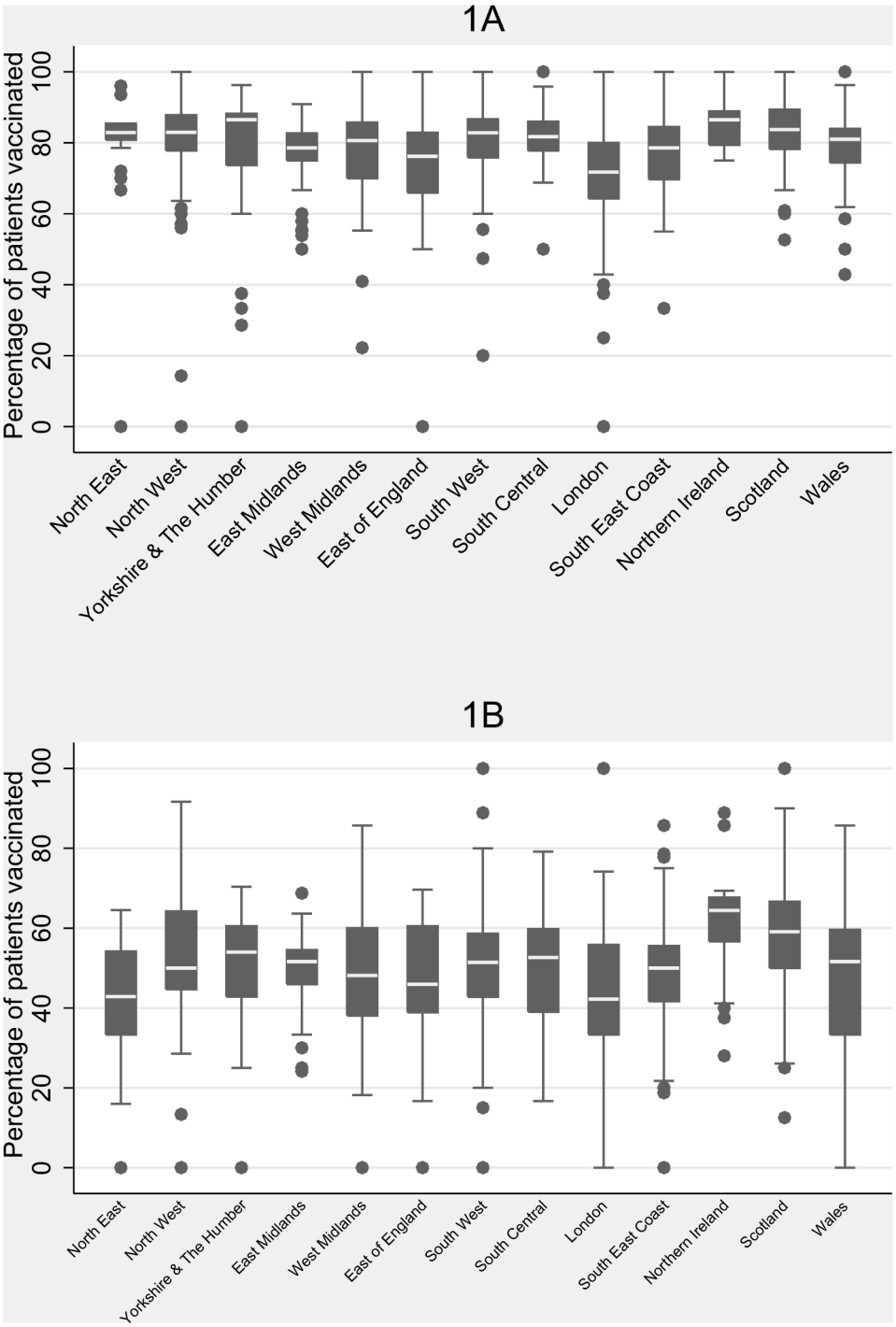


Fig 2. Box and whisker plots of influenza and pneumococcal vaccination uptake by region. Box and whisker plots showing the percentage of patients within a practice receiving at least one influenza vaccination (1A) or pneumonia vaccination (1B), by region. Box plots represent the median (central line), interquartile range (box), range, excluding outliers (whiskers) and outliers (dots) of the percentage of patients within a practice who receive at least one vaccination during follow-up.

doi:10.1371/journal.pone.0153848.g002

Table 4. Summary of studies of influenza and pneumococcus vaccination uptake in the UK.

Author	Type of study	N	Disease group	Influenza vaccine uptake	Pneumococcal vaccine uptake
Pradeep et al (2006)	Audit	64	Rheumatoid arthritis	63%	43%
Doe et al (2007)	Audit	169	Rheumatic diseases	79%	34%
Thomas et al (2004)	Audit	111	Rheumatic diseases	70%	33%
Bridges et al (2003)	Audit	129	Rheumatoid arthritis	56% (of those taking MTX (n = 59))	-
Clarke et al (2011)	Audit	71	Rheumatoid arthritis	~70%	-
Saravana et al (2004)	Audit	100	Rheumatic diseases	77%	-
Sowden et al (2007)	Audit	101	Rheumatic diseases	54%-93%	38%-64%
Hmamouchi et al (2015)	Cohort	43 (UK patients)	Rheumatoid arthritis	84%	44%

doi:10.1371/journal.pone.0153848.t004

they had any other risk factors for vaccination [9]. Another study used data from a cohort study, which included 43 UK patients, and described vaccination uptake [16]. The study found that 84% of UK patients had ever received an influenza vaccination, only 30% had optimal use, and 44% of UK patients had ever received a pneumococcal vaccination. These results were similar to this study. The Department of Health estimates that 52% of those on immunosuppressive medication aged 18–64 years, were vaccinated for influenza in 2013/14 and 73% of those aged 65 years and over were vaccinated for influenza [21]. This study has found slightly higher numbers had received at least one influenza vaccination; compared to previous audits, however when broken down by the number of expected and received vaccinations the patients receiving all vaccinations expected was lower than the previous audits at between 33%-50% for those with 5 years follow-up, which is suboptimal. The percentage of patients who received a pneumococcal vaccination during follow-up was similar to previous audits.

In previous studies patients were only expected to receive one pneumococcal vaccination, with no boosters. Although there are guidelines recommending boosters every 5 years [7], the green book for GPs in the UK recommend only vaccinating once [6], and the EULAR guidelines state that it is unknown whether boosters are required [5], so this may explain why patients have not received booster vaccinations and perhaps needs clarification.

This study also found that those with another indication for vaccination, in particular being aged 65 years or over had higher rates of vaccination, which was similar to previous audits. Only one previous study in the US had reported the proportion of patients with rheumatic diseases who received a pneumococcal vaccination prior to starting immunosuppressive therapy [15]. They found 37% of patients were vaccinated prior to starting immunosuppressive therapy, which was similar to this study where 42% of those vaccinated received a pneumococcal vaccination prior to starting DMARD therapy. There was wide variation by age, those below 65 years of age were much less likely to have been vaccinated prior to starting DMARDs (only 35% and 25% for influenza and pneumonia vaccination, respectively, for those below 65 years of age compared to 74% and 63% in those over 65 years of age). Future research is required to see whether this confers a clinical benefit in terms of reducing the incidence of infection. Because DMARDs can be initiated in hospitals with GP prescribing only after the first few months' hospital treatment, the proportions of patients vaccinated prior to DMARD therapy may be an over-estimate. It may be more difficult to vaccinate for influenza prior to starting

DMARDs, as vaccination takes place at a specific point in the year, however pneumonia vaccination can take place all year round, so should be easier to accomplish.

Interestingly there was some variation in uptake of vaccinations by region and the trend was similar to the variation seen in the NHS immunisation statistics for England 2013/14 [25]. Northern Ireland had the highest rates of vaccination and London had the lowest rates of vaccination, perhaps due to regional differences in the promotion of vaccination.

In the UK there are incentives for GPs to provide influenza vaccinations for those aged 65 years and over and those with coronary heart disease, a history of stroke or transient ischaemic attacks, diabetes and chronic obstructive pulmonary disease, through Quality and Outcomes Framework (QOF) targets. Pneumonia vaccinations for those aged 65 years and over, and influenza vaccinations for those at-risk but not on the QOF indicators (including those who are immunosuppressed), are implemented through enhanced services. These provide GPs with payment for immunisations, though do not specify targets for how many should be vaccinated. This may explain the low vaccination rates observed, particularly for pneumonia vaccinations, where the enhanced services only covers those aged 65 years or over and it does not cover at risk groups. The UK influenza and pneumococcal vaccination booklets are not specific regarding RA, and the decision on whether to vaccinate individual immunosuppressed patients is left to clinician discretion. Therefore, it may be beneficial for rheumatologists to provide more input into the vaccination process. For example, they may wish to consider administering vaccines themselves prior to initiating immunosuppressive therapy, or provide GPs with clear advice on when vaccines should be administered. Regardless, experience tells us it is essential that both approaches should be implemented and resourced or they will be ineffective.

The study has several advantages—it used a large sample of patients with RA therefore the study population is likely to be representative of patients with RA in the UK. The data used came from electronic medical records which were recorded at the time of the visit so there should not be inconsistencies in the way GPs recorded the data or in how patients reported their symptoms. The data is "real-world data" and contains information on administered vaccinations, rather than being self-reported, therefore is likely to be accurate and free from recall bias. In the UK vaccinations primarily take place in primary care therefore most vaccinations will have been identified using primary care electronic medical records. There are however some limitations to be considered when interpreting the results. There may be some misclassification with respect to the identification of diseases such as RA, within CPRD as these are coded by the practices. However, vaccinations should be coded accurately as the product codes are generally only added to the health records after administration by a clinician, hence representing administration and not prescription. In addition, the vaccination codes are used to identify QOF and enhanced services compliance to determine payment, so GPs have an incentive to ensure they are accurate. Studies have shown that patients with RA on biologic DMARD therapy were more likely to have received a pneumococcal vaccination [8, 16]. Biologic DMARD therapy is not captured on the CPRD database as these are prescribed in secondary care, therefore we do not know what influence this had on vaccination uptake. There was some missing data for BMI and smoking as this is collected opportunistically, however the amount of missing data was small, particularly for smoking.

In conclusion, despite international recommendations, this study has found that many patients with RA in the UK are not being immunised regularly for influenza, and often not at all for pneumonia. Many patients are not being immunised prior to starting DMARD therapy. The patients most often being missed are those who are below 65 years of age and who do not have another disease for which vaccination is recommended.

Supporting Information

S1 Table. Rheumatoid arthritis codelist.
(XLSX)

S2 Table. Disease-Modifying Anti-Rheumatic Drug codelist.
(XLSX)

S3 Table. Influenza vaccination codelists.
(XLSX)

S4 Table. Pneumococcal vaccination codelists.
(XLSX)

S5 Table. Smoking codelists.
(XLSX)

S6 Table. Clinical risk group codelists.
(XLSX)

Author Contributions

Conceived and designed the experiments: WD KW RC. Performed the experiments: RC. Analyzed the data: RC KW SP BB WD. Wrote the paper: RC KW SP BB WD.

References

1. Doran MF, Crowson CS, Pond GR, O'Fallon WM, Gabriel SE. Frequency of infection in patients with rheumatoid arthritis compared with controls: a population-based study. Arthritis and rheumatism. 2002; 46(9):2287–93. Epub 2002/10/02. doi: 10.1002/art.10524 PMID: 12355475.
2. Doran MF, Crowson CS, Pond GR, O'Fallon WM, Gabriel SE. Predictors of infection in rheumatoid arthritis. Arthritis and rheumatism. 2002; 46(9):2294–300. Epub 2002/10/02. doi: 10.1002/art.10529 PMID: 12355476.
3. Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. The Lancet infectious diseases. 2012; 12(1):36–44. Epub 2011/10/29. doi: 10.1016/S1473-3099(11)70295-X PMID: 22032844.
4. Moberley S, Holden J, Tatham DP, Andrews RM. Vaccines for preventing pneumococcal infection in adults. The Cochrane database of systematic reviews. 2013; 1:CD000422. Epub 2013/02/27. doi: 10.1002/14651858.CD000422.pub3 PMID: 23440780.
5. van Assen S, Agmon-Levin N, Elkayam O, Cervera R, Doran MF, Dougados M, et al. EULAR recommendations for vaccination in adult patients with autoimmune inflammatory rheumatic diseases. Annals of the rheumatic diseases. 2011; 70(3):414–22. Epub 2010/12/07. doi: 10.1136/ard.2010.137216 PMID: 21131643.
6. Pneumococcal: the green book, chapter 25. In: England PH, editor. 2013.
7. Centers for Disease C, Prevention. Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine for adults with immunocompromising conditions: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Morbidity and mortality weekly report. 2012; 61(40):816–9. Epub 2012/10/12. PMID: 23051612.
8. Feuchtenberger M, Kleinert S, Schwab S, Roll P, Scharbatke EC, Ostermeier E, et al. Vaccination survey in patients with rheumatoid arthritis: a cross-sectional study. Rheumatology international. 2012; 32(6):1533–9. Epub 2011/02/18. doi: 10.1007/s00296-011-1808-z PMID: 21327432.
9. Sowden E, Mitchell WS. An audit of influenza and pneumococcal vaccination in rheumatology outpatients. BMC musculoskeletal disorders. 2007; 8:58. Epub 2007/07/06. doi: 10.1186/1471-2474-8-58 PMID: 17610723; PubMed Central PMCID: PMC1949405.
10. Clarke AJ, Gulati P, Abraham SM. A cross-sectional audit of the uptake of seasonal and H1N1 influenza vaccination amongst patients with rheumatoid arthritis in a London hospital. Clinical and experimental rheumatology. 2011; 29(3):596. Epub 2011/07/05. PMID: 21722505.

11. Pradeep J, Watts R, Clunie G. Audit on the uptake of influenza and pneumococcal vaccination in patients with rheumatoid arthritis. Annals of the rheumatic diseases. 2007; 66(6):837–8. Epub 2007/05/22. doi: 10.1136/ard.2006.060285 PMID: 17513572; PubMed Central PMCID: PMC1954646.

12. Bridges MJ, Coady D, Kelly CA, Hamilton J, Heycock C. Factors influencing uptake of influenza vaccination in patients with rheumatoid arthritis. Annals of the rheumatic diseases. 2003; 62(7):685. Epub 2003/06/18. PMID: 12810440; PubMed Central PMCID: PMC1754592.

13. Doe S, Pathare S, Kelly CA, Heycock CR, Binding J, Hamilton J. Uptake of influenza vaccination in patients on immunosuppressant agents for rheumatological diseases: a follow-up audit of the influence of secondary care. Rheumatology. 2007; 46(4):715–6. Epub 2007/01/25. doi: 10.1093/rheumatology/kel410 PMID: 17244667.

14. Lanternier F, Henegar C, Mouthon L, Blanche P, Guillevin L, Launay O. Low influenza-vaccination rate among adults receiving immunosuppressive therapy for systemic inflammatory disease. Annals of the rheumatic diseases. 2008; 67(7):1047. Epub 2008/06/17. doi: 10.1136/ard.2007.081703 PMID: 18556445.

15. Desai SP, Turchin A, Szent-Gyorgyi LE, Weinblatt M, Coblyn J, Solomon DH, et al. Routinely measuring and reporting pneumococcal vaccination among immunosuppressed rheumatology outpatients: the first step in improving quality. Rheumatology. 2011; 50(2):366–72. Epub 2010/10/27. doi: 10.1093/rheumatology/keq297 PMID: 20974614; PubMed Central PMCID: PMC3021950.

16. Hmamouchi I, Winthrop K, Launay O, Dougados M. Low rate of influenza and pneumococcal vaccine coverage in rheumatoid arthritis: data from the international COMORA cohort. Vaccine. 2015; 33 (12):1446–52.

17. McCarthy EM, de Barra E, Bergin C, Cunnane G, Doran M. Influenza and pneumococcal vaccination and varicella status in inflammatory arthritis patients. Irish medical journal. 2011; 104(7):208–11.

18. Haroon M, Adeeb F, Eltahir A, Harney S. The uptake of influenza and pneumococcal vaccination among immunocompromised patients attending rheumatology outpatient clinics. Joint, bone, spine: revue du rhumatisme. 2011; 78(4):374–7. Epub 2010/12/21. doi: 10.1016/j.jbspin.2010.10.012 PMID: 21169045.

19. Saravana S. Uptake of influenza vaccination in rheumatology patients. Rheumatology. 2004; 43 (8):1055. Epub 2004/07/20. doi: 10.1093/rheumatology/keh263 PMID: 15258293.

20. Thomas LM, Wood A, Shenton S, Ledingham JM. Uptake of influenza vaccination in rheumatology patients: reply. Rheumatology. 2005; 44(1):131. Epub 2004/12/22. doi: 10.1093/rheumatology/keh452 PMID: 15611306.

21. Seasonal flu uptake (GP) 2013/14—Data on GP registered patients. In: England PH, editor. 2014.

22. Williams T, van Staa T, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. Therapeutic advances in drug safety. 2012; 3(2):89–99. Epub 2012/04/01. doi: 10.1177/2042098611435911 PMID: 25083228; PubMed Central PMCID: PMC4110844.

23. Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? Arthritis and rheumatism. 2008; 59(9):1314–21. Epub 2008/09/02. doi: 10.1002/art.24015 PMID: 18759262.

24. Influenza: the green book, chapter 19. In: England PH, editor. 2014.

25. NHS immunisation statistics 2013–14. In: Centre HaSCI, editor. 2014.

Publication 2: Pneumonia vaccination timing in relation to starting conventional synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis.

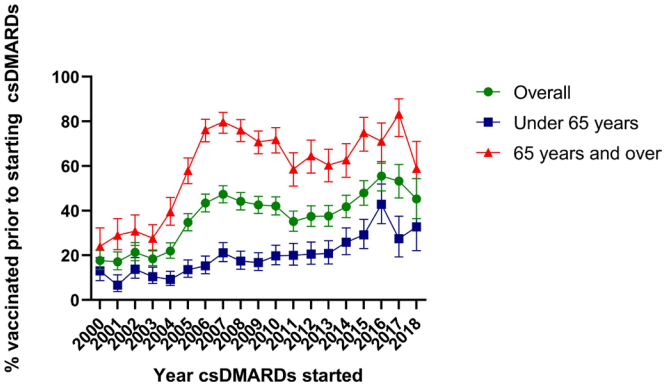


Figure 1 Timing of vaccination in relation to starting csDMARDs by year, overall and stratified by age. csDMARD, conventional synthetic disease-modifying anti-rheumatic drug.

Pneumonia vaccination timing in relation to starting conventional synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis

Patients with rheumatoid arthritis (RA) are at increased risk of infections, and pneumococcal vaccination is recommended. With some evidence that pneumococcal vaccinations are not as effective when administered after starting disease-modifying antirheumatic drugs (DMARDs), in particular methotrexate,[1] guidance on when best to vaccinate, in relation to DMARDs, has become more consistent in recent years. Early European League Against Rheumatism guidelines (2011) only referred to B-cell depleting biological DMARDs, but more recent guidelines (2019)[2] recommend vaccination prior to commencement of all DMARD types. Since 2011, British Society for Rheumatology (BSR) guidance advises vaccination prior to starting any DMARD.[3 4] The aims of this study were to explore the timing of pneumococcal vaccination in patients with RA in relation to starting conventional synthetic DMARDs (csDMARDs) and examine whether this has changed over time.

This was a cross-sectional study using data from the Clinical Practice Research Datalink GOLD (UK primary care electronic health records). The study period was from 1 January 2000 to 31 December 2018. To be included, patients were required to (1) have a diagnosis of RA identified using a validated algorithm[5]; (2) be prescribed csDMARDs up to a maximum of 3 months prior to, or anytime after, RA diagnosis and (3) have received a pneumococcal vaccination up to a maximum of 5 years prior to, or anytime after, starting csDMARDs. For each patient, it was determined if vaccination was prior to starting csDMARDs.

Of 21 461 patients with RA who started csDMARDs within the study window, 8205 (38.2%) were vaccinated and met the inclusion criteria. Nearly half (44.3%, n=3633) were age ≥65 years, 66.4% (n=5445) were female and 26.7% (n=2188) had a diagnosis for another disease where vaccination is also recommended. Overall, 2997 (36.5%) patients were vaccinated prior to starting csDMARDs. When stratified by age, of those vaccinated prior to starting csDMARDs, 88% (n=1911/2170) of those age ≥65 years and 72% (n=596/827) of those <65 years, were vaccinated prior to RA diagnosis. The frequency of vaccination was higher in the first year after starting csDMARDs, with 1779 (21.7%) vaccinated compared to 833 (10.2%) in the year preceding (online supplementary figure 1). However, 1000 (12.2%) were vaccinated >3 years prior, and 1844 (22.5%) were vaccinated >3 years after starting csDMARDs. By calendar year, the proportion vaccinated prior to starting csDMARDs increased over time, with the greatest increases seen between 2003 and 2007 where the proportion increased from 18% to 47%. When stratified by age (65 years or over, when UK guidelines recommend vaccinating everyone against pneumococcus), the proportions vaccinated prior to starting csDMARDs were higher overall in the ≥65 years old age group. In those <65 years, the proportions rose more steadily over time from 13.0% in 2000 to 29.2% in 2015 (figure 1).

This study has shown that, of patients with RA who received pneumococcal vaccinations, only around one-third of vaccinations occurred prior to starting csDMARDs. This is similar to a study in the USA where 41% were vaccinated prior to starting csDMARDs.[6] There was evidence that commencement of csDMARDs prompts vaccination, however the peak in vaccination was in the year after starting csDMARDs. Vaccination prior to starting csDMARDs has increased through time with greater increases seen in those aged ≥65 years. A marked increase followed the 2003 change in national guidelines recommending all adults≥65 years should receive a pneumococcal vaccination. Indeed, most patients ≥65 years old vaccinated prior to starting csDMARDs were vaccinated prior to RA diagnosis. In those <65 years old, there were still increases over time, which was positive. Encouragingly, overall, there was also a steady increase from 2011, when BSR vaccination

guidelines were published.[3] Although the proportion decreases in 2018, further data are required to determine if this is a long-term trend. Given recent guidelines, we encourage rheumatologists to promote awareness of the importance of vaccinations prior to csDMARD initiation through timely communications to patients and primary care physicians.

Ruth E Costello [1], **Jenny H Humphreys** [1], **Kevin L Winthrop** [2], **William G Dixon** [1,3]

[1]Centre for Epidemiology Versus Arthritis, The University of Manchester, Manchester, UK
[2]Division of Infectious Diseases, Oregon Health & Science University, Portland, Oregon, USA
[3]Department of Rheumatology, Salford Royal NHS Foundation Trust, Salford, UK

Correspondence to Mrs Ruth E Costello, Centre for Epidemiology Versus Arthritis, The University of Manchester, Manchester M13 9PL, UK; ruth.costello@manchester.ac.uk

Handling editor Josef S Smolen

Twitter William G Dixon @WGDixon

Contributors All authors made substantial contributions to the design and interpretation of the data. REC made substantial contributions to the analysis and drafted the manuscript. All authors revised the manuscript for important intellectual content, gave final approval of the version published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

Funding This work was supported by the Centre for Epidemiology Versus Arthritis (Grant number 21755) and supported by the NIHR Manchester Biomedical Research Centre.

Competing interests WGD has received consultancy fees from Bayer, Abbvie and Google unrelated to this work. All other authors have no conflicts of interest.

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.

OPEN ACCESS

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2020-217255).

To cite Costello RE, Humphreys JH, Winthrop KL, *et al*. *Ann Rheum Dis* 2020;**79**:1665–1666.

Received 28 February 2020
Revised 20 April 2020
Accepted 3 June 2020
Published Online First 26 June 2020

Ann Rheum Dis 2020;**79**:1665–1666. doi:10.1136/annrheumdis-2020-217255

ORCID iDs
Ruth E Costello http://orcid.org/0000-0003-2709-6666
Jenny H Humphreys http://orcid.org/0000-0003-2497-4080
Kevin L Winthrop http://orcid.org/0000-0002-3892-6947
William G Dixon http://orcid.org/0000-0001-5881-4857

REFERENCES

1 Subesinghe S, Bechman K, Rutherford AI, *et al*. A systematic review and Metaanalysis of antirheumatic drugs and vaccine immunogenicity in rheumatoid arthritis. *J Rheumatol* 2018;45:733–44.
2 Furer V, Rondaan C, Heijstek MW, *et al*. 2019 update of EULAR recommendations for vaccination in adult patients with autoimmune inflammatory rheumatic diseases. *Ann Rheum Dis* 2020;79:39–52.
3 BSR statement on vaccination in adult patients with rheumatic diseases 2011.
4 Ledingham J, Gullick N, Irving K, *et al*. BSR and BHPR guideline for the prescription and monitoring of non-biologic disease-modifying anti-rheumatic drugs. *Rheumatology*
5 Thomas SL, Edwards CJ, Smeeth L, *et al*. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis Rheum* 2008;59:1314–21.
6 Desai SP, Turchin A, Szent-Gyorgyi LE, *et al*. Routinely measuring and reporting pneumococcal vaccination among immunosuppressed rheumatology outpatients: the first step in improving quality. *Rheumatology* 2011;50:366–72.

Publication 3: Oral glucocorticoid therapy and all-cause and cause-specific mortality in patients with rheumatoid arthritis: a retrospective cohort study.

Eur J Epidemiol (2016) 31:1045–1055
DOI 10.1007/s10654-016-0167-1

CLINICAL EPIDEMIOLOGY

Oral glucocorticoid therapy and all-cause and cause-specific mortality in patients with rheumatoid arthritis: a retrospective cohort study

Mohammad Movahedi[1] · Ruth Costello[1] · Mark Lunt[1] · Stephen Richard Pye[1] · Jamie Christopher Sergeant[1,2] · William Gregory Dixon[1,3]

Received: 13 February 2016 / Accepted: 22 May 2016 / Published online: 2 June 2016

Abstract Previous studies of glucocorticoid (GC) therapy and mortality have had inconsistent results and have not considered possible perimortal bias—a type of protopathic bias where illness in the latter stages of life influences GC exposure, and might affect the observed relationship between GC use and death. This study aimed to investigate all-cause and cause-specific mortality in association with GC therapy in patients with rheumatoid arthritis (RA), and explore possible perimortal bias. A retrospective cohort study using the primary care electronic medical records. Oral GC exposure was identified from prescriptions. Mortality data were obtained from the UK Office for National Statistics. Multivariable Cox proportional hazards regression models assessed the association between GC use models and death. Several methods to explore perimortal bias were examined. The cohort included 16,762 patients. For ever GC use there was an adjusted hazard ratio for all-cause mortality of 1.97 (95 % CI 1.81–2.15). Current GC dose of below 5 mg per day (prednisolone equivalent dose) was not associated with an increased risk of death, but a dose–response association was seen for higher dose categories. The association between ever GC use and all-cause mortality was partly explained by perimortal bias. GC therapy was associated with an increased risk of mortality for all specific causes considered, albeit to a lesser extent for cardiovascular causes. GC use was associated with an increased risk of death in RA, at least partially explained by perimortal bias. Importantly, GC doses below 5 mg were not associated with an increased risk of death.

Keywords Rheumatoid arthritis · Glucocorticoids · Mortality · Steroids

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease which affects between 0.5 and 1 % of the adult population worldwide [1–3]. Oral glucocorticoid (GC) therapy was introduced as a treatment for patients with RA nearly 60 years ago [4] and is still used widely. Around one third of patients with RA are current users, and two thirds of patients have ever used GCs [5]. GCs improve symptoms of active RA through reducing joint pain, swelling and stiffness [6]. However, there are some concerns about their potential side effects including cardiovascular (CV) events, diabetes, infection, fracture, and cataracts [7–11], many of which are associated with an increased risk of mortality.

Previous studies have investigated the association between GC therapy and mortality, mostly focusing on all-cause mortality, though some have investigated CV mortality [5, 12–16]. Findings from these studies are not consistent. GCs have been associated with an increased risk of

Electronic supplementary material The online version of this article (doi:10.1007/s10654-016-0167-1) contains supplementary material, which is available to authorized users.

✉ William Gregory Dixon
Will.Dixon@manchester.ac.uk

[1] Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute for Inflammation and Repair, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK

[2] NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

[3] Health eResearch Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK

all-cause mortality in some [12–15, 17, 18] but not all studies [16, 19, 20], with similar inconsistency for CV mortality [12, 14, 16]. Very few studies have examined other cause-specific mortality. In studies that consider dose, some have suggested no association with doses <5 mg prednisolone equivalent [12, 13], reflecting either a lack of significant side effects at this dose or perhaps a favourable balance between side effects and positive anti-inflammatory properties.

There are important methodological issues when considering GC exposure and mortality, including confounding by indication—whereby GC therapy is given to patients with high disease severity and high disease severity is itself associated with increased mortality. However, studies have rarely considered a form of protopathic bias we will call 'perimortal bias', where illness in the latter stages of life influences GC exposure, and which consequently might affect the observed relationship between GC use and death. For example, if a patient were to develop cancer, GC therapy may be prescribed to treat the malignancy and a resultant association would be observed between GCs and (cancer-specific) mortality. The aim of this study was to investigate all-cause and cause-specific mortality in association with various models of oral GC exposure in patients with RA, and to explore and control for the possible existence of perimortal bias.

Methods

Database

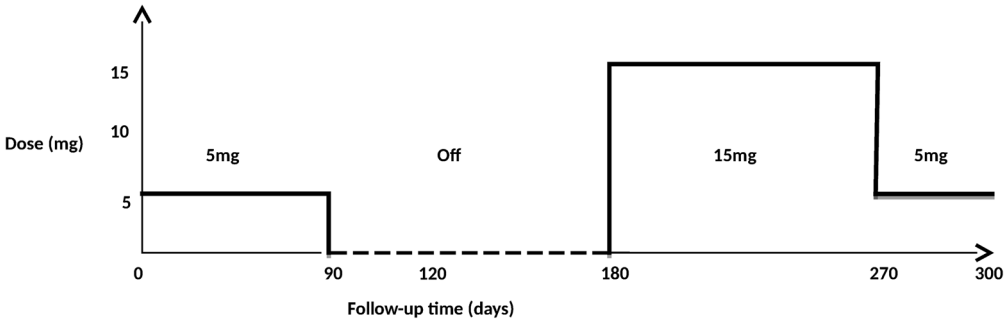
The Clinical Practice Research Datalink (CPRD) is a database of anonymised UK primary care electronic medical records covering 9 % of the population. There are 650 General Practitioner (GP) practices who contribute high-quality data, with over 5.5 million active patients who are broadly representative of the UK population [21, 22]. Information on the database includes patient demographics, medical diagnoses, clinical tests, hospital referrals, and drug prescriptions. Diagnoses on CPRD have been shown to have a high validity [23]. Selected practices consent to linkage to mortality data for England and Wales from the UK Office for National Statistics (ONS), and to Hospital Episode Statistics (HES), which provides information on hospital admissions.

Study population

Patients with RA were identified in the CPRD database using a validated algorithm [24]. To satisfy the algorithm patients needed either: more than one RA Read code, a seropositive/erosive RA or "rheumatoid arthritis" code (such as RA of knee), and no code for an alternative diagnosis after the last RA code; or a DMARD prescription with no Read code for an alternative indication in the 5 years prior to the first DMARD prescription. A study window of 1st January 1998–1st October 2011 was used. The cohort was restricted to the 340 GP practices eligible to be linked to ONS mortality and HES data, with data restricted to the period of mortality data linkage for each GP practice to ensure accurate vital status information. Patients entered the study on the latest of first RA code, date of ONS linkage or 01/01/1998. Patients under 16 years of age were excluded. The population was also restricted to patients with at least 1 year's information in CPRD prior to cohort entry, to allow assessment of prior GC exposure. Follow-up ended at transfer out of GP practice, GP practice data last collection, death, or 01/10/2011, whichever came first.

Exposure definition

The dose and duration of each GC prescription was derived from the available prescription information using a pre-specified algorithm (see Online Resource item A1). Doses of oral GCs were converted into a prednisolone-equivalent dosage (PED). Time-varying GC exposure was then defined in six ways: (1) ever use: a patient was considered a never user until the point of their first GC prescription when they became an ever user. This was the primary analyses. (2) Current use: a patient was considered a current user during their GC prescription and became a non-user during the periods without a GC prescription. (3) Current dose (5 mg/day): during a patient's GC prescription this was the dose divided by 5, during non-use this was zero. (4) Current dose category: a patient's current dose was categorised into the following categories: non-use, >0–4.9, 5–7.4, 7.5–14.9, 15–24.9 and 25 mg and over PED/day. (5) Cumulative dose since cohort entry (1000 mg/day): a patient's cumulative dose was calculated by summing the doses that had been prescribed up to that point and dividing by 1000, during non-use the cumulative dose would remain at the cumulative dose of prescriptions up to that point. (6) Cumulative dose category: a patent's cumulative dose was categorised into the following categories: non-use, >0–959, 960–3054, 3055–7299 and 7300 mg and over PED/day. An example of a patient's changing GC status through time is shown in Fig. 1. As time in hospital creates a gap in primary care records, because patients cannot attend the primary care practice, the GC exposure was set to the latest GC status prior to admission for the duration of any hospital inpatient stay identified using HES data.



GC exposure type	0-90 days	90-180 days	180-270 days	270-300 days
Ever GC status	Yes	Yes	Yes	Yes
Current status	Yes	No	Yes	Yes
Current GC dose (mg)	5	0	15	5
Current dose, category (mg)	5-7.4	0	15-24.9	5-7.4
Cumulative dose at end of each episode (mg)	450	450	1800	1950
Cumulative dose category (mg)	0-959.9	0-959.9	960-3054.9	960-3054.9

Fig. 1 Example of GC exposure definitions during follow-up for a hypothetical patient

Death ascertainment

The ONS defines cause of death by International Classification of Diseases version 2010 (ICD 10) codes with a specified underlying cause of death. We examined the underlying cause of death by the most frequent ICD 10 chapter headings of circulatory (ICD chapter I), neoplasms (ICD chapters C and D), respiratory diseases (ICD chapter J), and the remaining chapter headings grouped together in an "other causes" category. We also identified the leading causes of death in each chapter. Causes of death prior to 2001 were coded using ICD 9 and were later mapped to ICD 10.

Confounders

The following a priori potential confounders were included in the analyses: gender, age, body mass index (BMI), smoking status, socioeconomic status (SES) (Townsend quintile), prior 1 year cumulative GC dose at baseline, baseline Charlson comorbidity index [25], time-varying use of the DMARDs methotrexate, hydroxychloroquine, sulfasalazine and leflunomide and use of other DMARDs (penicillamine, azathioprine, cyclosporin, injectable gold) and time-varying use of non-steroidal anti-inflammatory drugs (NSAIDs) during follow-up. For a subgroup of the cohort who had the information available, the mean number of rheumatology outpatient visits per year and the mean number of GP visits per year was calculated and additionally adjusted for in a sensitivity analysis.

Statistical analysis

Baseline characteristics were tabulated for the whole cohort, and stratified by ever use at the end of follow-up, to examine if there were any differences between ever users and never users.

Mortality rates with 95 % confidence intervals (CI) were estimated by dividing the number of deaths by the total number of person-years follow-up.

Primary analyses examined the association between GC exposure and time until death, using Cox proportional hazards regression [26], using the six GC exposure definitions described above. Associations between GC exposure and mortality (both all-cause and cause-specific) were estimated through crude, and fully adjusted hazard ratios (HR), with 95 % CI. The proportional hazards assumption was checked by testing the Schoenfeld residuals. The association between oral GC use and cause-specific mortality was further explored using the Fine and Gray competing risks approach [27]. All data analysis was performed using Stata/MP Version 12.1 (StataCorp, Texas).

Missing data

The proportion of missing data for all confounders was assessed. If there was more than 5 % missing data the

variable was included in a fully adjusted model 1, and assessed in a complete case analysis. Any variable that was significantly associated with the outcome or changed the hazard ratio for the primary exposure by at least 10 % was included in the analyses, and therefore imputed. Other variables were excluded from the analysis. If there was <5 % missing data the variable was included in the model, and only complete cases were included in the analyses.

Exploring potential perimortal bias

Possible perimortal bias was explored in three ways. First, in order to explore whether GC therapy was being initiated in response to a terminal illness such as cancer, the distribution of cause-specific deaths in the first 6 months after GC initiation was compared to the distribution of cause-specific deaths more than 6 months after GC initiation in ever GC users. Second, the proportion of deaths was compared among two groups: (1) GC users who had oral GC therapy less than 6 months before death; and (2) GC users who had oral GC therapy more than 6 months before death but no GC use in the 6 months prior to death. Third, exposure during a 6 month period before death was excluded from the analyses to see if this had an impact on the results [28]. The same GC exposure models were used, although now based on the GC status at 6 months prior to death (see Figure A1 in Additional file 2).

The protocol for this study has been approved by Independent Scientific Advisory Committee for Medicines and Healthcare Regulatory Agency database research (Protocol number: 11_113RA4). As this study used routinely collected anonymised electronic health records consent was not required.

Results

There were 37,983 patients identified with a diagnosis of RA, of whom 21,355 were eligible for ONS linkage. After applying the exclusion criteria, the cohort reduced to 16,762 patients (Fig. 2). Table 1 summarizes the patient characteristics of the whole cohort, ever GC users and non-users. 70 % of patients were female, with similar proportions in the GC and non-GC groups. Mean age at baseline was 60.2 years [standard deviation (SD): 14.6].

There were 8367 (50 %) patients who received at least one prescription for oral GCs. These patients were on average 4 years older, more likely to have received GCs in the 1 year prior to RA (44 vs. 6 %, respectively) and had higher DMARD use during follow-up compared to non-users. The mean baseline Charlson comorbidity index was slightly higher in ever users compared with never users (1.39 vs. 1.25) (Table 1).

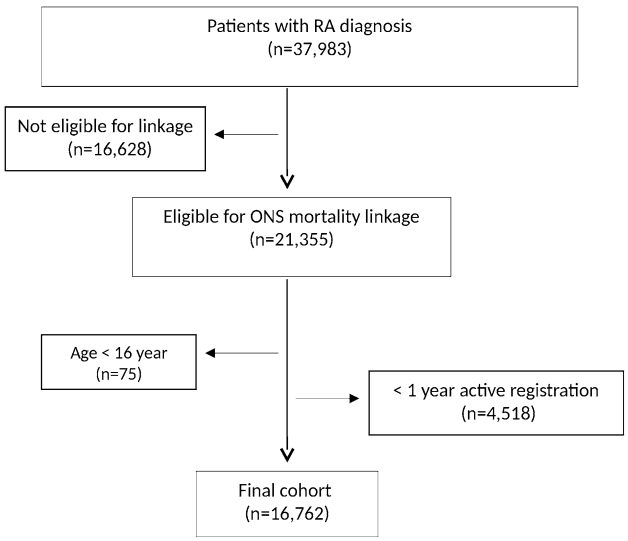


Fig. 2 Flowchart of the ONS linked patient cohort

During active GC prescriptions, the mean current daily dose (PED) was 7.5 mg (SD 6.9 mg). The mean cumulative dose (PED) among the 8367 patients who received GC therapy was 5.3 g (SD 6.0 g).

During a total of 111,099 person years, 2996 patients died (median follow-up of 6.1 years per person), giving an all-cause mortality rate of 27.0 deaths per 1000 person-years (pyrs) (95 % CI 26.0–28.0) (Table 2). In those never exposed to GCs the mortality rate was 15.5 deaths per 1000 pyrs, compared to 44.0 deaths per 1000 pyrs in those ever exposed to GCs.

Overall the most common cause of death was cardiovascular disease, followed by neoplasms and respiratory diseases. The underlying causes of death in the "other causes" category were mostly musculoskeletal (28.0 %). Figure 3 shows the cumulative incidence curves from Fine-Gray models [27] for the four categories of cause-specific mortality. Ever users had higher mortality rates in each cause-specific category compared to never users. For each category the mortality rate for ever users was higher than never users from the start of follow-up, and the mortality rate was consistent through follow-up for both exposed and unexposed groups.

Cardiovascular mortality rates were 15.8 deaths per 1000 pyrs in ever users compared to 6.4 deaths per 1000 pyrs in never users. Within this chapter, ischemic heart disease had the highest mortality rate for both ever and never users. Neoplasms had the second highest mortality rate for ever GC users. Conversely, the second highest mortality rate for never users was other causes of death. Respiratory diseases had the lowest mortality rate in both ever and never GC users (Table 2).

Table 1 Characteristics of the cohort, stratified by oral GC therapy status during follow-up

	All subjects	Never users	Ever users
Number of patients, n (%)	16,762	8395 (50.1)	8367 (49.9)
Follow-up time, total (person-years)	111,099	66,560	44,538
Females, n (%)	11,748 (70.1)	5945 (70.8)	5803 (69.4)
Age at baseline (years), mean (SD)	60.2 (14.6)	58.2 (14.9)	62.1 (14.1)
Body Mass Index at baseline			
Mean (SD)	26.8 (5.6)	26.9 (5.5)	26.8 (5.71)
Missing (%)	2763 (16.5)	1483 (17.7)	1280 (15.3)
Smoking status at baseline, n (%)			
Non smoker	7832 (46.7)	4115 (49.0)	3717 (44.4)
Ex smoker	3192 (19.0)	1489 (17.7)	1703 (20.4)
Current smoker	5227 (31.2)	2525 (30.1)	2702 (32.3)
Missing	511 (3.1)	266 (3.17)	245 (2.93)
Socioeconomic status quintile at baseline, n (%)			
First (least deprived)	3672 (21.9)	1871 (22.3)	1801 (21.5)
Second	4040 (24.1)	2031 (24.2)	2009 (24.5)
Third	3566 (21.3)	1746 (20.8)	1820 (21.8)
Fourth	3213 (19.2)	1601 (19.1)	1612 (19.3)
Fifth (most deprived)	2204 (13.2)	1112 (13.3)	1092 (13.1)
Missing	67 (0.4)	34 (0.4)	33 (0.4)
Prior history of GC use, n (%)	4138 (24.7)	484 (5.80)	3661 (43.8)
Charlson comorbidity index at baseline, mean (SD)	1.32 (0.70)	1.25 (0.64)	1.39 (0.76)
Methotroxate ever during follow-up, n (%)	8949 (53.4)	4020 (47.9)	4929 (58.9)
Hydroxycholoroquine ever during follow-up, n (%)	3728 (22.2)	1726 (20.6)	2002 (23.9)
Sulfasalazine ever during follow-up, n (%)	4793 (28.6)	2249 (26.8)	2544 (30.4)
Leflunomide ever during follow-up, n (%)	1465 (8.74)	455 (5.42)	1010 (12.1)
Other DMARDs ever during follow-up, n (%)[a]	4304 (25.7)	1683 (20.1)	2621 (31.3)

[a] Other DMARDS: penicillamine, azathioprine, cyclosporin, injectable gold

Table 3 shows the associations between oral GC use and risk of all-cause and cause-specific mortality, estimated through six alternative Cox models, adjusted for age, gender, smoking status, SES, prior cumulative dose of GC, baseline Charlson comorbidity index, time-varying NSAID use and time-varying DMARD use.

BMI was the only potential confounder with higher than five percent of missing data (Table 1). When it was included in a complete case analysis of model 1 it did not alter the hazard ratio for GC use and was not significantly associated with mortality and so was not included in the fully adjusted models. Smoking and SES had <5 % missing data and were included in the fully adjusted models. All models consistently showed that risk of death was associated with GC use and increased with higher dosages of GCs. There was a nearly twofold greater risk of all-cause mortality in ever users, compared to never users (HR 1.97, 95 % CI 1.81–2.15). For cause-specific mortality, ever users had over a three times higher risk of death from neoplasms compared to never users (HR 3.20, 95 % CI 2.66–3.86). For both all-cause and cause-specific mortality, a similar pattern was seen for current use, though the point estimates were lower. For each 5 mg increase in GC dose there was a 33 % increased risk of all-cause mortality compared to non-users (HR 1.33, 95 % CI 1.30–1.35). Similar increased risks were seen for each of the cause specific mortality categories, with the highest risk seen for neoplasms (HR 1.46, 95 % CI 1.42–1.49).

The categorisation of current daily dose showed that for all-cause mortality, CV mortality and mortality due to respiratory diseases, a dose below 5 mg per day was not associated with an increased risk of death. Furthermore, for neoplasms and 'other causes', a dose of below 7.5 mg per day was not associated with an increased risk of death. However, as current daily dose increased above these doses, so too did the risk of death. Comparing between the hazard ratios for cause-specific mortality, the risk of cardiovascular mortality was notably lower than for the other three categories of death, for current GC dose above 7.5 mg.

There was a 6 % increased risk of all-cause mortality for each 1000 mg/day increase in cumulative dose since cohort

Table 2 Underlying causes of death and crude mortality rates, overall and by ever GC use status

	All subjects		Never GC use[b]		Ever GC use	
	Events (%)	Mortality rate[a]	Events (%)	Mortality rate[a]	Events (%)	Mortality rate[a]
1 All-causes	2996	27.0 (26.0–28.0)	1034	15.5 (14.6–16.5)	1962	44.0 (42.1–46.0)
2 Cardiovascular diseases	1131 (100)	10.2 (9.60–10.8)	428 (100)	6.40 (5.84–7.07)	703 (100)	15.8 (14.7–17.0)
Ischemic heart diseases	581 (51.4)	5.23 (4.82–5.67)	207 (48.4)	3.11 (2.61–3.37)	374 (53.2)	8.39 (7.59–9.29)
Cerebrovascular diseases	247 (21.8)	2.22 (1.96–2.52)	121 (28.3)	1.82 (1.52–2.17)	126 (17.9)	2.83 (2.37–3.37)
Others	303 (26.8)	2.73 (2.44–3.05)	100 (23.3)	1.50 (1.24–1.83)	203 (28.9)	4.56 (3.97–5.23)
3 Neoplasms	639 (100)	5.75 (5.32–6.22)	191 (100)	2.87 (2.49–3.31)	448 (100)	10.1 (9.17–11.0)
Respiratory neoplasm	208 (32.6)	1.87 (1.63–2.14)	41 (21.5)	0.62 (0.45–0.84)	167 (37.3)	3.75 (3.22–4.36)
Digestive neoplasm	135 (21.1)	1.22 (1.03–1.44)	46 (24.1)	0.69 (0.52–0.92)	89 (19.9)	2.00 (1.62–2.46)
Others	296 (46.3)	2.66 (2.38–2.99)	104 (54.4)	1.56 (1.29–1.89)	192 (42.8)	4.31 (3.74–4.97)
4 Respiratory diseases	509 (100)	4.58 (4.20–5.00)	132 (100)	1.98 (1.67–2.35)	377 (100)	8.46 (7.65–9.36)
Respiratory infection	216 (42.4)	1.94 (1.70–2.22)	80 (60.6)	1.20 (0.97–1.50)	136 (36.1)	3.05 (2.58–3.61)
Lower respiratory diseases	205 (40.3)	1.85 (1.61–2.12)	32 (24.2)	0.48 (0.34–0.68)	173 (45.9)	3.88 (3.35–4.51)
Others	88 (17.3)	0.79 (0.64–0.98)	20 (15.2)	0.30 (0.19–0.47)	68 (18.0)	1.53 (1.20–1.94)
5 Others causes of death	717 (100)	6.45 (6.00–6.94)	283 (100)	4.25 (3.78–4.77)	434 (100)	9.74 (8.87–10.7)
Musculoskeletal diseases	201 (28.0)	1.81 (1.58–2.08)	67 (23.7)	1.00 (0.79–1.28)	134 (30.9)	3.01 (2.54–3.56)
Digestive diseases	158 (22.0)	1.42 (1.22–1.66)	68 (24.0)	1.02 (0.81–1.29)	90 (20.7)	1.64 (1.79–2.48)
Genitourinary diseases	75 (10.5)	0.68 (0.54–0.85)	22 (7.8)	0.33 (0.22–0.50)	53 (12.2)	1.19 (0.91–1.56)
Injury, poisoning and external causes	102 (14.2)	0.92 (0.76–1.11)	47 (16.6)	0.71 (0.53–0.93)	55 (12.7)	1.23 (0.95–1.61)
Others	181 (25.3)	1.63 (1.41–1.88)	79 (27.9)	1.19 (0.95–1.48)	102 (23.5)	2.29 (1.89–2.78)

[a] Mortality rates per 1000 patient-years

[b] Patients who had not yet used GCs could initially contribute person time to the 'never GC use' group, and then switch to 'ever GC use' person time on receipt of their first GC prescription

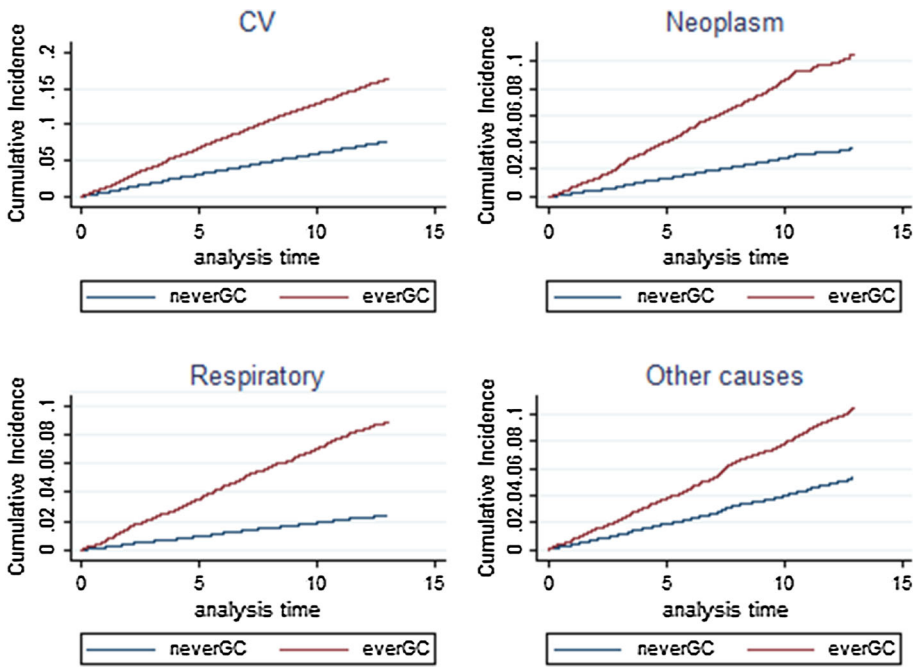


Fig. 3 Cumulative incidence curves by GC status

Table 3 Association between oral GC use and all-cause and cause-specific mortality (n = 16,187)

Model	Oral GC pattern	Adjusted hazard ratio (aHR) with 95 % CI[a]				
		All-cause mortality death = 2770	CV mortality death = 1039	Neoplasms death = 606	Respiratory diseases death = 468	Other causes death = 657
1	Ever use, (ref = never use)	1.97 (1.81–2.15)	1.66 (1.45–1.91)	3.20 (2.66–3.86)	2.64 (2.11–3.31)	1.39 (1.16–1.66)
2	Current use, (ref = non-use)	1.77 (1.62–1.93)	1.58 (1.37–1.83)	2.22 (1.84–2.68)	1.92 (1.57–2.36)	1.69 (1.41–2.02)
3	Current dose per 5 mg/day	1.33 (1.30–1.35)	1.21 (1.16–1.27)	1.46 (1.42–1.49)	1.36 (1.30–1.41)	1.25 (1.20–1.31)
4	Current dose category, (ref = non-use)					
	>0–4.9 mg	1.02 (0.87–1.20)	1.10 (0.85–1.41)	0.79 (0.51–1.22)	0.87 (0.57–1.33)	1.15 (0.85–1.57)
	5.0–7.4 mg	1.44 (1.26–1.64)	1.59 (1.31–1.94)	1.07 (0.75–1.52)	1.74 (1.30–2.32)	1.23 (0.93–1.63)
	7.5–14.9 mg	2.24 (1.98–2.54)	1.96 (1.59–2.42)	2.34 (1.75–3.13)	2.19 (1.62–2.97)	2.66 (2.09–3.38)
	15.0–24.9 mg	4.50 (3.61–5.62)	2.79 (1.80–4.31)	8.07 (5.41–12.0)	8.03 (5.31–12.2)	2.06 (1.09–3.90)
	≥25 mg	11.0 (8.87–13.6)	2.48 (1.23–4.99)	31.3 (23.5–41.9)	11.4 (6.84–19.0)	6.87 (4.01–11.8)
5	Cumulative dose since cohort entry (1000 mg/day)	1.06 (1.05–1.07)	1.05 (1.04–1.07)	1.06 (1.04–1.08)	1.07 (1.05–1.09)	1.07 (1.05–1.08)
6	Cumulative dose category (ref = non-use)					
	>0–959.9 mg	1.60 (1.42–1.81)	1.41 (1.16–1.72)	2.51 (1.97–3.21)	2.18 (1.61–2.95)	1.04 (0.79–1.36)
	960–3054.9 mg	1.83 (1.62–2.07)	1.38 (1.12–1.70)	3.84 (3.04–4.87)	2.24 (1.64–3.05)	1.16 (0.88–1.52)
	3055–7299.9 mg	2.11 (1.87–2.39)	1.91 (1.57–2.32)	3.31 (2.55–4.30)	2.65 (1.95–3.61)	1.48 (1.15–1.92)
	≥7300 mg	3.11 (2.74–3.52)	2.59 (2.11–3.18)	3.85 (2.90–5.10)	4.85 (3.59–6.55)	2.54 (1.98–3.25)

[a] Adjusted for gender, age, smoking status, SES, prior cumulative dose of GC, Charlson comorbidity index at baseline, time-varying NSAID use and time-varying DMARD use

entry (HR 1.06, 95 % CI 1.05–1.07). Similar increases in risk were seen for each cause-specific mortality category. Categorisation of cumulative dose showed a dose response increased risk of all-cause mortality in each category of cumulative dose, with risk of death increasing with increased categories of cumulative dose. The exception to this was for other causes of death, where there was not an increased risk of death from other causes with cumulative doses up to 3054.9 mg (Table 3). Additional adjustment for mean number of rheumatology outpatient visits per year and mean number of GP visits per year in general increased the risk of all-cause mortality and cause-specific mortality, but did not alter the significance, except for the lowest current dose category (0-4.9 mg) where a significantly reduced risk of mortality due to neoplasms was seen (Online resource Table A1).

Perimortal bias

The mortality rate in the first 6 months following GC therapy initiation was 56.5 deaths per 1000 pyrs, compared to 42.8 deaths per 1000 pyrs beyond 6 months after GC initiation. The rate of neoplasm deaths was higher in patients in the first 6 months following GC initiation (23.5 per 1000 pyrs compared to 8.7 per 1000 pyrs beyond 6 months) (Online Resource Table A2).

Of those who died (N = 2996), 1962 patients ever used GCs. Of these, 1576 patients used GCs during the 6 months prior to death and 368 last used GCs more than 6 months prior to death. Those who used GC in the 6 months prior to death had a higher proportion of deaths due to respiratory, neoplasms and other causes, but a lower proportion of CV deaths, compared to those patients who received GC therapy more than 6 months prior to death. For example 23.4 % of those who used GCs during the 6 months prior to death died from neoplasms, compared to 20.7 % in those who used GCs more than 6 months prior to death (Online Resource Table A3).

After the exclusion of GC information in the 6 months prior to death, the association between ever use and all-cause mortality was reduced but remained statistically significant (HR 1.64, 95 % CI 1.50–1.79). A similar reduction in hazard ratio was seen for cause-specific mortality, in particular neoplasm mortality where ever users had only a 76 % increased risk of death from neoplasms (HR 1.76, 95 % CI 1.47–2.10), compared to a threefold greater risk when the 6 months prior to death was included (HR 3.20, 95 % CI 2.66–3.86). In Model 4, the magnitude of risk was reduced for the highest dose category of >25 mg PED for all-cause and each cause-specific mortality. Excluding the exposure data from 6 months prior to death had the biggest impact on deaths caused by

neoplasms, with hazard ratios falling from 8.07 (95 % CI 5.41–12.0) to 3.42 (95 % CI 1.87–6.28) for 15–25 mg, and from 31.3 (95 % CI 23.5–41.9) to 5.66 (95 % CI 2.80–11.4) for >25 mg. Full results for models 1-6 following exclusion of GC information in the 6 months prior to detail are shown in Online Resource Table A4.

Unmeasured confounding

The cause-specific analyses found an association between oral GC use and death from other causes, supporting the possibility of unmeasured confounding. To explore this, a post hoc sensitivity analysis was conducted using the rule out approach [29, 30]. This approach finds the minimum effect an unmeasured confounder would need to have to remove statistical significance. It was found that an unmeasured confounding factor with 40 % prevalence would have to increase the relative risk of mortality by a factor 3 and at the same time increase the odds of GC exposure by a factor of 3.5 in order to fully remove the association found between ever use and mortality risk due to other causes (HR 1.39, 95 % CI 1.16–1.66). For each of the other causes of death the unmeasured confounders would need to increase the relative risk of mortality and the odds of GC exposure by too large an amount for them to explain the result fully. For example, an unmeasured confounding factor for CV mortality would have to increase the relative risk of CV mortality by a factor of 3 and increase the odds of GC exposure by a factor of 7.7 in order to remove the association found, which seems unlikely. Similarly, an unmeasured confounder with increased risk of death by a factor below 3 cannot plausibly explain the observed association between GC exposure and CV mortality.

Discussion

This study examined the association between oral GC therapy and mortality rates in a cohort of patients with RA in the UK. Ever GC use and current GC use was associated with an increased risk of all-cause mortality and cause-specific mortality, with a largely consistent dose–response effect. An increase in current dose of 5 mg per day was associated with an increased risk of death, however categorisation showed that taking <5 mg per day at the time of death did not increase the risk of all-cause mortality or cause-specific mortality, and taking <7.5 mg per day at the time of death did not increase this risk of death from neoplasms or other non-CV and non-respiratory causes. In addition, moderate to high doses of GC therapy were associated with a lesser risk of CV deaths compared to neoplasm, respiratory and other causes of death, which might suggest GC therapy has a less harmful effect on CV mortality.

The study showed that perimortal bias partially explained some of the results, especially at higher doses. Perimortal bias is important to consider for a number of reasons. GCs can be used to treat diseases that might develop through the course of follow-up, and where that disease is the leading cause of death. For example, if a patient were to develop a malignancy, they might start GC therapy as part of their cancer treatment which would lead to a positive association between GCs and (cancer-related) mortality. Similarly, end of life care might lead to a switch from disease-modifying anti-rheumatic drug (DMARD) therapy (that requires regular blood monitoring) to GC therapy, again generating an association between GC use and death.

When GC use in the 6 months prior to death was removed, the association between ever GC use and all-cause mortality remained significant, but the risks were reduced. This was mainly influenced by the large reduction in risk of death from neoplasms, where there is a clear possibility of perimortal bias: GCs are prescribed as a treatment for cancer [31]. Initial signals of possible perimortal bias were evident in the magnitude of the association between high-dose GCs and risk of death due to neoplasm (HR 31.3, 95 %CI 23.5–41.9).

The all-cause mortality rate for this study was 27 deaths per 1000 person-years, and the cardiovascular mortality rate was 10 deaths per 1000 person-years. This was higher than a recent cohort study in the UK (Norfolk Arthritis Registry (NOAR)) [32] where rates were 20–21 per 1000 person-years and 7–8 per 1000 person years for all-cause mortality and cardiovascular mortality respectively. This would be expected as NOAR includes patients with early inflammatory arthritis, whereas this study included patients with a higher baseline age and with RA only, and therefore more severe disease.

Our findings are in agreement with some previous studies [5, 12–15, 17] which have investigated all-cause mortality or CV mortality in association with GC use. Caplan et al. [5] found an increased risk of death with current GC use, with an adjusted odds ratio of 2.2 (95 % CI 1.9–2.7) and an increased risk of death with increasing duration of GC treatment. del Rincon et al. [12] found a GC dose-dependent increase in death from all causes (HR 1.07 per 1 mg/day (95 % CI 1.05–1.08) and CV cause with a similar point estimate. They also showed that there was a dose response association for cumulative dose for all-cause and CV mortality with a threshold of 40 g. Listing et al. [13] showed that GC doses higher than 5.0 mg per day were significantly associated with increased all-cause mortality, independent of disease activity. Treatment with prednisolone higher than 15 mg per day was associated

with 3.4-fold (95 % CI 2.01–5.86) increased risk of all-cause mortality compared with non-use. Our findings of probable perimortal bias, however, might suggest that the hazard ratios reported in these previous studies are overestimates of the true effect.

An important finding of this study was the absence of an association between both all-cause and cause-specific mortality and GC doses lower than 5 mg per day. This may reflect either a low risk of adverse events at this dose, or at least a favourable balance between the harms and the biologically plausible benefits through their anti-inflammatory effects [33]. This finding replicates similar findings from Listing et al. [13] and del Rincon et al. [12], which showed that doses lower than 5 and 8 mg PED, respectively, had no association with mortality risk.

The strengths of this study are firstly its size, with nearly 3000 deaths in 16,762 patients. This meant the study had greater power to detect differences in mortality rates and allowed us to explore cause-specific mortality. We were thus able to see an increased rate of respiratory deaths, accepting the possibility of perimortal bias but also likely driven by a causal increased risk of respiratory infection [9]. Second, the study used linkage to the national mortality register, providing robust and complete information on cause of death for all patients in the study. Third, time-varying covariates for DMARDs, NSAIDs and GCs, were used to allow more accurate estimation compared to time-independent variables for these drugs. Fourth, a range of patterns of GC use were explored including GC use, GC daily dose, and cumulative dose since cohort entry and their categorical variables compared with non GC use in association with risk of death. This approach allowed some consideration of the impact of dose, duration and timing of treatment on mortality risk. For example, the finding that the highest current dose category was associated with very high HRs for neoplasm, respiratory and other causes of death, whilst the highest quartile had notably lower HRs, suggests that high doses may be used at the end of life when cumulative exposure is less of an issue. We also explored possible perimortal bias which has not been considered in previous studies. Moreover, we examined oral GC therapy in association with cause-specific mortality beyond CV mortality which has not been investigated in previous studies.

There were some limitations with the study. The prescription data from the CPRD dataset are reliable in terms of drugs prescribed, but does not cover drugs prescribed in secondary care only, such as biologic DMARDs, or over the counter use of NSAIDs; although it has been shown that biologic DMARDs are not associated with an increased mortality compared to standard DMARDs [34]. In addition there may have been some exposure misclassification because of assumptions in data preparation, missing data, patient adherence, injectable steroids and hospital administered GC, although the latter is likely to be minimal as UK rheumatologists typically make recommendations for oral GC treatment to GPs. Like all observational studies, the impact of confounding and bias needed consideration. A range of possible confounders were adjusted for, including time-varying exposure to DMARDs and NSAIDS, and healthcare utilisation variables as surrogate measures of RA disease severity. Although we didn't have direct measures of disease severity, previous studies that did adjust for clinician-reported disease severity found a persistent association between GC use and mortality [13]. It is thus likely that there was some residual confounding by disease severity. In terms of possible residual or unmeasured confounders affecting the results, of which disease severity is one, sensitivity analyses showed that these would need to be very large to fully explain the results. So even though, for example, high cumulative disease severity has been shown to be associated with lymphoma [35] this would not fully explain the results seen. Adjusting for the Charlson comorbidity index at baseline was expected to control for key comorbidities that contribute to an increased risk of mortality. The main difference at baseline between GC users and non-GC users was prior GC use, which was much higher in GC users (44 vs. 6 %). It may have been that this group was particularly susceptible to death, and any association between GC use and death may have been exaggerated. However, prior GC use was adjusted for so the results should not be biased.

It is very challenging to understand the true causal relationship between oral GC use and mortality from an observational study due to the complex relationships between the indication for treatment (that changes through time) and the outcome, as well as the granularity of the data from a population necessarily large to support the analysis. Nonetheless, despite this blurring of causality by bias and confounding, some important messages emerge. Doses <5 mg were not associated with an increased risk of death. This absent risk is not explained by confounding by disease severity (where you would expect mortality to be higher in the treated compared to non-treated), or by perimortal bias where you would again expect an increased risk compared to non-use. The lower dose-specific hazard ratios for cardiovascular deaths compared to the hazard ratios seen for the other causes of death raises the interesting hypothesis that GC therapy might have a beneficial effect on cardiovascular mortality; yet a protective effect is impossible to conclude with certainty as there is a statistically significant increased risk for CV mortality with all doses above 5 mg. Disentangling these complex factors is impossible, but the large population observational research raises questions

that can feed back into more targeted studies, both basic science and epidemiological.

Conclusions

This study has found that GC use is associated with an increased risk of death in RA, both all-cause and cause-specific mortality, which is partially explained by peri-mortal bias. Importantly, doses of below 5 mg PED were not associated with an increased risk of death. There is a suggestion that GCs may have a less harmful effect on CV mortality compared to their association with other cause-specific mortality, but targeted research is required to examine this signal further.

Acknowledgments This work was supported by the Arthritis Research UK Centre for Epidemiology (20380). Dr Dixon was supported by a Medical Research Council Clinician Scientist Fellowship (G0902272).

Compliance with ethical standards

Conflict of interests The authors declare that they have no conflicts of interest.

Ethical approval and informed consent The protocol for this study has been approved by Independent Scientific Advisory Committee for Medicines and Healthcare Regulatory Agency database research (Protocol number: 11_113RA4). As this study used routinely collected anonymised electronic health records consent was not required.

References

1. Carmona L, Cross M, Williams B, Lassere M, March L. Rheumatoid arthritis. Best Pract Res Clin Rheumatol. 2010;24(6):733–45. doi:10.1016/j.berh.2010.10.001.
2. Neovius M, Simard JF, Askling J, Group As. Nationwide prevalence of rheumatoid arthritis and penetration of disease-modifying drugs in Sweden. Ann Rheum Dis. 2011;70(4):624–9. doi:10.1136/ard.2010.133371.
3. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford). 2002;41(7):793–800.
4. Hench PS, Kendall EC, Slocumb CH, Polley HF. Effects of cortisone acetate and pituitary ACTH on rheumatoid arthritis, rheumatic fever and certain other conditions—a study in clinical physiology. Arch Intern Med. 1950;85(4):545–666.
5. Caplan L, Wolfe F, Russell AS, Michaud K. Corticosteroid use in rheumatoid arthritis: prevalence, predictors, correlates, and outcomes. J Rheumatol. 2007;34(4):696–705.
6. Kirwan JR, Bijlsma JW, Boers M, Shea BJ. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. Cochrane Database Syst Rev. 2007;1:CD006356.
7. Davis JM 3rd, Maradit Kremers H, Crowson CS, Nicola PJ, Ballman KV, Therneau TM, et al. Glucocorticoids and cardiovascular events in rheumatoid arthritis: a population-based cohort study. Arthritis Rheum. 2007;56(3):820–30. doi:10.1002/art.22418.
8. Dessein PH, Joffe BI, Stanwix AE, Christian BF, Veller M. Glucocorticoids and insulin sensitivity in rheumatoid arthritis. J Rheumatol. 2004;31(5):867–74.
9. Dixon WG, Abrahamowicz M, Beauchamp ME, Ray DW, Bernatsky S, Suissa S, et al. Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case-control analysis. Ann Rheum Dis. 2012;. doi:10.1136/annrheumdis-2011-200702.
10. Van Staa TP, Leufkens HG, Abenhaim L, Zhang B, Cooper C. Oral corticosteroids and fracture risk: relationship to daily and cumulative doses. Rheumatology (Oxford). 2000;39(12):1383–9.
11. Miller DP, Watkins SE, Sampson T, Davis KJ. Long-term use of fluticasone propionate/salmeterol fixed-dose combination and incidence of cataracts and glaucoma among chronic obstructive pulmonary disease patients in the UK General Practice Research Database. Int J Chron Obstruct Pulmon Dis. 2011;6:467–76. doi:10.2147/COPD.S14247.
12. del Rincon I, Battafarano DF, Restrepo JF, Erikson JM, Escalante A. Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. Arthritis Rheumatol. 2014;66(2):264–72. doi:10.1002/art.38210.
13. Listing J, Kekow J, Manger B, Burmester GR, Pattloch D, Zink A, et al. Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNFalpha inhibitors and rituximab. Ann Rheum Dis. 2015;74(2):415–21. doi:10.1136/annrheumdis-2013-204021.
14. Sihvonen S, Korpela M, Mustonen J, Huhtala H, Karstila K, Pasternack A. Mortality in patients with rheumatoid arthritis treated with low-dose oral glucocorticoids. A population-based cohort study. J Rheumatol. 2006;33(9):1740–6.
15. Wolfe F, Mitchell DM, Sibley JT, Fries JF, Bloch DA, Williams CA, et al. The mortality of rheumatoid arthritis. Arthritis Rheum. 1994;37(4):481–94.
16. Wallberg-Jonsson S, Ohman ML, Dahlqvist SR. Cardiovascular morbidity and mortality in patients with seropositive rheumatoid arthritis in Northern Sweden. J Rheumatol. 1997;24(3):445–51.
17. Leigh JP, Fries JF. Mortality predictors among 263 patients with rheumatoid arthritis. J Rheumatol. 1991;18(9):1307–12.
18. Corbett M, Dalton S, Young A, Silman A, Shipley M. Factors predicting death, survival and functional outcome in a prospective study of early rheumatoid disease over fifteen years. Br J Rheumatol. 1993;32(8):717–23.
19. Mitchell DM, Spitz PW, Young DY, Bloch DA, McShane DJ, Fries JF. Survival, prognosis, and causes of death in rheumatoid arthritis. Arthritis Rheum. 1986;29(6):706–14.
20. Allebeck P, Rodvall Y, Allander E. Mortality in rheumatoid arthritis, particularly as regards drug use. Scand J Rheumatol. 1985;14(2):102–8.
21. CPRD. http//www.cprd.com. Last accessed Oct 2014.
22. Tate AR, Beloff N, Al-Radwan B, Wickson J, Puri S, Williams T, et al. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. JAMIA. 2014;21(2):292–8. doi:10.1136/amiajnl-2013-001847.
23. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. Br J Clin Pharmacol. 2010;69(1):4–14. doi:10.1111/j.1365-2125.2009.03537.x.

24. Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? Arthritis Rheum. 2008;59(9):1314–21.
25. Khan NF, Perera R, Harper S, Rose PW. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC Fam Pract. 2010;11:1. doi:10.1186/1471-2296-11-1.
26. Cox D. Regression models and life-tables. J R Stat Soc [SerB]. 1972;34(2):187–220.
27. Fine J, Gray R. A proportional hazards model for the subdistribution of competing riska in survival analysis. J Am Stat Assoc. 1999;94:496–509.
28. Tamim H, Monfared AA, LeLorier J. Application of lag-time into exposure definitions to control for protopathic bias. Pharmacoepidemiol Drug Saf. 2007;16(3):250–8. doi:10.1002/pds.1360.
29. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006;15(5):291–303. doi:10.1002/pds.1200.
30. Psaty BM, Koepsell TD, Lin D, Weiss NS, Siscovick DS, Rosendaal FR, et al. Assessment and control for confounding by indication in observational studies. J Am Geriatr Soc. 1999;47(6):749–54.
31. McKay L, Cidlowski JA. Corticosteriods in the treatment of neoplasms. Cancer Medicine. Hamilton, ON: BC Decker; 2003.
32. Humphreys JH, Warner A, Chipping J, Marshall T, Lunt M, Symmons DP, et al. Mortality trends in patients with early rheumatoid arthritis over 20 years: results from the Norfolk Arthritis Register. Arthritis Care Res. 2014;66(9):1296–301. doi:10.1002/acr.22296.
33. Pincus T, Sokka T, Stein CM. Are long-term very low doses of prednisone for patients with rheumatoid arthritis as helpful as high doses are harmful? Ann Intern Med. 2002;136(1):76–8.
34. Lunt M, Watson KD, Dixon WG, British Society for Rheumatology Biologics Register Control Centre C, Symmons DP, Hyrich KL, et al. No evidence of association between anti-tumor necrosis factor treatment and mortality in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Arthritis Rheum. 2010;62(11):3145–53. doi:10.1002/art.27660.
35. Baecklund E, Iliadou A, Askling J, Ekbom A, Backlin C, Granath F, et al. Association of chronic inflammation, not its treatment, with increased lymphoma risk in rheumatoid arthritis. Arthritis Rheum. 2006;54(3):692–701. doi:10.1002/art.21675.

Publication 4: The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study.

BMC Rheumatology

RESEARCH ARTICLE Open Access

The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study

Ruth E. Costello[1], Antonia Marsden[2], Mohammad Movahedi[1,3], Mark Lunt[1], Jenny H. Humphreys[1], Richard Emsley[4] and William G. Dixon[1*]

Abstract

Background: Patients with rheumatoid arthritis (RA) have increased cardiovascular (CV) and mortality risk. Patients with RA are also frequently prescribed glucocorticoids (GCs) which have been associated with increased risk of mortality. In addition, for patients who have concomitant diabetes mellitus (DM), GCs are known to worsen glycaemic control and hence may further increase CV and mortality risk. This study aimed to understand the relationship between GCs, DM and mortality in patients with RA.

Methods: This was a retrospective cohort study of patients with incident RA identified from UK primary care electronic medical records. Patients with linkage to Office for National Statistics (ONS) for mortality data ($N = 9085$) were included. DM was identified through Read codes, prescriptions and blood tests, and GC use was identified through prescriptions. Mortality rate ratios (RR) and rate differences (RD) were calculated across the different exposure groups. Cox proportional hazards regression models were used to estimate interaction on the multiplicative and additive scales.

Results: In those without DM GC use had a 4.4-fold increased all-cause mortality RR (95% confidence interval (CI): 3.77 to 5.07) compared to non-use, whilst those with DM had a lower RR for GC use (2.99 (95% CI: 2.32, 3.87)). However, those with DM had a higher RD associated with GC use because of their higher baseline risk. In those with DM, GC use was associated with an additional 44.9 deaths/1000 person-years (pyrs) (95% CI: 32.9 to 56.8) compared to non-use, while in those without DM GC use was associated with an additional 34.4 deaths/1000 pyrs (95% CI: 30.1 to 38.7) compared to non-use, while in those without DM GC use was associated with an additional 36.2 deaths/1000 pyrs (95% CI: 31.6 to 40.8). A similar pattern was seen for CV mortality. The adjusted Cox proportional hazards model showed no evidence of multiplicative interaction, but additive interaction indicated a non-significant increased risk. For CV mortality there was no interaction on either scale.

Conclusions: GC use was associated with higher mortality rates in people with comorbid DM compared to people without DM, despite apparently reassuring similar relative risks. Clinicians need to be aware of the higher baseline risk in patients with DM, and consider this when prescribing GCs in patients with RA and comorbid DM.

Keywords: Rheumatology, General diabetes, Epidemiology

* Correspondence: will.dixon@manchester.ac.uk
[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
Full list of author information is available at the end of the article

Background

Rheumatoid arthritis (RA) is an inflammatory disease that is thought to affect around 1% of the UK population [1] and is associated with a significantly higher rate of cardiovascular (CV) mortality compared to the general population [2]. Glucocorticoids (GC) have been widely used as a treatment for RA since their discovery in the 1950s [3] and continue to be used in around half of patients with RA [4]. Although GCs have many benefits, they also have risks associated with them, including possible increased risk of CV events and mortality [5, 6]. In addition, GCs are known to increase the risk of diabetes mellitus (DM) [7, 8] and are associated with poor glucose control [9], meaning they may also affect the long-term outcome of DM (including CV events and mortality) [10, 11]. This has not been investigated in patients with RA. Further, it is not known how the additional burden of DM and then GC therapy influence the cardiovascular and mortality risk in patients with RA. Therefore an important unanswered question is whether GC treatment in RA is associated with worse outcomes in patients with comorbid DM, compared to patients without DM.

As we think that the baseline risk of CV and all-cause mortality for patients with RA and DM will be higher than those with RA only, to investigate the impact of GCs it is appropriate to look at the absolute risks as well as the relative risks. The aims of this study were: 1) to compare the event rates for all-cause mortality and CV mortality, by GC use status and DM status, and 2) to examine whether DM modifies, on either the multiplicative or additive scales, the effect of GCs on all-cause mortality and CV mortality.

Methods

Setting

This was a retrospective cohort study using data from the Clinical Practice Research Datalink (CPRD) which was linked to mortality data from the Office of National Statistics (ONS). The CPRD is a large database of primary care electronic medical records that covers around 7% of the UK population and has been shown to be broadly representative of the UK population. Consenting practices in England have linkage to the ONS mortality data, which represents around 58% of all CPRD practices [12]. CPRD provide indicators of when a practice's data was up to research standard, and whether a patient's data meets their acceptability standards. For this study, only data from practices that consented to ONS linkage were used if the data met acceptability standards and was up to research standard.

Study population

The study period began at the start of ONS coverage (1st January 1998) and ended 1st October 2011. Patients with incident RA during the study period were identified from CPRD using a validated algorithm where patients have to have either at least 2 Read codes for RA and no alternative diagnosis after their last RA code or a Read code for RA and at least 2 product (medication) codes for Disease-Modifying Anti-Rheumatic Drugs (DMARDs) and no alternative diagnosis for the DMARDs in the previous 5 years [13]. Patients entered into the study upon RA diagnosis and participation ended at death, the date the patient left the practice or at the end of the study period. All patients were registered with the practice for a year prior to RA diagnosis, to ensure patients were truly incident cases.

Exposures

Patients were identified as having type 2 DM if they had either (1) a Read code for type 2 DM; (2) at least two prescriptions for oral anti-diabetic medication, either on 2 different dates or the same date with 2 types of medication; or (3) fasting blood sugar ≥7.0 mmol/litre, random glucose test ≥11.1 mmol/litre, glucose tolerance test ≥11.1 mmol/litre or a glycosylated haemoglobin (HbA1C) ≥7% [7]. Patients with polycystic ovary syndrome (PCOS) treated with metformin were excluded as it was possible they were incorrectly identified as diabetic because of taking anti-diabetic medication. Diagnosis of DM was time-varying and could be prior to diagnosis of RA whereby a person would be flagged as diabetic throughout follow-up, or during follow-up whereby a person would be flagged as diabetic from the point of DM diagnosis. Where the diagnosis was made on the basis of two sequential prescriptions, the date of onset was allocated as the date of the second prescription to avoid immortal time bias.

Oral GC therapy was identified using product codes from prescription data. Patients were classified by current/recent use of GCs, whereby a person was classified as exposed for the duration of each GC prescription and for 6 months after the end of the prescription.

Outcomes

All-cause and CV mortality were identified through linkage to ONS data with date of death and cause of death provided. Cause of death was recorded on ONS using International Statistical Classification of Diseases and Related Health Problems (ICD) version 10 codes. Deaths prior to 2001 were recorded using ICD-9 codes and these were mapped to ICD-10 codes. There also were 31 deaths recorded on CPRD but not on ONS and these were included in the all-cause mortality analyses. CV mortality was identified using ICD-10 codes under the circulatory chapter heading as the underlying cause of death.

Covariates

Age at RA diagnosis was calculated using year of birth and year of RA diagnosis. Gender was given on the CPRD database. Baseline Charlson comorbidity index was determined using an adaption of the index for CPRD data where diseases were identified through Read codes for diagnosis at any point prior to RA diagnosis [14]. DMARD types and non-steroidal anti-inflammatory drugs (NSAIDs) were identified using product codes and were time-varying. GC use in the year preceding baseline was determined from GC prescriptions prior to baseline. Baseline smoking category (ever or never) was determined using Read codes and product codes at any point up to RA diagnosis, or in the 3 months after RA diagnosis. Prior macrovascular disease was defined as diseases of large blood vessels including myocardial infarction, stroke, peripheral artery disease or amputation [15] and were identified through Read codes prior to RA diagnosis. Body mass index (BMI) at baseline was calculated using median height and weight measurements from the 5 years prior to baseline. All code lists can be found in Additional file 1.

Analysis

For both outcomes, mortality rates were estimated (with 95% confidence intervals (CI)), stratified by time-varying DM status and time-varying current/recent use of GCs. As mentioned earlier, the baseline risk of CV and all-cause mortality for patients with RA and DM will be higher than those with RA only. Therefore, to investigate the impact of GCs both rate ratios (RR) and rate differences (RD) between GC users and non-GC users were calculated for those with and without DM separately.

When estimating the effect of both GC exposure and DM status, the presence of interaction was measured on both the multiplicative scale, corresponding to the RR, and on the additive scale, corresponding to the RD. Interaction on the additive scale can give more meaningful comparisons as it is not dependent on baseline risks [16]. Crude and adjusted Cox proportional hazards (PH) regression models were fitted with an interaction term for time-varying DM and time-varying current/recent use of GCs. Multiplicative interaction was assessed via the inclusion of an interaction term in the Cox model.

Additive interaction cannot be estimated directly from the Cox model as it depends on the baseline hazard function [17]. However, we can estimate the Relative Excess Risk due to Interaction (RERI) and Ratio of Absolute Effects (RAE): 1) RERI [17, 18] assesses if there is a difference in the hazard differences. The RERI is equal to 0 if the additive interaction effect is equal to 0. Therefore, if it is statistically significantly different from zero then this is interpreted as a statistically significant difference in the hazard differences between those with and without DM, and indicates the direction of the effect. 2) RAE is defined as the ratio of hazard differences in patients with DM compared to those without DM (See Additional file 2 for further information). Departure from 1 indicates a difference in the two groups and it was calculated here in addition to the RERI as it gives an indication of the magnitude of the difference in subgroup absolute effects, unlike the RERI. Both measures are calculated after the Cox model as a function of the model parameters.

Missing data

Ever smoking at baseline and baseline BMI had 753 (8%) and 3849 (42%) missing data, respectively. Multiple imputation with 57 imputations was used to replace these missing values. The number of imputations was based on the fraction of missing information. Forty-nine patients did not have a Townsend score, however this was not imputed as it was not used in the final models.

Results

There were 15,833 patients identified who had a diagnosis of RA and were registered at their practice for at least 1 year prior to diagnosis, 6748 were excluded due to either inconsistent follow-up dates, being age 18 years or under at diagnosis, being registered at a practice that did not consent to ONS linkage or having a diagnosis of PCOS and being treated with metformin, resulting in 9085 patients in the final cohort (Fig. 1). The cohort had a mean follow-up of 5.2 years (standard deviation 3.5 years).

At baseline there were 1034 patients with DM, and 761 patients developed DM during follow-up. Compared to those without DM at baseline, those with DM at baseline were older (DM: mean 64 years vs non-DM: mean 59 years) had a greater proportion of males (DM: 37% vs non-DM: 30%) and ever smokers (DM: 58% vs non-DM: 50%), had more GC use prior to baseline (DM: 31% vs non-DM: 23%), had more macrovascular disease at baseline (DM: 11% vs non-DM: 4%) and had a higher BMI (DM: 30 vs non-DM: 27) (Table 1). 50% of patients had used GC at any point during follow-up. Those with prior DM had slightly higher average GC dose over follow-up (DM: 4.9 mg prednisolone equivalent dose (PED) vs non-DM: 4.4 mg PED). Across both those with and without DM those who ever used GC were older and had more prior macrovascular disease.

All-cause mortality

During follow-up there were 1,005 deaths. Mortality rates differed according to the presence of DM and the use of GC therapy. For those with DM, the mortality rate was 67.4 (95% CI 57.1 to 79.5) per 1000 person-years (pyrs) in those with GC exposure and 22.5 (95% CI 18.7 to 27.1) per 1000 pyrs in those without GC exposure. For those without DM, the mortality rate was 44.6 (95% CI

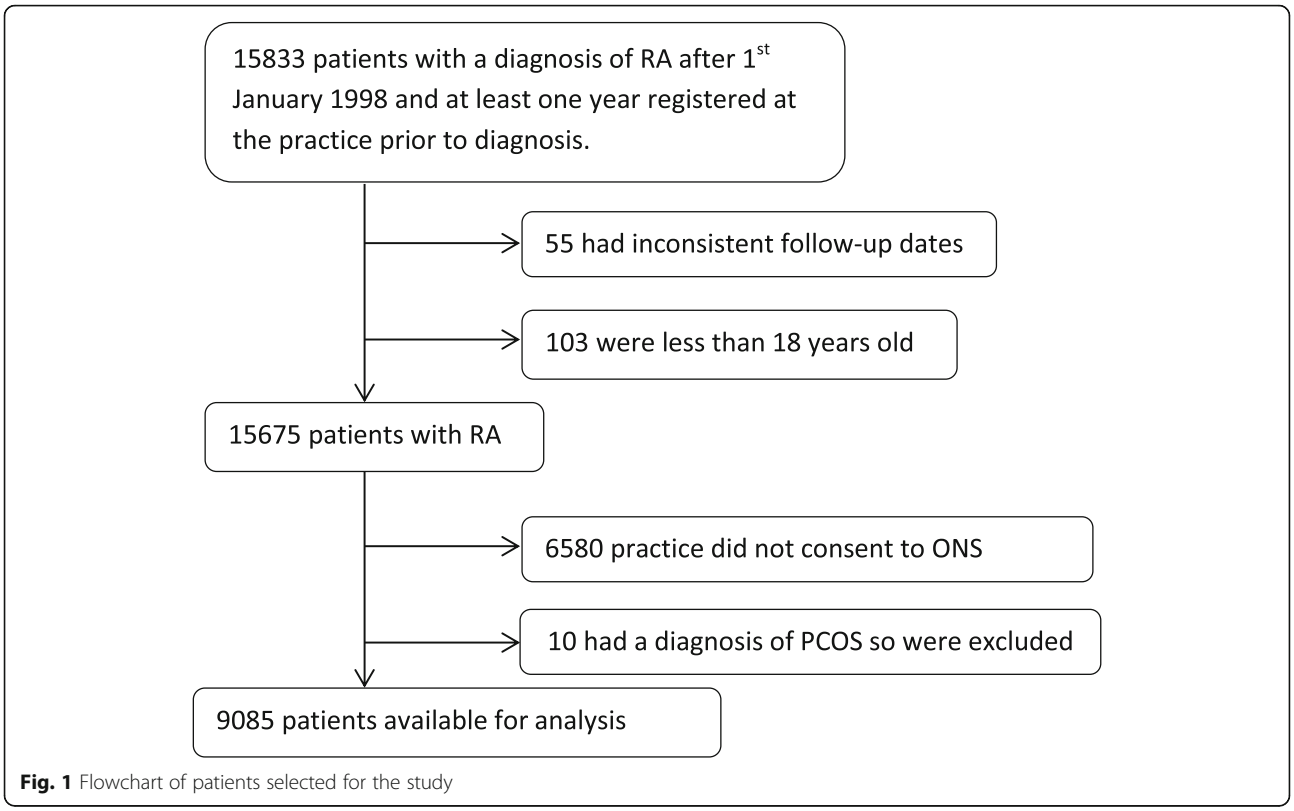


Fig. 1 Flowchart of patients selected for the study

40.6 to 48.9) per 1000 pyrs in those exposed to GCs and 10.2 (95% CI: 9.1 to 11.4) per 1000 pyrs in those without GC exposure. The risk ratio for GC use was slightly lower for those with DM (DM RR 2.99 (95% confidence interval (CI) 2.32 to3.87) compared to those with no DM RR 4.37 (95% CI 3.77 to 5.07)). However, despite this lower RR, those with DM had a *higher RD* compared to those without DM (DM RD: 44.9 (95% CI: 32.9 to 56.8) vs no DM RD: 34.4 (95% CI: 30.1 to 38.7 per 1000 pyrs) (Table 2).

The unadjusted Cox PH model for all-cause mortality showed current/recent GC use and DM interacted on the multiplicative scale (0.69 (95% CI 0.51, 0.91)). Adjustment removed this significant interaction (0.86 (95% CI: 0.64–1.15)) (Table 3). In both the unadjusted and adjusted models both the RERI and RAE indicated increased risk for those with DM and current/recent GC use but were not statistically significant (adjusted RAE: 1.22 (95% CI: 0.86 to 1.72) (Table 3).

CV mortality

There were 384 CV deaths during follow-up. A similar pattern was seen for CV mortality, where a slightly lower RR was seen for those with DM compared to those without DM, but the RD was higher for those with DM (Table 2). The unadjusted and adjusted Cox models showed that DM did not interact with ever GC use on the multiplicative scale, the additive interaction indicated increased risk but was not statistically significant (Table 3).

Discussion

In this study, we have shown that in patients with RA and DM, the RR of GC use on all-cause and CV mortality was slightly lower than in patients with RA alone. This might seem reassuring at first glance, suggesting the impact of GC therapy in patients with DM is no worse than in patients without DM. However, the RD was notably higher in those with DM compared to those without. The higher baseline mortality rate for those with DM is thus resulting in a greater number of excess deaths despite the slightly lower RR. When examined together in an adjusted Cox PH model, current/recent use of GC in those with DM was associated with a non-significant absolute increased hazard of all-cause mortality compared to those without DM, but not a relative increased hazard. A similar pattern was seen for CV mortality. The increased absolute hazard for all-cause mortality indicates the greater public health impact of people with RA using GCs if they have DM. This increase is not seen on the multiplicative scale because the comparison made is relative to other patients with DM who have a higher risk of mortality prior to using GCs. Notably, most studies only assess effect modification or interaction on the multiplicative scale, despite

Table 1 Baseline characteristics by diabetes mellitus status and ever use of glucocorticoids during follow-up ($N = 9085$)

	DM at baseline			No DM at baseline			DM during FU[a]		
	All subjects	Never users	Ever users	All subjects	Never users	Ever users	All subjects	Never users	Ever users
	$N = 1034$	$N = 512$	$N = 522$	$N = 8051$	$N = 4026$	$N = 4025$	$N = 761$	$N = 269$	$N = 492$
Females, n (%)	652 (63.1)	325 (63.5)	327 (62.6)	5600 (69.6)	2878 (71.5)	2722 (67.6)	503 (66.1)	176 (65.4)	327 (66.5)
Age at baseline (years), mean (standard deviation (SD))	64.42 (13.0)	63 13.6)	65.81 (12.3)	58.51 (14.7)	55.94 (14.6)	61.07 (14.4)	64.63 (12.8)	62.4 (12.9)	65.86 (12.5)
Body Mass Index in year prior to baseline									
mean (SD)	29.77 (6.5)	29.78 (6.7)	29.77 (6.3)	27.14 (5.5)	27.17 (5.5)	27.1 (5.4)	29.86 (6.9)	30.71 (6.8)	29.35 (6.9)
Missing (%)	180 (17.4)	87 (17.0)	93 (17.8)	3669 (45.6)	1853 (46.0)	1816 (45.1)	224 (29.4)	68 (25.3)	156 (31.7)
Smoking status at baseline, n (%)									
Never smoker	407 (39.4)	203 (39.7)	204 (39.1)	3337 (41.5)	1748 (43.4)	1589 (39.5)	234 (30.8)	80 (29.7)	154 (31.3)
Ever smoker	600 (58.0)	293 (57.2)	307 (58.8)	3988 (49.5)	1949 (48.4)	2039 (50.7)	505 (66.4)	182 (67.7)	323 (65.7)
Missing	2 (0.2)	1 (0.2)	1 (0.2)	41 (0.5)	20 (0.5)	21 (0.5)	5 (0.7)	1 (0.4)	4 (0.8)
SES quintile at baseline (In subset), n (%)									
First (least deprived)	214 (20.7)	102 (19.9)	112 (21.5)	1830 (22.7)	910 (22.6)	920 (22.9)	165 (21.7)	60 (22.3)	105 (21.3)
Second	224 (21.7)	98 (19.1)	126 (24.1)	1993 (24.8)	1023 (25.4)	970 (24.1)	182 (23.9)	65 (24.2)	117 (23.8)
Third	215 (20.8)	121 (23.6)	94 (18.0)	1731 (21.5)	839 (20.8)	892 (22.2)	156 (20.5)	62 (23.1)	94 (19.1)
Fourth	229 (22.2)	119 (23.2)	110 (21.1)	1470 (18.3)	741 (18.4)	729 (18.1)	161 (21.2)	49 (18.2)	112 (22.8)
Fifth (most deprived)	150 (14.5)	71 (13.9)	79 (15.1)	986 (12.3)	493 (12.3)	493 (12.3)	92 (12.1)	32 (11.9)	60 (12.2)
Missing	2 (0.2)	1 (0.2)	1 (0.2)	41 (0.5)	20 (0.5)	21 (0.5)	6 (0.8)	1 (0.4)	5 (1.0)
Charlson comorbidity index at baseline, mean (SD)	2.57 (0.8)	2.48 (0.8)	2.66 (0.8)	1.32 (0.7)	1.23 (0.6)	1.42 (0.7)	2.63 (0.8)	2.39 (0.7)	2.76 (0.8)
Prior history of macrovascular diseases, n (%)	113 (10.9)	41 (8.0)	72 (13.8)	297 (3.7)	108 (2.7)	189 (4.7)	77 (10.1)	16 (6.0)	61 (12.4)
History of GC use in year prior to baseline, n (%)	325 (31.4)	42 (8.2)	283 (54.2)	1861 (23.1)	256 (6.4)	1605 (39.9)	299 (39.3)	3 (1.1)	296 (60.2)
Duration of diabetes at baseline (yrs)									
Mean (SD)	4.59 (3.7)	4.88 (3.8)	4.31 (3.6)	N/A	N/A	N/A	N/A	N/A	N/A
Number of anti-DM medication prior to baseline									
0	525 (50.8)	264 (50.6)	261 (51.0)	N/A	N/A	N/A	N/A	N/A	N/A
1	482 (46.6)	249 (47.7)	233 (45.5)	N/A	N/A	N/A	N/A	N/A	N/A
2	26 (2.5)	8 (1.5)	18 (3.5)	N/A	N/A	N/A	N/A	N/A	N/A
3	1 (0.1)	1 (0.2)	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A
Prescribed insulin prior to follow-up, n (%)	155 (15.0)	86 (16.8)	69 (13.2)	N/A	N/A	N/A	N/A	N/A	N/A
DMARDs prescribed during follow-up, n (%)									
Methotrexate	633 (61.2)	309 (60.4)	324 (62.1)	5012 (62.3)	2386 (59.3)	2626 (65.2)	353 (46.4)	110 (40.9)	243 (49.4)
Hydroxychloroquine	261	132	129	2304	1129	1175	138	49 (18.2)	89 (18.1)

Table 1 Baseline characteristics by diabetes mellitus status and ever use of glucocorticoids during follow-up (*N* = 9085) *(Continued)*

	DM at baseline			No DM at baseline			DM during FU[a]		
	All subjects	Never users	Ever users	All subjects	Never users	Ever users	All subjects	Never users	Ever users
	N = 1034	*N* = 512	*N* = 522	*N* = 8051	*N* = 4026	*N* = 4025	*N* = 761	*N* = 269	*N* = 492
	(25.2)	(25.8)	(24.7)	(28.6)	(28.0)	(29.2)	(18.1)		
Sulfasalazine	372 (36.0)	171 (33.4)	201 (38.5)	3333 (41.4)	1591 (39.5)	1742 (43.3)	207 (27.2)	80 (29.7)	127 (25.8)
Leflunomide	75 (7.3)	28 (5.5)	47 (9.0)	719 (8.9)	264 (6.6)	455 (11.3)	57 (7.5)	13 (4.8)	44 (8.9)
Other	70 (6.8)	15 (2.9)	55 (10.5)	549 (6.8)	117 (2.9)	432 (10.7)	54 (7.1)	6 (2.2)	48 (9.8)
Average GC dose during follow up, mean (SD)	4.93 (14.0)	0	9.8 (18.5)	4.38 (6.3)	0	8.75 (6.4)	5.02 (6.3)	0	7.77 (6.4)

[a]Characteristics at time of diabetes mellitus diagnosis

recommendations to use both the multiplicative and additive scales [16, 19].

To our knowledge no previous studies have looked at the effect of both GCs and DM on mortality in patients with RA. Studies have looked at short term diabetic outcomes with GC use, investigating its effects on glucose intolerance or metabolic syndrome in patients with RA [8, 20]. Two studies have investigated longer term outcomes of GC use in patients with DM but not RA. One looked at mortality 14 years after diagnosis and found that after adjustment for age and gender there was not increased mortality in patients with DM who had GC treatment compared to those who did not, however only small numbers of patients had GC treatment in this study (35/1334) [10]. The other study aimed to describe the adverse effects of GC treatment in patients with DM, but did not discuss mortality [11]. We and others have previously shown GC therapy to be associated with higher all-cause mortality rates in patients with RA. However, a causal association is difficult to establish as several biases are at play in an observational study including 'peri-mortal bias' [21].

This was a large study that used electronic medical records that are a rich source of medical information. CPRD data has been shown to be broadly representative of the UK population, so results should be generalisable to the UK RA population [12]. However, there are some limitations with the study. Although we used a validated algorithm to identify patients with RA there could still be some misclassification. Further misclassification may

Table 2 Mortality rates, rate ratios and rate difference by diabetes mellitus and glucocorticoid use status

Outcome		✓RA ✓DM ✓GC	✓RA ✓DM ×GC	✓RA ×DM ✓GC	✓RA ×DM ×GC
All-cause mortality	Number of events	140	112	442	311
	Follow-up time	2077.5	4975.3	9914.0	30463.4
	IR per 1000 person-years (95% confidence interval (CI))	67.4 (57.1 to 79.5)	22.5 (18.7 to 27.1)	44.6 (40.6 to 48.9)	10.2 (9.1 to 11.4)
	Rate ratio (95% CI)	2.99 (2.32 to 3.87)		4.37 (3.77 to 5.07)	
	Rate difference, per 1000 person-years (95% CI)	44.9 (32.9 to 56.8)		34.4 (30.1 to 38.7)	
CV mortality	Number of events	46	48	150	140
	Follow-up time	2077.5	4975.3	9920.7	30497.3
	IR, per 1000 person-years. (95% CI)	22.14 (16.6 to 29.6)	9.65 (7.3 to 12.8)	15.1 (12.9 to 17.7)	4.6 (3.9 to 5.4)
	Rate ratio	2.30 (1.50, 3.51)		3.29 (2.60 to 4.18)	
	Rate difference (per 1000 person-years)	12.49 (5.54, 19.45)		10.53 (7.99 to 13.07)	

Table 3 Multiplicative and additive interaction[a] between diabetes mellitus and ever glucocorticoid use

Outcome	Multiplicative interaction	Additive interaction	
	Hazard ratio (95% confidence interval)	RERI (95% confidence interval)	RAE (95% confidence interval)
All-cause mortality			
Unadjusted	0.69 (0.51 to 0.91)	0.94 (−0.29 to 2.16)	1.27 (0.95 to 1.70)
All-cause mortality			
Adjusted[b]	0.86 (0.64 to 1.15)	0.41 (−0.36 to 1.18)	1.22 (0.86 to 1.72)
CV mortality			
Unadjusted	0.70 (0.44 to 1.11)	0.40 (−1.19 to 2.00)	1.17 (0.64 to 2.14)
CV mortality			
Adjusted[b]	0.93 (0.60 to 1.48)	0.11 (−0.75 to 0.96)	1.11 (0.48 to 2.57)

[a] On the multiplicative scale significant interaction is different than 1, on the additive scale significant interaction for the RERI is different than 0 and for the RAE is different than 1

[b]Adjusted for age, gender, Charlson comorbidity index, baseline BMI, baseline smoking status, DMARDs, prior GC, prior macrovascular disease and NSAIDs

result from medication being based on prescription data rather than dispensing data. However, any differences between prescribed medication and medication dispensed are unlikely to differ by DM status. To allow examination of interaction a simple model of oral GC exposure was used, therefore it was not possible to examine the impact of GC dose or intramuscular GCs. This study focuses on type 2 DM, as GCs induce insulin resistance similar to type 2 diabetes. Results are likely to be similar with type 1 diabetes, but given the different pathogenetic mechanisms, further work would be required to confirm this. There could be confounding by indication, as RA disease severity has been shown to confound the relationship between GCs and CVD in RA [22]. However, there is no measure of disease activity available on CPRD and we would not expect the confounding to differentially affect those with or without DM. There may be known unmeasured confounding, there were no measures of biologic DMARD use in this study as biologics are only prescribed in secondary care in the UK. This may be important as biologics have been shown to be associated with reduced CVD [23]. Unfortunately we were not able to use methods to explore unmeasured confounding as most are applied to relative risks rather than additive interaction terms.

Conclusions

This study gives an indication that GC therapy may be associated with a higher number of deaths in patients with RA and comorbid type 2 DM. Rheumatologists should consider DM status when prescribing GCs to patients with RA given this potential impact of GC therapy on glucose control and mortality.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s41927-019-0105-4.

Additional file 1. Codelists for disease and drug definition.

Additional file 2. The RAE measure.

Abbreviations

BMI: Body mass index; CI: Confidence interval; CPRD: Clinical practice research datalink; CV: Cardiovascular; DM: Diabetes mellitus; DMARDs: Disease modifying anti-rheumatic drugs; GCs: Glucocorticoids; Hba1c: Glycosylated haemoglobin; ICD: International Statistical Classification of Diseases and Related Health Problems; NSAIDs: Non-steroidal anti-inflammatory drugs; ONS: Office for National Statistics; PCOS: Polycystic ovary syndrome; pyrs: Person years; RA: Rheumatoid arthritis; RAE: Ratio of Absolute Effects; RD: Rate difference; RERI: Relative Excess Risk due to Interaction; RR: Rate ratio

Acknowledgements

Not applicable.

Authors' contributions

REC contributed to study design, conducting the analyses, interpreting the data, drafting the manuscript, critically revising the manuscript, approving the final version and agrees to be accountable for all aspects of the work. AM, RE and WGD contributed to study design, interpreting the data, critically revising the manuscript, approving the final version and agrees to be accountable for all aspects of the work. MM contributed to acquiring the data, interpreting the data, critically revising the manuscript, approving the final version and agrees to be accountable for all aspects of the work. ML and JHH contributed to interpreting the data, critically revising the manuscript, approving the final version and agrees to be accountable for all aspects of the work.

Funding

This work was supported by the Centre for Epidemiology Versus Arthritis (21755) and the NIHR Biomedical Research Centre. The funder was not involved in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

Clinical Practice Research Datalink (CPRD) data can be accessed with an appropriate licence from the CPRD and with approval from the Independent Scientific Advisory Committee. Licences are available from CPRD: Clinical Practice Research Datalink, The Medicines and Healthcare products Regulatory Agency, 10th Floor, 10 South Colonnade, Canary Wharf, London E14 4PU, England or http://www.cprd.com.

Ethics approval and consent to participate

The protocol for this study has been approved by Independent Scientific Advisory Committee for Medicines and Healthcare Regulatory Agency database research (Protocol number: 11_113RA4).

Consent for publication

Not applicable.

Competing interests

WGD has received consultancy fees from Bayer and Google outside of the submitted work. All other authors have no competing interests.

Author details

[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, Manchester Academic

Health Science Centre, The University of Manchester, Manchester, UK. [2]Centre for Biostatistics, School of Health Sciences, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. [3]Ontario Best Practices Research Initiative, University Health Network, Toronto, Ontario, Canada. [4]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

Received: 18 August 2019 Accepted: 15 November 2019
Published online: 19 February 2020

References

1. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology. 2002;41(7):793–800.
2. Avina-Zubieta JA, Choi HK, Sadatsafavi M, Etminan M, Esdaile JM, Lacaille D. Risk of cardiovascular mortality in patients with rheumatoid arthritis: a meta-analysis of observational studies. Arthritis Rheum. 2008;59(12):1690–7.
3. Hench PS, Kendall EC, Slocumb CH, Polley HF. Effects of cortisone acetate and pituitary acth on rheumatoid arthritis, rheumatic fever and certain other conditions - a study in clinical physiology. Arch Intern Med. 1950;85(4):545–666.
4. Black RJ, Joseph RM, Brown B, Movahedi M, Lunt M, Dixon WG. Half of U.K. patients with rheumatoid arthritis are prescribed oral glucocorticoid therapy in primary care: a retrospective drug utilisation study. Arthritis Res Ther. 2015;17:375.
5. del Rincon I, Battafarano DF, Restrepo JF, Erikson JM, Escalante A. Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. Arthritis Rheumatol. 2014;66(2):264–72.
6. Davis JM 3rd, Maradit Kremers H, Crowson CS, Nicola PJ, Ballman KV, Therneau TM, et al. Glucocorticoids and cardiovascular events in rheumatoid arthritis: a population-based cohort study. Arthritis Rheum. 2007;56(3):820–30.
7. Movahedi M, Beauchamp ME, Abrahamowicz M, Ray DW, Michaud K, Pedro S, et al. Risk of incident diabetes mellitus associated with the dosage and duration of oral glucocorticoid therapy in patients with rheumatoid arthritis. Arthritis Rheumatol. 2016;68(5):1089–98.
8. Hoes JN, van der Goes MC, van Raalte DH, van der Zijl NJ, den Uyl D, Lems WF, et al. Glucose tolerance, insulin sensitivity and beta-cell function in patients with rheumatoid arthritis treated with or without low-to-medium dose glucocorticoids. Ann Rheum Dis. 2011;70(11):1887–94.
9. Da Silva JA, Jacobs JW, Kirwan JR, Boers M, Saag KG, Ines LB, et al. Safety of low dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. Ann Rheum Dis. 2006;65(3):285–93.
10. Olivarius Nde F, Siersma V, Dyring-Andersen B, Drivsholm T, Hansen LJ, Henriksen JE. Patients newly diagnosed with clinical type 2 diabetes during oral glucocorticoid treatment and observed for 14 years: all-cause mortality and clinical developments. Basic Clin Pharmacol Toxicol. 2011;108(4):285–8.
11. Reynolds RM, Labad J, Sears AV, Williamson RM, Strachan MW, Deary IJ, et al. Glucocorticoid treatment and impaired mood, memory and metabolism in people with diabetes: the Edinburgh type 2 diabetes study. Eur J Endocrinol. 2012;166(5):861–8.
12. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol. 2015;44(3):827–36.
13. Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? Arthritis Rheum. 2008;59(9):1314–21.
14. Khan NF, Perera R, Harper S, Rose PW. Adaptation and validation of the Charlson index for read/OXMIS coded databases. BMC Fam Pract. 2010;11:1.
15. Kontopantelis E, Springate DA, Reeves D, Ashcroft DM, Rutter MK, Buchan I, et al. Glucose, blood pressure and cholesterol levels and their relationships to clinical outcomes in type 2 diabetes: a retrospective cohort study. Diabetologia. 2015;58(3):505–18.
16. VanderWeele TJ, Knol MJ. A tutorial on interaction. Epidemiol Methods. 2014;3(1):33–72.
17. Li R, Chambless L. Test for additive interaction in proportional hazards models. Ann Epidemiol. 2007;17(3):227–36.
18. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
19. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. Int J Epidemiol. 2012;41(2):514–20.
20. den Uyl D, van Raalte DH, Nurmohamed MT, Lems WF, Bijlsma JW, Hoes JN, et al. Metabolic effects of high-dose prednisolone treatment in early rheumatoid arthritis: balance between diabetogenic effects and inflammation reduction. Arthritis Rheum. 2012;64(3):639–46.
21. Movahedi M, Costello R, Lunt M, Pye SR, Sergeant JC, Dixon WG. Oral glucocorticoid therapy and all-cause and cause-specific mortality in patients with rheumatoid arthritis: a retrospective cohort study. Eur J Epidemiol. 2016;31(10):1045–55.
22. van Sijl AM, Boers M, Voskuyl AE, Nurmohamed MT. Confounding by indication probably distorts the relationship between steroid use and cardiovascular disease in rheumatoid arthritis: results from a prospective cohort study. PLoS One. 2014;9(1):e87965.
23. Agca R, Heslinga SC, Rollefstad S, Heslinga M, McInnes IB, Peters MJ, et al. EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. Ann Rheum Dis. 2017;76(1):17–28.

Publisher's Note

The RAE measure

Suppose we wish to determine if the effect of treatment, T, on some outcome, Y, is different across patient subgroups defined by variable M, where T and M are binary. In a time-to-event data setting, the additive interaction effect is defined as the difference in risk differences across subgroups of M,

$$INT_A = \big(\lambda(l; T=1, M=1, X) - \lambda(l; T=0, M=1, X)\big) - \big(\lambda(l; T=1, M=0, X) - \lambda(l; T=0, M=0,)\big)$$

where $\lambda(l, \dots)$ denotes the hazard function at time l. If this measure departs from 0, there is a difference in the hazard differences in the subgroups of M.

The most common model for time-to-event data is the Cox proportional hazards model. This is a semi-parametric model which incorporates a baseline hazard, $\lambda_0(l)$, which is a function of time and is not estimated. However, if we were to fit this model including the treatment, the moderator, interaction between the two and any additional covariates, X_k, $k = 1, .., n$ (equation 1) we would not be able to directly estimate the additive interaction effect as a function of the estimated regression coefficients as this measure depends on the baseline hazard function $\lambda_0(l)$ (equation 2).

$$\lambda(l; T, X) = \lambda_0(l) e^{\beta_0 + \beta_1 T + \beta_2 M + \beta_3 TM + \sum_{k=1}^{n} c_k X_k} \quad (1)$$

$$\begin{aligned} \widehat{INT}_A &= (\lambda_0(l) e^{\hat{\beta}_0 + \hat{\beta}_1 T + \hat{\beta}_2 M + \hat{\beta}_3 TM + \sum_{k=1}^{n} \hat{c}_k X_k} - \lambda_0(l) e^{\hat{\beta}_0 + \hat{\beta}_2 M + \sum_{k=1}^{n} \hat{c}_k X_k}) \\ &- (\lambda_0(l) e^{\hat{\beta}_0 + \hat{\beta}_1 T + \sum_{k=1}^{n} \hat{c}_k X_k} - \lambda_0(l) e^{\hat{\beta}_0 + \sum_{k=1}^{n} \hat{c}_k X_k}) \\ &= \lambda_0(l) e^{\sum_{k=1}^{n} \hat{c}_k X_k} (e^{\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3} - e^{\hat{\beta}_1} - e^{\hat{\beta}_2} + 1) \end{aligned} \quad (2)$$

However, we can also compare the hazard differences in subgroups of M by considering the ratio of the hazard differences; defined the Ratio of Absolute Effects (RAE) measure.

$$RAE = \frac{\lambda(l; T = 1, M = 1, X) - \lambda(l; T = 0, M = 1, X)}{\lambda(l; T = 1, M = 0, X) - \lambda(l; T = 0, M = 0, X)}$$

If this measure departs from 1, there is a difference in the hazard differences in the subgroups of M. An $RAE = a$ implies that the absolute treatment effect in patients with $M = 1$ is a times that in patients with $M = 0$. An $RAE > 1$ suggests either a larger positive absolute effect or a smaller negative effect in patients with $M = 1$ compared to patients with $M = 0$. If $RAE \approx 1$, there is no suggestion of treatment effect modification by M on the additive scale. Moreover if $RAE < 0$, the estimated absolute treatment effect is in the opposite direction in the two subgroups.

When the RAE is estimated as a function of the regression coefficients of the Cox model, the baseline hazard function cancels out; thus, unlike the additive interaction effect, the RAE measure can be calculated from this model.

$$\begin{aligned}\widehat{RAE} &= \frac{\lambda_0(l)e^{\hat{\beta}_0+\hat{\beta}_1+\hat{\beta}_2+\hat{\beta}_3+\sum_{k=1}^{n}\hat{c}_k X_k} - \lambda_0(l)e^{\hat{\beta}_0+\hat{\beta}_2+\sum_{k=1}^{n}\hat{c}_k X_k}}{\lambda_0(l)e^{\hat{\beta}_0+\hat{\beta}_1+\sum_{k=1}^{n}\hat{c}_k X_k} - \lambda_0(l)e^{\hat{\beta}_0+\sum_{k=1}^{n}\hat{c}_k X_k}} \\ &= \frac{e^{\hat{\beta}_1+\hat{\beta}_2+\hat{\beta}_3} - e^{\hat{\beta}_2}}{e^{\hat{\beta}_1} - 1}\end{aligned}$$

The standard error for this measure can be estimated using the delta method. As the RAE is a ratio, it is unlikely to be normally distributed. Therefore, to obtain a 95% confidence interval for the $\widehat{RAE}$, one can calculate $\log(RAE)$ and it's standard error, calculate a 95% confidence interval for $\log(RAE)$ assuming it is approximately normally distriuted (95% CI: estimate $\pm 1.95 \times$ standard error) and then exponentiate the upper and lower limits.

BMC Rheumatology

CORRECTION

Open Access

Correction to: The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study

Ruth E. Costello[1], Antonia Marsden[2], Mohammad Movahedi[1,3], Mark Lunt[1], Jenny H. Humphreys[1], Richard Emsley[4] and William G. Dixon[1*]

Correction to: BMC Rheumatol 4, 4 (2020)
https://doi.org/10.1186/s41927-019-0105-4

Following publication of the original article [1], the authors noted several errors in the reported values in Table 2, the 'Results' section of the abstract, and in the first sentence of the "All-cause mortality" sub-section. The correct Table and text are given below with the corrected values highlighted in bold for the Abstract.

The original article has been updated.

Abstract

Results: In those without DM GC use had a 4.4-fold increased all-cause mortality RR (95% confidence interval (CI): **3.77 to 5.07**) compared to non-use, whilst those with DM had a lower RR for GC use (**2.99 (95% CI: 2.32, 3.87**)). However, those with DM had a higher RD associated with GC use because of their higher baseline risk. In those with DM, GC use was associated with an additional **44.9** deaths/1000 person-years (pyrs) (95% CI: **32.9 to 56.8**) compared to non-use, while in those without DM GC use was associated with an additional **34.4** deaths/1000 pyrs (95% CI: **30.1 to 38.7**). A similar pattern was seen for CV mortality. The adjusted Cox proportional hazards model showed no evidence of multiplicative interaction, but additive interaction indicated a non-significant increased risk. For CV mortality there was no interaction on either scale.

In the "All-cause mortality" sub-section it now correctly reads "During follow-up there were 1005 deaths" rather than 1002 deaths.

Author details
[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [2]Centre for Biostatistics, School of Health Sciences, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. [3]Ontario Best Practices Research Initiative, University Health Network, Toronto, Ontario, Canada. [4]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

Published online: 02 March 2021

Reference

1. Costello RE, Marsden A, Movahedi M, et al. The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study. BMC Rheumatol. 2020;4:4. https://doi.org/10.1186/s41927-019-0105-4.

The original article can be found online at https://doi.org/10.1186/s41927-019-0105-4.

* Correspondence: will.dixon@manchester.ac.uk
[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
Full list of author information is available at the end of the article

Table 2 Mortality rates, rate ratios and rate difference by diabetes mellitus and glucocorticoid use status

Outcome		✓RA ✓DM ✓GC ✗GC	✓RA ✓DM ✗GC	✓RA ✗DM ✓GC ✗GC	✓RA ✗DM ✗GC
All-cause mortality	Number of events	140	112	442	311
	Follow-up time	2077.5	4975.3	9914.0	30463.4
	IR per 1000 person-years (95% confidence interval (CI))	67.4 (57.1 to 79.5)	22.5 (18.7 to 27.1)	44.6 (40.6 to 48.9)	10.2 (9.1 to 11.4)
	Rate ratio (95% CI)	2.99 (2.32 to 3.87)		4.37 (3.77 to 5.07)	
	Rate difference, per 1000 person-years (95% CI)	44.9 (32.9 to 56.8)		34.4 (30.1 to 38.7)	
CV mortality	Number of events	46	48	150	140
	Follow-up time	2077.5	4975.3	9920.7	30497.3
	IR, per 1000 person-years. (95% CI)	22.14 (16.6 to 29.6)	9.65 (7.3 to 12.8)	15.1 (12.9 to 17.7)	4.6 (3.9 to 5.4)
	Rate ratio	2.30 (1.50, 3.51)		3.29 (2.60 to 4.18)	
	Rate difference (per 1000 person-years)	12.49 (5.54, 19.45)		10.53 (7.99 to 13.07)	

Publication 5: Glucocorticoid use is associated with an increased risk of hypertension.

RHEUMATOLOGY

Rheumatology 2021;60:132–139
doi:10.1093/rheumatology/keaa209
Advance Access publication 27 June 2020

Original article

Glucocorticoid use is associated with an increased risk of hypertension

Ruth E. Costello [1], Belay B. Yimer[1], Polly Roads[1], Meghna Jani [1,2] and William G. Dixon [1,2]

Abstract

Objectives. Patients with RA are frequently treated with glucocorticoids (GCs), but evidence is conflicting about whether GCs are associated with hypertension. The aim of this study was to determine whether GCs are associated with incident hypertension in patients with RA.

Methods. A retrospective cohort of patients with incident RA and without hypertension was identified from UK primary care electronic medical records (Clinical Practice Research Datalink). GC prescriptions were used to determine time-varying GC use, dose and cumulative dose, with a 3 month attribution window. Hypertension was identified through either: blood pressure measurements >140/90 mmHg, or antihypertensive prescriptions and a Read code for hypertension. Unadjusted and adjusted Cox proportional hazards regression models were fitted to determine whether there was an association between GC use and incident hypertension.

Results. There were 17 760 patients in the cohort. A total of 7421 (42%) were prescribed GCs during follow-up. The incident rate of hypertension was 64.1 per 1000 person years (95% CI: 62.5, 65.7). The Cox proportional hazards model indicated that recent GC use was associated with a 17% increased hazard of hypertension (hazard ratio 1.17; 95% CI: 1.10, 1.24). When categorized by dose, only doses above 7.5 mg were significantly associated with hypertension. Cumulative dose did not indicate a clear pattern.

Conclusion. Recent GC use was associated with incident hypertension in patients with RA, in particular doses ≥7.5 mg were associated with hypertension. Clinicians need to consider cardiovascular risk when prescribing GCs, and ensure blood pressure is regularly monitored and treated where necessary.

Key words: rheumatoid arthritis, cardiovascular, epidemiology, immunosuppressants, primary care rheumatology

CLINICAL SCIENCE

Rheumatology key messages

- Glucocorticoid use increases the risk of hypertension in patients with RA.
- Glucocorticoid doses of ≥7.5 mg in particular are associated with hypertension.
- Blood pressure should be monitored in patients with RA prescribed glucocorticoids.

Introduction

RA is a chronic inflammatory condition, affecting around 1% of the general population [1]. Patients with RA are at an increased risk of all-cause mortality compared with the general population [2]. Cardiovascular (CV) disease is a major driver of this: a meta-analysis showed that patients with RA have a 50% increased risk of CV mortality compared with the general population [3]. This increased risk of CV disease [4] is due not only to traditional risk factors such as smoking and hypertension,

[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, University of Manchester, Manchester and [2]Department of Rheumatology, Salford Royal NHS Foundation Trust, Salford, UK

Submitted 13 December 2019; accepted 31 March 2020

Correspondence to: Ruth Costello, Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, The University of Manchester, Oxford Road, Manchester M13 9PT, UK. E-mail: ruth.costello@manchester.ac.uk

but also to disease-related factors such as disease activity, which increases inflammation [5, 6], and potentially to medication used to manage RA, for example NSAIDs [7] or glucocorticoids (GCs).

GCs are frequently prescribed in RA, with up to two-thirds of patients with RA ever prescribed GCs [8, 9]. This reflects their powerful anti-inflammatory effects, yet their use is associated with a wide range of adverse effects, such as fractures, infections, insomnia and weight gain [10]. Another less well studied but widely cited side effect of GCs is hypertension. Hypertension has been captured as one of many adverse events in clinical trials [11–14]. In placebo controlled trials of patients with a variety of rheumatic conditions (RA, polymyalgia rheumatica, GCA) there were 3–28 hypertension events per 100 patient years in those using chronic medium dose GCs (7.5 to <30 mg/day). However, the range of reported hypertension events is wide compared with other GC adverse events [15]. There have been very few studies focussed specifically on GC-induced hypertension in RA. Observational studies specifically investigating hypertension and GC use have had conflicting results: some studies have described medium to high dose GCs being associated with hypertension [16, 17], while other studies found no association [18, 19]. As hypertension may further increase CV risk, it is important to evaluate whether GCs increase the risk of hypertension and if so, how this might relate to dose. Therefore, the aim of this study was to determine whether GCs are associated with increased risk of incident hypertension in a cohort of patients with incident RA.

Methods

Design

This was a retrospective cohort study using data from the Clinical Practice Research Datalink (CPRD), a database of UK primary care electronic medical records. The data covers around 7% of the UK population and it has been shown to be broadly representative of the general population [20]. This study used only data from practices that were considered up to research standard (a CPRD measure indicating when practice data is up to research quality based on mortality rates and continuity of data). The study period was from 1 January 1992 until 31 June 2019. The protocol for this study has been approved by the Independent Scientific Advisory Committee (Protocol number: 11_113RA6).

Study population

All patients with incident RA diagnosed during the study period were identified using a validated algorithm [21]. Patients were excluded if they had a diagnosis of hypertension (criteria for diagnosis described in the outcome section below) before the RA diagnosis date or were aged <18 years at RA diagnosis. Patients were followed up from RA diagnosis until leaving the practice, death or the end of the study period.

Exposure

Oral GC prescriptions were identified through product codes. The data were prepared using a published algorithm [22] and the assumptions made are described in Supplemental Data S1 available at *Rheumatology* online. People were considered GC users for the duration of each prescription. GC dose for each prescription was converted to prednisolone equivalent doses [23]. Dose was then categorized as non-use, >0–4.9, 5–7.4, 7.5–14.9 and ≥15 mg/day. Cumulative dose was calculated by multiplying daily GC dose by the number of days prescribed, and then summing this value for all prescriptions up to that time point. Values were divided by 1000 to give cumulative dose in grams (g) rather than milligrams (mg). Categories of cumulative dose were then defined as non-use, >0 to <2.5, 2.5 to <5, 5 to <10 and ≥10 g.

Outcome

A validated definition of hypertension was used [24] where a person was considered to have hypertension from the earliest of either: (i) two consecutive systolic blood pressure (SBP) readings ≥140 mmHg within a year, (ii) two consecutive diastolic blood pressure (DBP) readings ≥90 mmHg within a year, (iii) a hypertension Read code (see [25] and Supplemental Data S1, available at *Rheumatology* online), and on therapy with antihypertensive medications (angiotensin-converting enzyme inhibitors, alpha blockers, angiotensin receptor blockers, beta blockers, calcium channel blockers and diuretics) prescribed on at least two different dates within 6 months either side of the Read code. For criteria (i) and (ii), a person was considered hypertensive from the second BP reading as a person would not be considered hypertensive based on one BP reading. For criteria (iii), a person was considered hypertensive from the earliest of Read code or antihypertensive prescription start date. Follow-up was censored at the point of hypertension diagnosis.

Confounders

The following covariates were included in the analyses: baseline age; gender; baseline BMI calculated using height and nearest weight measurement (if present within 5 years prior to baseline); baseline smoking status, classified as ever or never using Read codes and smoking cessation prescription codes; time-varying conventional synthetic DMARD use and time-varying prescribed NSAID use, identified using product codes where patients were considered exposed for the duration of their prescription; and Charlson comorbidity index at baseline, determined using a validated algorithm [26], where patients were considered to have the comorbidity if they had a Read code at any point from registration with the practice or up to research standard date, whichever was latest, until baseline. All these covariates were considered *a priori* confounders and were included in the analysis. All code lists can be found in Supplemental Data S1, available at *Rheumatology* online.

Missing data

Baseline BMI and smoking status had 43% and 17% missing data, respectively. Data were imputed using multiple imputation with 47 imputations, this number was based on the fraction of missing information.

Risk attribution model

A risk attribution model was used whereby a person was considered at risk of hypertension for 3 months after the estimated GC, DMARD and NSAID prescription end dates. This allowed for uncertainty around the start and stop dates, infrequent BP assessment and for potential long lasting effects of these drugs. All GC exposure models used this risk attribution model, therefore GC use and GC dose will be described as recent GC use and recent GC dose. In sensitivity analyses the attribution model was explored by running the same analyses with a GC exposure risk attribution model of 1 month and then 6 months, to see if this affected the results.

Analysis

The baseline characteristics of the cohort were described stratified by whether GC was ever prescribed during follow-up. Incidence rates overall and by GC status were calculated. Cox proportional hazards regression models (unadjusted, age and gender adjusted, and adjusted for all confounders) were used to examine whether recent GC use, categories of GC dose and categories of cumulative GC dose were associated with incident hypertension.

Accounting for possible surveillance bias

As hypertension is a potential side effect of GCs, it is plausible that people prescribed GCs may have their BP measured more often than people not prescribed GCs and therefore may have more opportunity for hypertension to be identified (a surveillance bias). To investigate this, the frequency of BP measurements was compared in the first 2 years since diagnosis stratified by the level of GC exposure. As follow-up length varied, follow-up was censored at 2 years or at hypertension diagnosis if this was prior to 2 years to allow comparison between groups. As GC use had been measured in a time-varying manner a summary variable was created to describe level of GC use over the 2 years. GC exposure was classified as 'no GC use', 'intermittent GC use', if they had $<$80% of follow-up with GC use in the first 2 years since diagnosis or 'continuous GC use' if they had $\geq$80% GC use in the first 2 years.

Sensitivity analyses

CPRD data can be linked to secondary care data and area-based datasets where practices consent to linkage, with 58% of all practices currently consenting to linkage [20]. For those practices, data were linked to Hospital Episodes Statistics outpatient data and practice level deprivation data. This allowed additional adjustment for healthcare utilization and socioeconomic status in a subpopulation. Healthcare utilization was measured as a proxy for disease severity where a person was considered to have high disease activity if they had more than three rheumatology outpatient visits per year. Socioeconomic status was measured using quintiles of English Index of Multiple Deprivation (IMD) 2015. Further sensitivity analyses using a stricter definition of hypertension were conducted, where only those with a Read code for hypertension and at least two antihypertensive medication prescriptions within 6 months either side of the Read code were considered hypertensive.

Patient and public involvement

Patients were not involved in the design, conduct or reporting of this study.

Results

Cohort characteristics

Of 31 657 patients with a diagnosis of RA, 13 897 (44%) had hypertension prior to RA diagnosis, resulting in 17 760 patients who were included in this cohort (supplementary Fig. S1, available at *Rheumatology* online). Those included in the cohort had a mean age 56.3 years (S.D. 12.7) and were predominantly female (68%, $N = 12\,101$). Of those, 41.8% ($N = 7421$) were prescribed GCs during follow-up, and these patients were slightly older (mean age 57.7 *vs* 55.3 years of those never prescribed GCs), were predominantly female, had a history of smoking and had more comorbidities compared with those not prescribed GCs during follow-up (Table 1).

There were 6243 cases of incident hypertension over 97 547 person years (pyrs) of follow-up, giving an incident rate of 64.1 per 1000 pyrs (95% CI: 62.5, 65.7). Cases were most frequently first identified through consecutive high SBP measurements alone ($N = 4018$, 64%), followed by consecutive high SBP and DBP measurements ($N = 1134$, 18%) and consecutive high DBP measurements alone ($n = 504$, 8%). Only 7% ($N = 449$) were identified first through antihypertensive prescriptions and Read codes alone (Fig. 1). Of those identified through high BP measurements, 60% ($N = 3396/5656$) were subsequently prescribed antihypertensive medication.

Glucocorticoid association with hypertension

In those exposed to GCs there were 1321 cases of incident hypertension with an incidence rate of 87.6 per 1000 pyrs. In those unexposed there were 4922 cases with an incidence rate of 59.7 per 1000 pyrs. (Table 2).

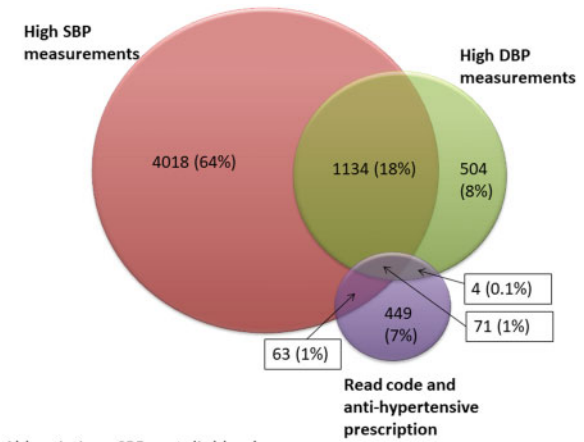
The unadjusted Cox proportional hazards model for recent GC use showed GC use was associated with a 44% increased hazard of hypertension [hazard ratio (HR) 1.44; 95% CI: 1.35, 1.53]; when fully adjusted this was attenuated to 17% increased hazard but remained statistically significant (HR 1.17; 95% CI: 1.10, 1.24). The unadjusted model for categories of recent exposure

Table 1 Baseline characteristics of cohort overall and stratified by glucocorticoid use during follow-up

	Overall	Never prescribed GCs during follow-up	Ever prescribed GCs during follow-up
N	17 760	10 339 (%)	7421 (%)
Baseline age [mean (S.D.)]	56.31 (12.7)	55.31 (12.4)	57.72 (13.1)
Female gender (%)	12 101 (68.1)	7139 (69.0)	4962 (66.9)
Baseline ever smoker (%)[a]	8817 (60.0)	4936 (57.5)	3881 (63.4)
Baseline BMI [mean (S.D.)][a]	26.89 (5.45)	26.95 (5.44)	26.79 (5.47)
Baseline BMI category (%)			
Underweight	219 (2.2)	104 (1.8)	115 (2.7)
Normal	3864 (38.8)	2238 (38.7)	1626 (38.8)
Overweight	3541 (35.5)	2072 (35.8)	1469 (35.1)
Obese	2084 (20.9)	1217 (21.0)	867 (20.7)
Morbidly obese	261 (2.6)	152 (2.6)	109 (2.6)
Baseline Charlson comorbidity index (%)			
0	13 760 (77.5)	8435 (81.6)	5325 (71.8)
1	2845 (16.0)	1333 (12.9)	1512 (20.4)
2	786 (4.4)	388 (3.8)	398 (5.4)
3 or more	369 (2.1)	183 (1.8)	186 (2.5)
IMD quintile (%)[a]			
1	1415 (15.4)	755 (15.1)	660 (15.7)
2	1765 (19.2)	960 (19.2)	805 (19.1)
3	1872 (20.3)	1009 (20.2)	863 (20.5)
4	1920 (20.9)	1059 (21.2)	861 (20.4)
5	2233 (24.3)	1206 (24.2)	1027 (24.4)
GC use prior to RA diagnosis (%)	3383 (19.0)	628 (6.1)	2755 (37.1)
Cumulative GC dose in year prior to baseline [mean (S.D.)]	334.75 (1242.3)	55.80 (366.6)	723.37 (1802.0)

[a]There were missing data for the following variables: ever smoking: $N = 3057$ (17.2%); baseline BMI: $N = 7791$ (43.9%); IMD 2010: $N = 8555$ (48.2%). GC: glucocorticoid; IMD: English Index of Multiple Deprivation.

Fig. 1 Venn diagram showing how hypertension was identified



Abbreviations: SBP: systolic blood pressure, DBP: diastolic blood pressure.

dosage showed all GC dosage categories were associated with hypertension. When fully adjusted, only doses of ≥7.5 mg were statistically significant, indicating increased hazard of hypertension (7.5–14.9 mg: HR 1.18; 95% CI: 1.08, 1.29; ≥15 mg: HR 1.36; 95% CI: 1.18, 1.56). Doses <7.5 mg had increased hazard but were not statistically significant. The unadjusted model for categories of cumulative dose showed all categories were significantly associated with hypertension, but when fully adjusted there was no clear pattern. Only the category of 5–9.99 g was statistically significant, though ≥10 g had a similar point estimate (Table 3). Point estimates for the covariates in the adjusted models were in the expected direction, with leflunomide having the biggest effect and NSAIDs having a similar magnitude of effect on hypertension as recent GC use (supplementary Table S1, available at *Rheumatology* online).

Possible surveillance bias

When the cohort follow-up was censored to 2 years, most patients (73%) had at least 2 years' follow-up. The majority of the cohort did not use GCs during this period ($n = 12\,124$, 68.3%), 3461 (19.5%) had intermittent use and 2175 (12.3%) had continuous use. There were no differences in the frequency of BP measurements between the groups (Table 4 and Fig. 2), suggesting that surveillance bias was not present.

Sensitivity analyses

There were 5860 patients with linkage to Hospital Episodes Statistics outpatient data, of whom 1487 developed incident hypertension giving an incident rate of 59.9 per 1000 pyrs (95% CI: 57.0, 63.0). Additional adjustment for our proxy for disease activity and IMD

Table 2 Number of cases and rate of hypertension by GC status

	Exposed to GCs	Unexposed to GCs	Overall
Total number[a]	7421	16 850	17 760
Follow-up time (days)	15 076	82 382	97 457
Cases of hypertension	1321	4922	6243
Incident rate, per 1000 person-years (95% CI)	87.6 (83.0, 92.4)	59.7 (58.1, 61.4)	64.1 (62.5, 65.7)

[a]As GC use is time-varying people could be in both categories, therefore total number across both categories is greater than the total number of people in the study. GC: glucocorticoid.

Table 3 Unadjusted and adjusted Cox proportional hazards regression model

	Unadjusted [HR (95% CI)]	Age and gender adjusted [HR (95% CI)]	Fully adjusted[a] [HR (95% CI)]
Recent GC use	1.44 (1.35, 1.53)	1.23 (1.16, 1.31)	1.17 (1.10, 1.24)
Recent GC dose			
No GC use	Reference	Reference	Reference
>0–4.9 mg	1.35 (1.21, 1.53)	1.13 (1.01, 1.28)	1.10 (0.98, 1.24)
5–7.4 mg	1.40 (1.22, 1.60)	1.11 (0.97, 1.27)	1.07 (0.93, 1.23)
7.5–14.9 mg	1.44 (1.33, 1.57)	1.26 (1.16, 1.38)	1.18 (1.08, 1.29)
≥15 mg	1.60 (1.40, 1.84)	1.45 (1.27, 1.66)	1.36 (1.18, 1.56)
Cumulative dose			
No GC use	Reference	Reference	Reference
>0–2.49 g	1.14 (1.05, 1.23)	1.04 (0.96, 1.12)	1.00 (0.92, 1.08)
2.5–4.99 g	1.16 (1.06, 1.27)	1.04 (0.95, 1.13)	0.99 (0.90, 1.08)
5–9.99 g	1.36 (1.24, 1.48)	1.18 (1.08, 1.30)	1.12 (1.02, 1.22)
≥10 g	1.35 (1.24, 1.49)	1.16 (1.06, 1.27)	1.07 (0.97, 1.17)

[a]Adjusted for baseline age, gender, baseline BMI, baseline ever smoking, Charlson comorbidity index, time-varying synthetic DMARD use and time-varying NSAID use. HR: hazard ratio; GC: glucocorticoid.

2015 did not substantively change the results: the recent GC use HR was slightly lower (HR 1.14; 95% CI: 1.00, 1.29) and only doses ≥15 mg were statistically significant. Though the dose category 7.5–14.9 mg just missed significance, this was the same regardless of the additional adjustment for disease activity and IMD 2015 (supplementary Table S2, available at *Rheumatology* online). When the attribution window was increased to 6 months the results were broadly similar (supplementary Table S3, available at *Rheumatology* online). When the attribution window was reduced to 1 month the results were broadly similar, though the lowest category of GC dose (>0–4.9 mg) was just statistically significant (HR:1.16; 95% CI: 1.02, 1.31) (supplementary Table S4, available at *Rheumatology* online). There were 2002 cases of hypertension using the strict hypertension definition (two or more antihypertensive prescriptions within 6 months either side of a Read code). Although there were only 449 patients initially identified through this strict definition, many of those who were first identified through BP measurements alone later went on to meet the criteria using the strict definition. The results using this strict definition of hypertension were similar, the HR was slightly lower for recent GC use (HR 1.13; 95% CI: 1.01, 1.27). Doses >7.5 mg were not statistically significant, although they remained in the direction of increased risk (supplementary Table S5, available at *Rheumatology* online).

Discussion

This study found that GC use was associated with a 17% overall increased risk of hypertension in patients with incident RA and without hypertension at RA diagnosis. When GC use was stratified by dose categories, doses <7.5 mg were not found to be associated with hypertension, indicating that low doses were less of a concern, although the point estimates were in the direction of increased risk for all categories of GC dose. There was no clear pattern seen for cumulative dose, but this may be due to the nature of the measure itself, as a small cumulative dose may represent a person prescribed a low dose for a long period or a person prescribed a high dose for a short period, making it difficult to draw conclusions in terms of the entire exposed period. Additionally, 40% of patients prescribed GCs with hypertension (defined by consecutive high SBP or DBP readings) were not prescribed an antihypertensive at any point during the study duration. Whilst some may have been offered lifestyle advice, left untreated this has important implications in terms of addressing modifiable

FIG. 2 Number of blood pressure measurements over 2 years, by glucocorticoid use category

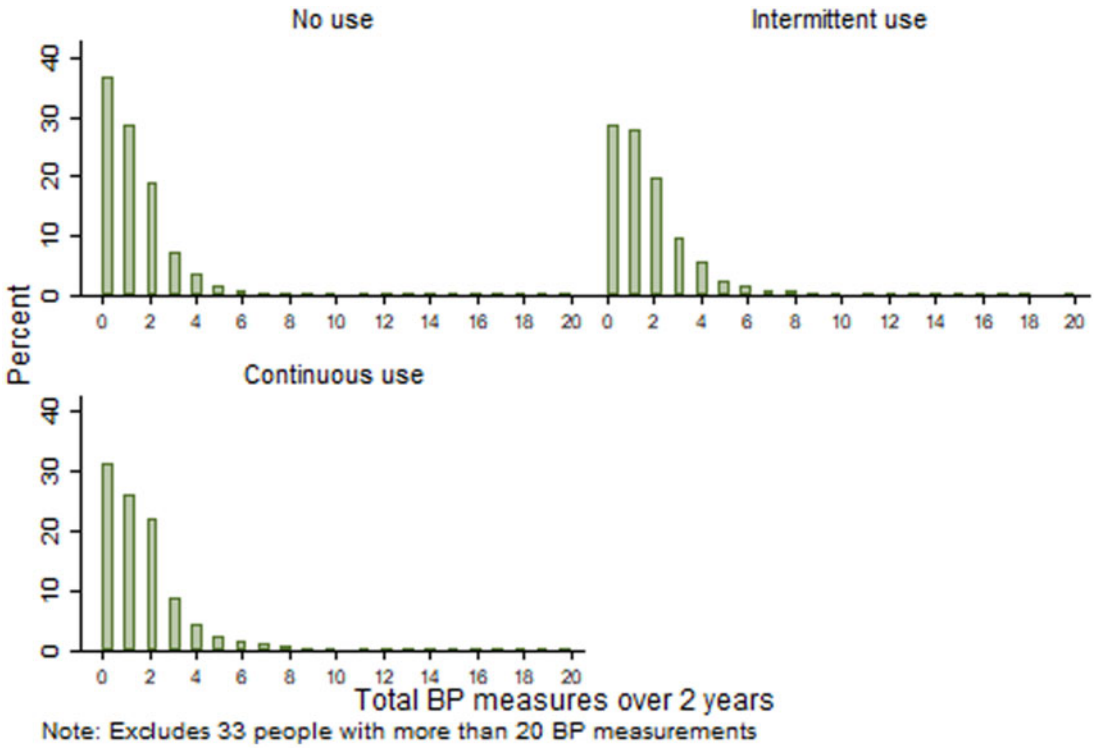


TABLE 4 Frequency of blood pressure measurements by categories of GC use over 2 years

GC use category	*N* (%)	At least 1 BP measurement [*n* (%)]	Median number of measurements (IQR)	More than 2 BP measurements [*n* (%)]	Maximum number of measurements
No use	12 124 (68.3)	7714 (65.6)	1 (0–2)	1995 (16.5)	34
Intermittent use	3461 (19.5)	2477 (71.6)	1 (0–2)	841 (24.3)	39
Continuous use	2175 (12.3)	1492 (68.6)	1 (0–2)	448 (20.6)	25

GC: glucocorticoid; BP: blood pressure; IQR: interquartile range.

risk factors in an RA population already at increased risk of CV disease.

Differences in the frequency of BP measurement by GC exposure were not seen, providing reassurance that surveillance bias does not explain the findings. Importantly, around 30% of the cohort did not have their BP measured during the first 2 years after diagnosis. EULAR recommends monitoring and treatment of CV risk factors in RA [27] and hypertension in GC-treated patients [15]. This study highlights that this may not be the case overall in RA with regards to monitoring and treating high BP in primary care. Given this finding, it is important for primary care physicians (and rheumatologists) to be aware that GCs increase the risk of hypertension, and to monitor patients' BP more vigilantly while GCs are prescribed.

Previous studies

These results concur with a single-centre cross-sectional study, where long-term (<6 months use) medium dose (≤7.5 mg) prednisolone was associated with hypertension [16], and a study of patients in a German registry where patients who were prescribed GC doses >7.5 mg for >6 months had higher proportions of self-reported 'increase in blood pressure' [17]. However, our results do not concur with another study that used CPRD data to investigate adverse effects associated with GC use, including hypertension. They did not find an association between GC use and hypertension; however, only a Read code was used to identify hypertension, so cases may have been missed and may explain why their results were different from this study [18].

Incidence of GC-associated hypertension

This study provides an estimate of incidence of hypertension associated with GC use, which allows more informed decisions for the patient. A UK study using primary care electronic records has estimated the incidence of hypertension in patients with RA [28]. This study found a lower incident rate of hypertension, 336.2 per 10 000 pyrs, and a higher proportion being treated (85%) compared with our study (60%). However, this study only identified hypertension using Read codes and/or antihypertensive prescriptions, which means patients with high BP but not coded or treated are missed, which may explain the differences found compared with our study.

Strengths and limitations

This was a large retrospective cohort study using routinely collected data with a number of strengths. The use of prescription data allowed more precise measurement of time-varying GC use, and a variety of attribution models were used to test the impact of our assumptions when preparing the data. Hypertension diagnosis has not been consistently defined across the few studies using CPRD data, and in our study hypertension was identified through BP measurements or a Read code and antihypertensive prescriptions. This definition has been validated in Spanish primary care electronic health records [24] and allowed a more robust identification of the outcome. As anti-hypertensive medication can be prescribed for other indications, it was important to use both Read code for hypertension and antihypertensive medication prescriptions to ensure antihypertensive medication was not prescribed for another indication.

Alongside these strengths there are some limitations. Misclassification of medication use is a possibility; as CPRD data only contains prescriptions, we do not know if these medications were dispensed. However, we used a number of attribution models to allow for potential differences in when prescriptions would be dispensed. This study was designed specifically to examine incident hypertension and thus included only patients without prior hypertension. Further work is needed to understand how GCs may affect BP in those already diagnosed with hypertension. Although we need to be careful of over-interpretation of covariate point estimates [29], the variables adjusted for were in the expected direction. However, there are some variables that cannot be measured in CPRD: disease severity is not available. However, currently there is no evidence that high disease activity is associated with high BP, suggesting that confounding by indication is less of a concern [30, 31]. There is not a validated proxy for disease severity in CPRD; however, we have conducted a sensitivity analysis using a pragmatic proxy for disease severity and this did not alter the results. As biologics are prescribed in secondary care this is not well captured in CPRD. TNF inhibitors have been shown to reduce BP [11]; however, it has been shown that those prescribed biologics are more likely to have received GCs [32]. As we would expect GCs to increase BP, if TNF inhibitors are prescribed more frequently in those prescribed GCs we would expect the effect of GCs on BP to be underestimated. Therefore any unmeasured confounding would not explain our positive findings.

Conclusions

This study found that GC use was associated with incident hypertension in patients with RA, and in particular doses >7.5 mg were associated with hypertension. There was an incidence rate of 64.1 per 1000 pyrs. BP was not frequently monitored in primary care and a large proportion of RA patients on GCs with high BP readings were untreated. Given that patients with RA are already at increased risk of CV disease, it is important that these patients should have their BP checked regularly and treated appropriately.

Funding: This work was supported by the Centre for Epidemiology Versus Arthritis (Grant number 21755) and supported by the National Institute for Health Research Manchester Biomedical Research Centre.

Disclosure statement: W.G.D. has received consultancy fees from Google and Beyer unrelated to this work. The other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Symmons D, Turner G, Webb R *et al*. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002;41:793–800.

2 Dadoun S, Zeboulon-Ktorza N, Combescure C *et al*. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. Jt Bone Spine 2013;80:29–33.

3 Aviña-Zubieta JA, Choi HK, Sadatsafavi M *et al*. Risk of cardiovascular mortality in patients with rheumatoid arthritis: a meta-analysis of observational studies. Arthritis Rheum 2008;59:1690–7.[10.1002/art.24092]

4 Meune C, Touze E, Trinquart L, Allanore Y. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. Rheumatology (Oxford) 2009; 48:1309–13.

5 Innala L, Moller B, Ljung L *et al*. Cardiovascular events in early RA are a result of inflammatory burden and traditional risk factors: a five year prospective study. Arthritis Res Ther 2011;13:R131.

6 Naranjo A, Sokka T, Descalzo MA *et al*. Cardiovascular disease in patients with rheumatoid arthritis: results from the QUEST-RA study. Arthritis Res Ther 2008;10:R30.

7 Agca R, Heslinga SC, Rollefstad S *et al*. EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. Ann Rheum Dis 2017;76:17–28.

8 Sokka T, Toloza S, Cutolo M *et al*. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA Study. Arthritis Res Ther 2009;11:R7.

9 Black RJ, Joseph RM, Brown B *et al*. Half of U.K. patients with rheumatoid arthritis are prescribed oral glucocorticoid therapy in primary care: a retrospective drug utilisation study. Arthritis Res Ther 2015;17:375.

10 Hoes JN, Jacobs JWG, Buttgereit F, Bijlsma J. Current view of glucocorticoid co-therapy with DMARDs in rheumatoid arthritis. Nat Rev Rheumatol 2010;6: 693–702.

11 Klarenbeek NB, van der Kooij SM, Huizinga TJ *et al*. Blood pressure changes in patients with recent-onset rheumatoid arthritis treated with four different treatment strategies: a post hoc analysis from the BeSt trial. Ann Rheum Dis 2010;69:1342–5.

12 Hafstrom I, Rohani M, Deneberg S *et al*. Effects of low-dose prednisolone on endothelial function, atherosclerosis, and traditional risk factors for atherosclerosis in patients with rheumatoid arthritis–a randomized study. J Rheumatol 2007;34:1810–6.

13 van Everdingen AA, Jacobs JW, Siewertsz Van Reesema DR, Bijlsma JW. Low-dose prednisone therapy for patients with early active rheumatoid arthritis: clinical efficacy, disease-modifying properties, and side effects: a randomized, double-blind, placebo-controlled clinical trial. Ann Intern Med 2002;136:1–12.

14 Capell HA, Madhok R, Hunter JA *et al*. Lack of radiological and clinical benefit over two years of low dose prednisolone for rheumatoid arthritis: results of a randomised controlled trial. Ann Rheum Dis 2004;63:797–803.

15 Duru N, van der Goes MC, Jacobs JWG *et al*. EULAR evidence-based and consensus-based recommendations on the management of medium to high-dose glucocorticoid therapy in rheumatic diseases. Ann Rheum Dis 2013;72:1905–13.

16 Panoulas VF, Douglas KM, Stavropoulos-Kalinoglou A *et al*. Long-term exposure to medium-dose glucocorticoid therapy associates with hypertension in patients with rheumatoid arthritis. Rheumatology (Oxford) 2008;47:72–5.

17 Huscher D, Thiele K, Gromnica-Ihle E *et al*. Dose-related patterns of glucocorticoid-induced side effects. Ann Rheum Dis 2009;68:1119–24.

18 Wilson JC, Sarsour K, Gale S *et al*. Incidence and risk of glucocorticoid-associated adverse effects in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2019;71:498–511.

19 Jackson SH, Beevers DG, Myers K. Does long-term low-dose corticosteroid therapy cause hypertension? Clin Sci 1981;61:381–3.

20 Herrett E, Gallagher AM, Bhaskaran K *et al*. Data resource profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol 2015;44:827–36.

21 Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? Arthritis Rheum 2008;59: 1314–21.

22 Pye SR, Sheppard T, Joseph RM *et al*. Assumptions made when preparing drug exposure data for analysis have an impact on results: an unreported step in pharmacoepidemiology studies. Pharmacoepidemiol Drug Saf 2018;27:781–8.

23 NICE. Glucocorticoid therapy [Internet]. Treatment summary. https://bnf.nice.org.uk/treatment-summary/glucocorticoid-therapy.html (19 March 2020, date last accessed).

24 de Burgos-Lunar C, Salinero-Fort MA, Cárdenas-Valladolid J *et al*. Validation of diabetes mellitus and hypertension diagnosis in computerized medical records in primary health care. BMC Med Res Methodol 2011; 11:146.

25 NICE. Quality and Outcomes Framework guidance for GMS contract 2013/14. London: NHS Employers, 2013; March:67–85.

26 Khan NF, Perera R, Harper S, Rose PW. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC Fam Pract 2010;11:1.

27 van der Goes MC, Jacobs JW, Boers M, Andrews T *et al*. Monitoring adverse events of low-dose glucocorticoid therapy: EULAR recommendations for clinical trials and daily practice. Ann Rheum Dis 2010;69:1913–9.

28 Jafri K, Bartels CM, Shin D, Gelfand JM, Ogdie A. Incidence and management of cardiovascular risk factors in psoriatic arthritis and rheumatoid arthritis: a population-based study. Arthritis Care Res (Hoboken) 2017;69:51–7.

29 Westreich D, Greenland S. The Table 2 Fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol 2013;177:292–8.

30 Siebert S, Lyall DM, Mackay DF *et al*. Characteristics of rheumatoid arthritis and its association with major comorbid conditions: cross-sectional study of 502 649 UK Biobank participants. RMD Open 2016;2: e000267.

31 Panoulas VF, Douglas KM, Milionis HJ *et al*. Prevalence and associations of hypertension and its control in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007;46:1477–82.

32 Low ASL, Symmons DPM, Lunt M *et al*. Relationship between exposure to tumour necrosis factor inhibitor therapy and incidence and severity of myocardial infarction in patients with rheumatoid arthritis. Ann Rheum Dis 2017;76:654–60.

Extract of supplemental data 1:

Drug preparation algorithm

Decisions used for glucocorticoids, disease-modifying antirheumatic drugs:

- 4: Implausible quantity: set to population average
- 9: Missing quantity: set to population average
- 15: Implausible ndd: set to population average
- 20: Missing ndd: set to population average
- 27: Clean duration: set to 6 months if > 6 months
- 34: Select stop date if multiple: If one available use it, if two available and equal use that date, if 2 available uses mean, if 3 available uses mean of closest 2 if within 30 days.
- 41: Missing stop dates: use individual mean if unavailable use population mean
- 43: Multiple prescriptions for same product on same day: use mean ndd and mean length
- 50: Overlapping prescriptions: move later to next available time
- 53: Sequential prescriptions with short gaps: change stop gap to start of next prescriptions if gap is <= 30 days.

Decisions used for NSAIDs:

- 4: Implausible quantity: set to population average
- 9: Missing quantity: set to population average
- 15: Implausible ndd: set to population average
- 20: Missing ndd: set to population average
- 27: Clean duration: set to 6 months if > 6 months
- 34: Select stop date if multiple: If one available use it, if two available and equal use that date, if 2 available uses mean, if 3 available uses mean of closest 2 if within 30 days.
- 41: Missing stop dates: use individual mean if unavailable use population mean
- 43: Multiple prescriptions for same product on same day: use mean ndd and mean length
- 49: Overlapping prescriptions: do nothing, allow to overlap
- 53: Sequential prescriptions with short gaps: change stop gap to start of next prescriptions if gap is <= 30 days.

Publication 6: Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community.

BMJ Open

Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community

Ruth Costello,[1] Rikesh Patel,[1] Jennifer Humphreys,[1] John McBeth,[1,2] William G Dixon[1,2,3]

To cite: Costello R, Patel R, Humphreys J, *et al*. Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community. *BMJ Open* 2017;**7**:e014603. doi:10.1136/bmjopen-2016-014603

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2016-014603).

Received 6 October 2016
Revised 21 December 2016
Accepted 10 January 2017

For numbered affiliations see end of article.

Correspondence to
Dr William G Dixon;
Will.Dixon@manchester.ac.uk

ABSTRACT

Objectives: To identify the side effects most important to glucocorticoid (GC) users through a survey of a UK online health community (Healthunlocked.com).

Design: Online cross-sectional survey.

Setting: Participants were recruited through Healthunlocked.com, an online social network for health.

Participants: Adults who were currently taking GCs, or had taken GCs in the past month.

Method: Responders scored the importance of listed side effects from 1 to 10, with 10 being of high importance to them. For each side effect, histograms were plotted, and the median rating and IQR were determined. Side effects were ranked by median ranking (largest to smallest) and then IQR (smallest to largest). The scores were categorised as low (scores 1–3), medium (scores 4–7) and high (scores 8–10) importance.

Results: 604 responders completed the survey. Histograms of side effect scores showed a skew towards high importance for weight gain, a U-shaped distribution for cardiovascular disease (CVD), diabetes, eye disease and infections, and a skew towards low importance for acne. When ranked, the side effect of most importance to responders was weight gain (median score=9, IQR 6–10) followed by insomnia and moon face with equal median score (8) and IQR (5–10). Three serious side effects, CVD, diabetes and infections, were ranked of lower importance overall but had wide ranging scores (median score=8, IQR 1–10).

Conclusions: The three most highly rated side effects were not clinically serious but remained important to patients, perhaps reflecting their impact on quality of life and high prevalence. This should be taken into consideration when discussing treatment options and planning future GC safety studies.

Strengths and limitations of this study

- This survey used a novel recruitment method, through an online social network for health, which resulted in over 600 UK respondents who were taking glucocorticoids for a variety of conditions.
- Only a few studies have previously investigated which glucocorticoid side effects are most important to patients.
- The sample was mainly female and over 50 years of age, which may represent bias in the type of people who participate in studies.

INTRODUCTION

Glucocorticoids (GC) continue to be widely used to treat inflammatory diseases since their discovery over 60 years ago.[1] In the UK, around 1% of the population have been prescribed oral GCs, most commonly in the context of respiratory disease.[2] For certain conditions, such as vasculitis, systemic lupus erythematosus and polymyalgia rheumatica, GCs are used in nearly all patients.[3 4]

GCs have many side effects, ranging from potentially life-threatening such as cardiovascular events and infections,[5–7] to less clinically serious effects such as bruising, skin thinning and fat redistribution. Understandably, research to date has focused more on the serious side effects, but these 'less serious' side effects may be important to the patient and have the potential to markedly impair a patient's quality of life. Furthermore, patients may elect not to take GC therapy because of concerns about possible side effects. To date, only a few studies have investigated which side effects are important to patients.[8–10] Although osteoporosis was in the top three most important side effects in two of the three studies, the findings in general have not been consistent. For example, in one study 'diabetes/glucose intolerance' was ranked third most important,[8] while in another 'trouble with blood glucose levels/diabetes' was 12th of side effects that bothered patients a lot.[10] Two of these studies were in patients with specific diseases, adrenal insufficiency (where GCs are used to replace deficient endogenous GCs)[9] and

immune thrombocytopenic purpura (ITP),[10] while the third studied patients with rheumatic diseases.[8] Observed differences between these studies might be explained by the use of GC therapy for the treatment of inflammatory disease versus replacement therapy. To understand which side effects are most important to patients across several disease groups, the aim of this study was to identify the side effects most important to GC users through a survey of multiple disease communities within a UK online health social media platform (Healthunlocked.com (HU)).

METHODS

Setting

HU is a social network for health where patients, caregivers and health advocates can discuss issues related to their health through online message boards and private messages. Discussions take place within communities set up by patient charities and condition communities from NHS Choices. It is the largest health-related social network in Europe with 4 million visitors per month. The HU platform allows rapid access to hundreds of potential GC users by embedding a survey in posts with a particular title word, or tagged with a given word or phrase.

Design

A short survey about GC use, timing of GC administration and perceptions of side effects was designed by the research team specifically for this study. The survey was drafted by WD to include information about people's beliefs about the importance of range of known serious and non-serious GC-associated side effects. The number of items was selected to balance the burden of data entry with collecting opinion on a range of side effects, including items scored of high, intermediate and low importance in previous studies.[8–10] The draft was further refined with input from rheumatologists, endocrinologists and epidemiologists (RP, JH, JMcB, RC plus wider consultation with local colleagues (see acknowledgements)). This resulted in some rewording of questions and two additional side effects were added: 'changes in mood' and 'round face or "moon" face'. The survey was then piloted with 13 members of an existing musculoskeletal Research User Group (RUG), comprised of patients with musculoskeletal disease and their carers who meet quarterly to help support research studies. RUG members were asked to comment on comprehension, ease of completion and provide any general feedback. The survey was finalised based on the pilot testing responses. No additional GC-associated adverse events were suggested for inclusion in the survey by the patient group (see online supplementary material S1). The testing supported our decision to ask participants to score rather than rank each item. One reviewer commented, "I always find it hard to do the thing where they ask you to rank items—in this case, rank this list of side effects from highest to lowest (importance to you) and so I think the system you have used is better. And anyway, a heart attack is surely never going to anywhere other than at the top of the list of undesirable outcomes."

The survey popped up on HU posts that included either the title word 'steroid' or the tags 'glucocorticoid', 'prednisolone', 'prednisone', 'steroid' or 'dexamethasone' and was restricted to UK users. When the survey popped up, the community group for the post was recorded automatically for each responder. To avoid recall bias, only respondents who were currently using, or had used GCs in the last month were eligible. To determine eligibility, a stem question asked whether the respondent was currently taking oral steroids, or had taken oral steroids within the last month. If the response was 'No', the survey ended. If the response was 'Yes', the survey continued. The survey started in December 2015 was live for 3 months or until 1000 surveys were completed, whichever came first. No formal sample size was calculated. Recruitment targets were instead based on discussions with HU about anticipated response rates over a 3-month period.

The perception of GC side effects was examined by asking respondents, 'Please score each side effect, even if you have not experienced it, on a scale where 1= very little importance and 10= high importance to you'. Side effects were listed alphabetically as follows: acne, cardiovascular disease (eg, heart attack), changes in mood, diabetes, eye disease (cataracts, glaucoma), high blood pressure, indigestion, infection (eg, pneumonia), insomnia (unable to sleep), palpitations (racing heart), reduced bone strength (osteoporosis, fractures), round face or 'moon' face, skin changes (bruising, thin skin, stretch marks) and weight gain. Experience of side effects was examined by asking respondents, 'Have you had any of these side effects whilst taking steroids?' Respondents could indicate any that applied.

Statistical analysis

The scores for each side effect were plotted on histograms, and the median score and IQR was determined. Side effects were ranked by median score (largest to smallest) and then IQR (smallest to largest) for those with the same median, to identify the most important side effects to patients. The scores for each side effect were categorised as low importance (scores 1–3), medium importance (scores 4–7) and high importance (scores 8–10). Side effect scores were then stratified by community group and experience of side effects. Median side effect scores and IQR, stratified by experience, were displayed in a box and whisker plot. Respondents with missing data for side effect scores were not included in the analysis.

RESULTS

Patient characteristics

The survey was live for 3 months, it popped up for 17 999 visitors, and 1311 (7.1%) clicked on the survey.

Of those, 756 (58%) agreed to take part in the survey, 664 (51%) were eligible and 604 (46%) provided complete data (see online supplementary figure S1).

Patients came from five community groups: British Lung Foundation (BLF) (N=54), ITP support (N=17), Lupus UK (N=82), National Rheumatoid Arthritis Society (NRAS) (N=229) and Polymyalgia Rheumatica and Giant Cell Arteritis UK (PMRGCAUK) (N=221). The majority of completers were over 50 years old (81%) and women (86%). Those who dropped out part way through the survey (n=60) were not significantly different from those who completed the survey in terms of age, gender and community (table 1).

Survey responses

Figure 1 shows histograms of scores for each side effect. Comparing across histograms, weight gain scores show a pronounced skew towards high importance. Cardiovascular disease (CVD), diabetes, eye disease and infections scores have a U-shaped distribution of scores. Acne scores show a pronounced skew towards low importance.

When ranked, weight gain was the side effect of most importance (median score =9, IQR 6–10), with 64% of weight gain scores categorised as high importance. Insomnia, moon face, high blood pressure (BP), reduced bone strength, eye disease, CVD, diabetes and infection all had the same median score of 8; however, the range of scores varied. Insomnia and moon face were ranked joint second as they had the smallest range of scores (IQR 5–10). Insomnia, like weight gain, had only 12% of respondents who rated it as low importance, whereas all other side effects with a median of 8 were rated as low importance by at least 20% of participants. Side effects with a median score below 8 had <50% of scores categorised as high importance (table 2).

When stratified by community group the rankings remained similar to the overall rankings for all communities except the PMRGCAUK community group, where the side effects most important to respondents were eye disease, CVD and insomnia, with weight gain fourth (table 3).

When stratified by prior experience, participants who had previously experienced the side effect of interest reported higher median scores, with smaller IQRs. The side effects most important to those who had experienced them were diabetes, eye disease and CVD, all scoring a median of 10. The side effects most important to those who had not experienced them were reduced bone strength, CVD and eye disease (table 4, see online supplementary figure S2). Although weight gain had the highest rank overall, it was ranked only fourth in those who had and eighth in those who had not experienced it prior to completing the survey, with median scores and IQRs of 9 (7–10) and 6 (2–9), respectively. The most commonly experienced side effects were, in order, weight gain, round face, insomnia, changes in mood, skin changes and indigestion, all of which were experienced by over half of the 604 respondents.

Table 1 Characteristics of survey responders who completed the survey and those who dropped out during the survey (N=664)

	Completed survey (n=604) N (%)	Dropped out during survey (n=60) N (%)
Community group		
BLF	54 (8.9)	10 (16.7)
ITP support	17 (2.8)	1 (1.7)
Lupus UK	82 (13.6)	6 (10)
NRAS	229 (37.9)	22 (36.7)
PMRGCAUK	221 (36.6)	19 (31.7)
Missing	1 (0.2)	2 (3.3)
Age (years)		
Under 39	40 (6.6)	4 (6.7)
40–49	77 (12.7)	5 (8.3)
50–59	201 (33.3)	14 (23.3)
60–69	181 (30)	19 (31.7)
70 years or over	105 (17.4)	16 (26.7)
Missing	0 (0)	2 (3.3)
Gender		
Male	79 (13.1)	6 (10)
Female	522 (86.4)	49 (81·7)
Missing	3 (0.5)	5 (8·3)
Total	604	60

BLF, British Lung Foundation; ITP, immune thrombocytopenia; NRAS, National Rheumatoid Arthritis Society; PMRGCAUK, Polymyalgia Rheumatica and Giant Cell Arteritis UK.

DISCUSSION

It is known that oral GCs have many side effects, but few studies have investigated which matter the most to patients. This survey found that overall weight gain, insomnia and moon face were the side effects ranked highest by patients, despite them being less clinically serious. The importance of side effects to respondents was different depending on whether they had been experienced, with clinically serious side effects (diabetes, eye disease and CVD) being most important to respondents who had experienced them. As these clinically serious side effects had not been experienced by the majority of respondents, they dropped in the rankings overall. Weight gain, scored at 9 out of 10 for those who had experienced it and 6 out of 10 for those who had not, ranking at fourth position in both groups, but rose to the top ranking overall because of its high prevalence having been experienced by 442/604 (73%) participants. Participants from the PMRGCAUK community rated eye disease as most important, with CVD second and insomnia and weight gain joint third. This contrasted to all other communities where weight gain was the most important side effect overall. This group may be taking a higher dose of GC, compared with the other communities, which may explain the difference. Alternatively, respondents from this community may be

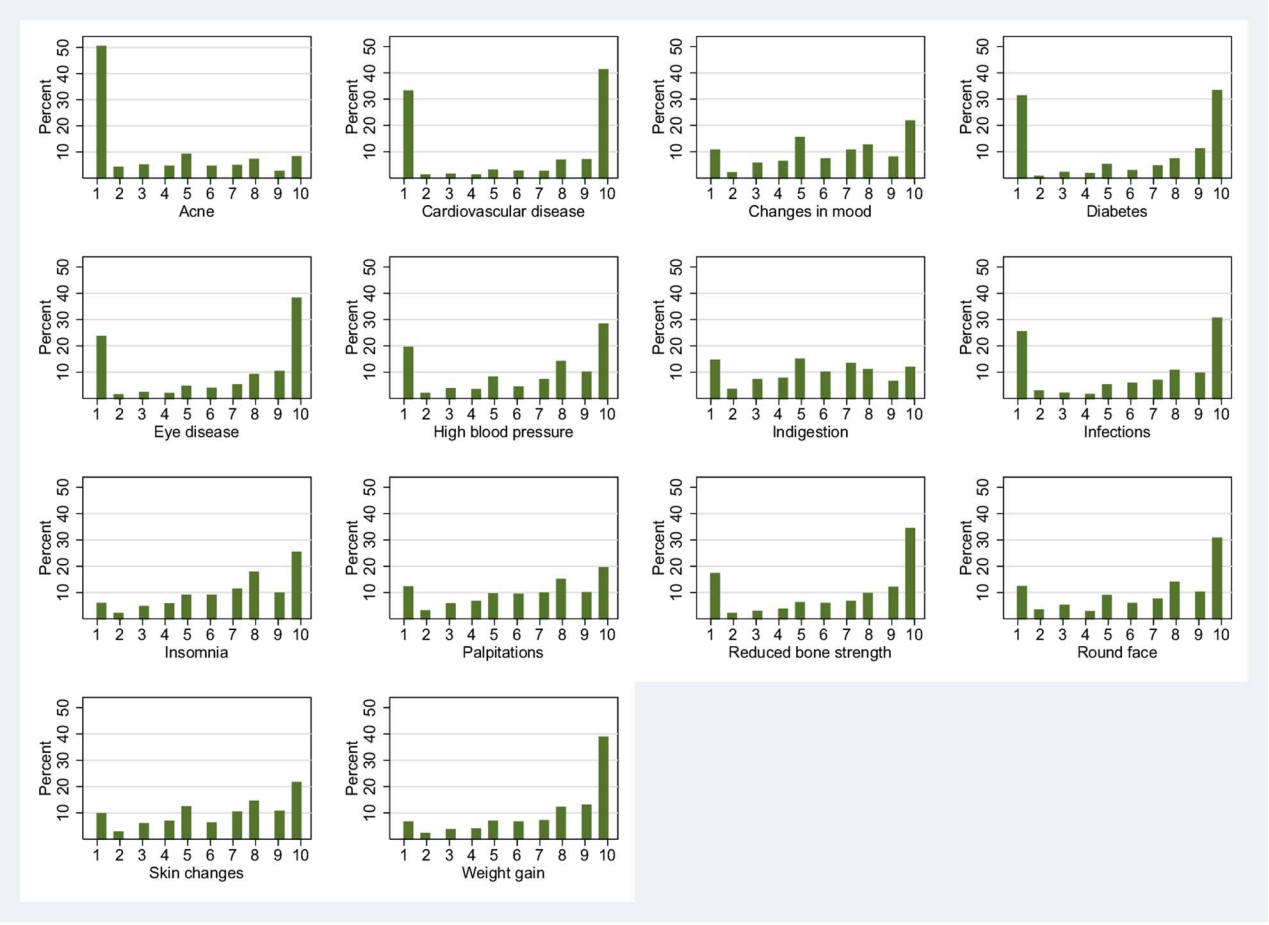


Figure 1 Histograms of side effect ratings. 1= rating of lowest importance, 10= rating of highest importance.

Table 2 Median, IQR, rank and categories of side effect scores

Symptom	Median (IQR)	Rank	Low (score 1–3) N (%)	Medium (score 4–7) N (%)	High (score 8–10) N (%)
Weight gain	9 (6–10)	1	74 (12.3)	145 (24)	385 (63.7)
Insomnia	8 (5–10)	2	75 (12.4)	210 (34.8)	319 (52.8)
Moon face	8 (5–10)	2	125 (20.7)	149 (24.7)	330 (54.6)
High blood pressure	8 (4–10)	4	150 (24.8)	138 (22.8)	316 (52.3)
Reduced bone strength	8 (4–10)	4	133 (22)	134 (22.2)	337 (55.8)
Eye disease	8 (3–10)	6	164 (27.2)	93 (15.4)	347 (57.5)
Cardiovascular disease	8 (1–10)	7	216 (35.8)	56 (9.3)	332 (55)
Diabetes	8 (1–10)	7	206 (34.1)	86 (14.2)	312 (51.7)
Infections	8 (1–10)	7	182 (30.1)	116 (19.2)	306 (50.7)
Changes in mood	7 (5–9)	10	110 (18.2)	239 (39.6)	255 (42.2)
Skin changes	7 (5–9)	10	109 (18)	214 (35.4)	281 (46.5)
Palpitations	7 (4–9)	12	125 (20.7)	212 (35.1)	267 (44.2)
Indigestion	6 (3–8)	13	152 (25.2)	276 (45.7)	176 (29.1)
Acne	1 (1–6)	14	359 (59.4)	138 (22.8)	107 (17.7)

older, and thus could be more concerned about diseases more prevalent at this higher age. Awareness of potential ocular involvement of giant cell arteritis (GCA) may also make the possible occurrence of further eye disease particularly concerning.

Clinicians and patients make treatment decisions after weighing the benefits against the possible harms, and for each benefit or harm, considering its probability, its nature, and a value judgement of how important it is to the individual.[11] While many studies have estimated the frequency of side effects, few have considered how important they are to patients.[12–15] This is relevant because patients' value judgements about a given side effect will influence their decisions about treatment and

Table 3 Median, IQR and rank of side effect scores, stratified by community group

Symptom	BLF (N=54) Median (IQR)	Rank	ITP support (N=17) Median (IQR)	Rank	Lupus UK (N=82) Median (IQR)	Rank	NRAS (N=229) Median (IQR)	Rank	PMRGCAUK (N=221) Median (IQR)	Rank
Acne	1 (1–5)	14	2 (1–7)	14	3 (1–6)	14	1 (1–6)	14	1 (1–6)	13
Cardiovascular disease	6.5 (1–10)	11	5 (1–10)	12	8 (1–10)	8	8 (1–10)	7	9 (1–10)	2
Changes in mood	7 (4–10)	9	6 (5–8)	8	8 (4–10)	6	7 (4–9)	10	7 (5–9)	9
Diabetes	7 (1–10)	10	4 (1–9)	13	7 (1–9)	12	8 (1–10)	7	8 (1–10)	8
Eye disease	7.5 (1–10)	6	8 (1–9)	5	8 (3–10)	7	8 (3–10)	4	9 (4–10)	1
High blood pressure	7.5 (4–10)	5	6 (4–8)	10	8 (5–9)	4	8 (3–10)	4	8 (3–10)	7
Indigestion	6 (4–8)	12	7 (5–8)	6	6 (3–8)	13	6 (4–8)	13	6 (3–8)	13
Infections	8.5 (3–10)	2	5 (1–9)	11	8 (5–10)	5	8 (2–10)	6	7 (1–10)	12
Insomnia	8 (7–10)	3	8 (7–10)	2	7 (5–9)	10	7 (5–9)	9	8 (6–10)	3
Palpitations	7 (4–9)	7	6 (4–7)	8	7.5 (5–9)	9	7 (4–9)	10	7 (4–9)	11
Reduced bone strength	8 (5–10)	4	8 (7–10)	2	9 (5–10)	2	8 (4–10)	3	8 (4–10)	6
Round face	6 (3–10)	13	8 (5–8)	2	8.5 (5–10)	3	8 (5–10)	2	8 (5–10)	5
Skin changes	7 (3–8)	7	7 (3–8)	7	7 (4–9)	11	7 (5–10)	10	7 (5–9)	9
Weight gain	9 (5–10)	1	9 (8–9)	1	9 (7–10)	1	9 (6–10)	1	8 (6–10)	3

BLF, British Lung Foundation; ITP, immune thrombocytopenia; NRAS, National Rheumatoid Arthritis Society; PMRGCAUK, Polymyalgia Rheumatica and Giant Cell Arteritis UK.

Table 4 Median, IQR and rank of side effect scores, stratified by experience of side effect

Symptom	Experienced side effect N	Median (IQR)	Rank	Did not experience side effect N	Median (IQR)	Rank
Acne	47	6 (5–8)	14	557	1 (1–5)	14
Cardiovascular disease	29	10 (6–10)	3	575	8 (1–10)	2
Changes in mood	356	8 (5–10)	10	248	5 (2–8)	11
Diabetes	66	10 (8–10)	1	538	7 (1–10)	4
Eye disease	117	10 (7–10)	2	487	8 (1–10)	2
High blood pressure	203	9 (7–10)	4	401	6 (1–9)	9
Indigestion	304	7 (5–9)	13	300	5 (1–7)	11
Infections	133	8 (6–10)	9	471	7 (1–10)	5
Insomnia	381	8 (7–10)	8	223	6 (4–8)	6
Palpitations	259	8 (5–10)	10	345	6 (3–8)	7
Reduced bone strength	162	9 (6–10)	7	442	8 (3–10)	1
Round face	383	9 (7–10)	4	221	5 (1–8)	13
Skin changes	348	8 (5–10)	10	256	5 (3–8)	10
Weight gain	442	9 (7–10)	4	162	6 (2–9)	8

adherence.[13 16] Three prior studies have investigated patient perspectives of GC side effects specifically. The most cited of these is a study comparing the perspectives of 140 patients and 110 rheumatologists. They found osteoporosis was 'the most worrisome' side effect for patients, followed by CVD, diabetes, weight gain and renal dysfunction.[8] The other two studies were interested in specific disease groups. One study of patients with ITP found the most bothersome side effects of those experienced, in line with our findings, were moon face, weight gain and insomnia.[10] Another study of patients with adrenal insufficiency found the most worrisome side effect was osteoporosis, followed by obesity and fatigue.[9] In all studies, weight gain was one of the top five most worrisome side effects, which is in agreement with our findings. Weight gain is known to adversely affect body image and self-esteem, although there are no studies, to the best of our knowledge, examining the impact of GC-associated weight gain on quality of life. A few studies have reported on weight gain following GC therapy.[14 17 18] However, studies are often not designed to measure this as an outcome and as a result, fail to address the sort of questions that patients may be interested in, such as the extent of weight gain with

specific doses, or the likelihood of weight loss following discontinuation. Despite the importance to patients of insomnia in this and other studies, it is interesting to note that very few studies have investigated insomnia in patients taking GC therapy.

This study used a novel method of recruiting survey respondents, through a social networking website, which was easy to conduct and resulted in a sample of just over 600 UK-based respondents who were taking GCs for a variety of conditions. However, there were limitations of the study. The sample was mainly female and over 50 years of age: this may be partly due to the disease demographic, but may also represent a selection bias in the types of people who are more likely to participate in studies,[19] or participate in a social network. This selection factor may have influenced our findings if perceptions of the importance of side effects could be different between the sexes. For example, female participants may be more inclined to see weight gain as important. It may also affect the generalisability of the results, as the scores may not represent the views of the whole population, for example, young men are not well represented. It relied on self-report to identify steroid users. However, a previous study showed high agreement between self-reported medication use and pharmacy records, so it is unlikely there will be large misclassification due to self-report.[20] We did not collect information about comorbidities in participants and were thus unable to examine how this may have influenced beliefs. For example, a patient with prevalent hypertension may have considered high BP or CVD to be particularly important to them as a GC-associated side effect. Nonetheless, our results reflect the patients' experiences and how they rate the importance of serious and non-serious outcomes. It was particularly interesting to note the distribution of responses in the high-ranking serious and non-serious conditions. For weight gain and insomnia, only 12% of participants scored them as low importance. Yet although CVD and diabetes had a median score of 8 (like insomnia), there was a U-shaped distribution of scores where more than 20% of participants scored them as low importance despite their seriousness. It may be that education influenced scores: if respondents were not aware of the risks of CVD with GCs, for example, they may not have scored CVD as important to them. Unfortunately, we did not collect information on education. Another explanation may be that some respondents may have had optimism bias,[21] where respondents believed that the serious side effects would not happen to them. This could also result in the wide variation of scores for serious side effects.

In conclusion, this study has shown that weight gain, insomnia and moon face were the top three most important side effects to patients taking GCs. Despite this, they are not widely studied with many unanswered questions. Research should be informed by patients, and targeted to provide patients with better information about these side effects of high importance.

Author affiliations
[1]Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, School of Biological Sciences, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
[2]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[3]Health eResearch Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK

Acknowledgements We would like to thank the following colleagues for their advice during the development of the survey: Louise Hunter, Meghna Jani, Rebecca Joseph, Catharine Morgan, David Ray and Suzanne Verstappen. We would also like to thank the Research User Group for their feedback on the draft survey.

Contributors WGD conceived the idea. WGD, JMcB, JH and RC were responsible for the design of the study. RC conducted the analysis and drafted the manuscript. All authors interpreted the results, critically revised the manuscript for important intellectual content and approved the final manuscript.

Funding This work was supported by the Arthritis Research UK Centre for Epidemiology (20380). WGD was supported by an MRC clinician scientist fellowship (G0902272).

Competing interests None declared.

Ethics approval The study received ethics approval from the University of Manchester Research Ethics Committee (Ref: 15496).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data are available and can be accessed by contacting Professor William Dixon (will.dixon@manchester.ac.uk).

REFERENCES

1. Hench PS, Kendall EC, Slocumb CH, *et al*. Effects of cortisone acetate and pituitary ACTH on rheumatoid arthritis, rheumatic fever and certain other conditions—a study in clinical physiology. *Arch Intern Med* 1950;85:545–666.
2. van Staa TP, Leufkens HG, Abenhaim L, *et al*. Use of oral corticosteroids in the United Kingdom. *QJM* 2000;93:105–11.
3. Bertsias G, Ioannidis JP, Boletis J, *et al*. EULAR recommendations for the management of systemic lupus erythematosus. Report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics. *Ann Rheum Dis* 2008;67:195–205.
4. Dejaco C, Singh YP, Perel P, *et al*. 2015 Recommendations for the management of polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. *Ann Rheum Dis* 2015;74:1799–807.
5. Davis JM III, Maradit Kremers H, Crowson CS, *et al*. Glucocorticoids and cardiovascular events in rheumatoid arthritis: a population-based cohort study. *Arthritis Rheum* 2007;56:820–30.
6. Dixon WG, Suissa S, Hudson M. The association between systemic glucocorticoid therapy and the risk of infection in patients with rheumatoid arthritis: systematic review and meta-analyses. *Arthritis Res Ther* 2011;13:R139.
7. Van Staa TP, Leufkens HG, Abenhaim L, *et al*. Oral corticosteroids and fracture risk: relationship to daily and cumulative doses. *Rheumatology* 2000;39:1383–9.
8. van der Goes MC, Jacobs JW, Boers M, *et al*. Patient and rheumatologist perspectives on glucocorticoids: an exercise to improve the implementation of the European League Against Rheumatism (EULAR) recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Ann Rheum Dis* 2010;69:1015–21.

9. Forss M, Batcheller G, Skrtic S, *et al*. Current practice of glucocorticoid replacement therapy and patient-perceived health outcomes in adrenal insufficiency—a worldwide patient survey. *BMC Endocr Disord* 2012;12:8.
10. Guidry JA, George JN, Vesely SK, *et al*. Corticosteroid side-effects and risk for bleeding in immune thrombocytopenic purpura: patient and hematologist perspectives. *Eur J Haematol* 2009;83:175–82.
11. Herxheimer A. Communicating with patients about harms and risks. *PLoS Med* 2005;2:e42.
12. Cooper V, Metcalf L, Versnel J, *et al*. Patient-reported side effects, concerns and adherence to corticosteroid treatment for asthma, and comparison with physician estimates of side-effect prevalence: a UK-wide, cross-sectional study. *NPJ Prim Care Respir Med* 2015;25:15026.
13. Morrison E, Crosbie D, Capell HA. Attitude of rheumatoid arthritis patients to treatment with oral corticosteroids. *Rheumatology* 2003;42:1247–50.
14. Curtis JR, Westfall AO, Allison J, *et al*. Population-based assessment of adverse events associated with long-term glucocorticoid use. *Arthritis Rheum* 2006;55:420–6.
15. Hoes JN, Jacobs JW, Buttgereit F, *et al*. Current view of glucocorticoid co-therapy with DMARDs in rheumatoid arthritis. *Nat Rev Rheumatol* 2010;6:693–702.
16. Dixon WG, Bansback N. Understanding the side effects of glucocorticoid therapy: shining a light on a drug everyone thinks they know. *Ann Rheum Dis* 2012;71:1761–4.
17. Berthon BS, MacDonald-Wicks LK, Wood LG. A systematic review of the effect of oral glucocorticoids on energy intake, appetite, and body weight in humans. *Nutr Res* 2014;34:179–90.
18. Wung PK, Anderson T, Fontaine KR, *et al*. Effects of glucocorticoids on weight change during the treatment of Wegener's granulomatosis. *Arthritis Rheum* 2008;59:746–53.
19. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007;17:643–53.
20. Curtis JR, Westfall AO, Allison J, *et al*. Agreement and validity of pharmacy data versus self-report for use of osteoporosis medications among chronic glucocorticoid users. *Pharmacoepidemiol Drug Saf* 2006;15:710–18.
21. Weinstein ND. Unrealistic optimism about future life events. *J Pers Soc Psychol* 1980;39:806–20.

Timing and outcomes of steroid use

We'd like to invite you to take this short survey, which aims to investigate the time at which patients take their steroids tablets (E.g. prednisolone, betamethasone, deflazacort, calcort, dexamethasone, hydrocortisone, methylprednisolone and medrone) and what side effects patients view as being important.

Before you decide whether you want to take part, it is important for you to understand why the survey is being done and what your participation will involve. Please take time to read the following information carefully:

● The survey is being funded by the Medical Research Council and conducted by the University of Manchester, in partnership with HealthUnlocked.
● All information you provide will be anonymous and treated in the strictest confidence and according to legal and ethical guidelines of the UK Data Protection Act 1998.
● The information you provide will be stored at the University of Manchester for 10 years.
● If you decide to take part you are still free to withdraw at any time during the survey. However, as data collection is anonymous and sent when the survey is completed, your information cannot be identified to withdraw after survey completion.

The total completion of the survey should not take more than 2 minutes

☐ **Tick this box if you understand the above information and agree to take part**

Do you take oral steroid tablets (or have taken them within the last month)?

Examples of steroid tablets include prednisolone, betamethasone, deflazacort, calcort, dexamethasone, hydrocortisone, methylprednisolone and medrone.

- ◯ Yes
- ◯ No

1. How old are you?

- ◯ under 19
- ◯ 20 - 29
- ◯ 30 - 39
- ◯ 40 - 49
- ◯ 50 - 59
- ◯ 60 - 69
- ◯ 70 or more

2. Which gender best describes you?

- ◯ Male
- ◯ Female

3. Do you take your daily dose of steroids once per day, or do you split the dose over two or more times through the day?

- ◯ Once per day
- ◯ Two or more

4. What time do you normally take your steroid tablets?

Please enter the nearest time when taken most often

	hh	mm	AM/PM
Time		:	-

4. What time do you normally take your steroid tablets?

Please enter the nearest time when taken most often.

First dose:

	hh	mm	AM/PM
Time		:	-

Second dose:

	hh	mm	AM/PM
Time		:	-

Third dose:

	hh	mm	AM/PM
Time		:	-

5. We are interested to learn how important a range of possible side effects is to you.

Please score each side effect, <u>even if you have not experienced it</u>, on a scale where 1=very little importance, and 10=high importance to you.

	1	2	3	4	5	6	7	8	9	10
Acne	○	○	○	○	○	○	○	○	○	○
Cardiovascular disease (e.g. heart attack)	○	○	○	○	○	○	○	○	○	○
Changes in mood	○	○	○	○	○	○	○	○	○	○
Diabetes	○	○	○	○	○	○	○	○	○	○
Eye disease (cataracts, glaucoma)	○	○	○	○	○	○	○	○	○	○
High blood pressure	○	○	○	○	○	○	○	○	○	○
Indigestion	○	○	○	○	○	○	○	○	○	○
Infection (e.g. pneumonia)	○	○	○	○	○	○	○	○	○	○
Insomnia (unable to get to sleep)	○	○	○	○	○	○	○	○	○	○
Palpitations (racing heart)	○	○	○	○	○	○	○	○	○	○
Reduced bone strength (osteoporosis, fractures)	○	○	○	○	○	○	○	○	○	○
Round face, or ‘moon’ face	○	○	○	○	○	○	○	○	○	○
Skin changes (bruising, thin skin, stretch marks)	○	○	○	○	○	○	○	○	○	○
Weight gain	○	○	○	○	○	○	○	○	○	○

Other (please specify)

6. Have you had any of these side effects whilst taking steroids?

Please select all that apply

- [ ] Acne
- [ ] Cardiovascular disease (e.g. heart attack)
- [ ] Changes in mood
- [ ] Diabetes
- [ ] Eye disease (cataracts, glaucoma)
- [ ] High blood pressure
- [ ] Indigestion
- [ ] Infection (e.g. pneumonia)
- [ ] Insomnia (unable to get to sleep)
- [ ] Palpitations (racing heart)
- [ ] Reduced bone strength (osteoporosis, fractures)
- [ ] Round face, or ‘moon’ face
- [ ] Skin changes (bruising, thin skin, stretch marks)
- [ ] Weight gain
- [ ] None of the above / Other (please specify)

Publication 7: Representativeness of a digitally engaged population and a patient organisation population with rheumatoid arthritis and their willingness to participate in research: a cross-sectional study

RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL ARTICLE

Representativeness of a digitally engaged population and a patient organisation population with rheumatoid arthritis and their willingness to participate in research: a cross-sectional study

Ruth Costello,[1] Clare Jacklin,[2] Matthew Jameson Evans,[3] John McBeth,[1] William G Dixon[1,4,5]

To cite: Costello R, Jacklin C, Jameson Evans M, *et al*. Representativeness of a digitally engaged population and a patient organisation population with rheumatoid arthritis and their willingness to participate in research: a cross-sectional study. *RMD Open* 2018;**4**:e000664. doi:10.1136/rmdopen-2018-000664

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/10.1136/rmdopen-2018-000664).

Received 12 February 2018
Accepted 25 May 2018

For numbered affiliations see end of article.

Correspondence to
William G Dixon;
will.dixon@manchester.ac.uk

ABSTRACT

Objectives To describe (1) the representativeness of (a) users of an online health community (HealthUnlocked.com (HU)) with rheumatoid arthritis (RA) and (b) paid members of an RA patient organisation, the National Rheumatoid Arthritis Society (NRAS), compared with the general RA population; and (2) the willingness of HU users with RA to participate in types of research (surveys, use of an app or activity tracker, and trials).

Methods A pop-up survey was embedded on HU to determine the characteristics of users and their willingness to participate in research. An anonymous data set of NRAS member characteristics was provided by the NRAS (N=2044). To represent the general RA population, characteristics of people with RA were identified from the Clinical Practice Research Datalink (CPRD) (N=20 594). Cross-sectional comparisons were made across the three groups.

Results Compared with CPRD, HU respondents (n=615) were significantly younger (49% aged below 55 years compared with 23% of CPRD patients), significantly more deprived (21% in the most deprived Townsend quintile compared with 12% of CPRD patients) and had more recent disease, with 62% diagnosed between 2010 and 2016 compared with 37% of CPRD patients. NRAS members were more similar to the CPRD, but significantly under-represented those aged 75 years or over and over-represented those aged 55–75 years compared with the CPRD. High proportions of HU users were willing to participate in future research of all types.

Conclusions NRAS members were broadly representative of the general RA population. HU users were younger, more deprived and more recently diagnosed. HU users were willing to participate in most types of research.

Key messages

What is already known about this subject?

- Studies are starting to recruit participants online and through patient organisations, but we do not know how representative these groups are.

What does this study add?

- Patient organisation members with rheumatoid arthritis (RA) were broadly representative of the general RA population, and online health community (OHC) users with RA were younger, more recently diagnosed and from more deprived areas.
- A high proportion of OHC users were willing to take part in all types of research (surveys, use of an app or activity tracker, and trials).

How might this impact on clinical practice?

- Future studies may be able to recruit more efficiently from OHCs and patient organisations with confidence in how these populations represent the study population.

INTRODUCTION

Large population studies often require significant numbers of participants to generate enough statistical power. This often requires multisite recruitment through rheumatology departments. A study of trials conducted in 2002–2008 found only 55% recruited to their prespecified sample size.[1] This leads to an underpowered study and possible inconclusive results.

Study recruitment may be improved in both numbers and efficiency by recruiting patients directly. This may be coordinated via patient organisations or, as patients are increasingly online,[2] through the internet. For example,

studies have recruited through social media,[3 4] recruited through online forums,[5–7] advertised on health websites[4] or advertised based on health-related search terms on Google.[8] However the representativeness of online health communities (OHCs) and patient organisations, particularly in a rheumatoid arthritis (RA) population, is not clear.

The aims of this study were to describe (1) the representativeness of paid members of a patient organisation with prevalent RA and users of an OHC with RA when compared with the general RA population, and (2) the types of studies that OHC users with RA would participate in.

METHODS

Design

This cross-sectional study compared the characteristics of adults with RA from the National Rheumatoid Arthritis Society (NRAS) members who had paid for membership and visitors to the NRAS community group on HealthUnlocked.com (HU) with adults with RA identified from the Clinical Practice Research Datalink (CPRD), a database of anonymised UK primary care electronic medical records. As the CPRD is broadly representative of the UK population,[9] adults with RA identified from the CPRD were considered representative of adults with RA in the UK.

Patient organisation population

The NRAS is a patient organisation for people living with RA. When people join NRAS or renew their membership, they can provide demographic and medical information. An anonymised data set of all members, past and present up until 1 May 2016, was provided by the NRAS. For consistency with the other data sets, and to avoid selection bias, only current NRAS members were used. The data set contained (self-reported) year of RA diagnosis, ethnicity, current age, gender, employment status, and ever use of disease-modifying antirheumatic drugs (DMARDs), biologics and glucocorticoids (GC). To be included in the analyses, respondents had to be residents in the UK to allow comparison with the other UK data sets.

HU population

HU is Europe's largest OHC, with over 4.5 million visitors per month.[10] The NRAS has a community group on HU for people with RA with, on average, 169 000 visitors per month. Anybody can visit the NRAS community on HU irrespective of a diagnosis of RA, NRAS membership or following the NRAS HU community. As people join HU without providing demographic information, a survey was developed to determine self-reported RA diagnosis, year of RA diagnosis, medications used, willingness to participate in different types of research (including questionnaires of varying durations, using an app, wearing an activity tracker and different types of trial), demographics (age, gender, employment, postcode and ethnicity) and the types of electronic devices owned (details of survey development in online supplementary file 1). After review by a combined patient and public involvement group and agreement with the NRAS, the finalised survey (online supplementary figure 1) was embedded in all posts within the NRAS HU community and popped up for completion when these posts were viewed by someone with a UK IP address. Prior to starting the survey, respondents confirmed they were over 18 years of age. The survey then started with an eligibility question to determine self-reported RA. The survey started on 6 May 2016 and was live for 3 months or until 1000 people had completed the survey, whichever was soonest. Postcode was converted to Townsend Deprivation Index[11] by a health data scientist outside of the research team prior to analysis.

CPRD population

A prevalent cohort of patients with a diagnosis of RA prior to 1 June 2016 was identified using a validated algorithm.[12] Eligibility criteria were (1) aged 18 years or over at RA diagnosis, (2) registered at a practice on 1 May 2016 and (3) data met the CPRD quality standards.[9] Age, gender, year of RA diagnosis, ethnicity, ever DMARD and GC use, and Townsend Deprivation Index (for practices that consented to linkage) were identified for these patients (covariate definitions in online supplementary file 1).

Analysis

For each data set, the characteristics were categorised and tabulated to match the HU survey responses to allow comparison between data sets. A Z-test for the difference in proportions within each category of each characteristic was calculated comparing NRAS with CPRD, and HU with CPRD, where CPRD data were available. The characteristics of those who would definitely or probably take part in each type of research are reported. Logistic regression was used to identify any characteristics that were independently associated with definite or probable participation in each type of research.

Missing data

To be included in this analysis, individuals had to have information on at least age and gender. For CPRD employment status was available for less than 5% of patients so it was not used in this analysis. For NRAS members, postcode and therefore Townsend Deprivation Index were unavailable. For all variables, except age and gender, when the variable was available for the data set, the percentage of missing data is reported.

RESULTS

Data sets

NRAS

The NRAS provided a data set of 4505 current and past members. Of those, 1498 were not currently members, 22 were from overseas and 941 did not have information

on age and gender, resulting in a data set of 2044 current members with RA.

HealthUnlocked survey

The HU survey was live for 74 days between 6 May 2016 and 12 August 2016 and had 100 112 pop-ups to unique IP addresses. There were 2647 pop-ups clicked, 900 respondents agreed to take part, 750 respondents were eligible, and 135 did not provide age and gender, resulting in 615 respondents available for analysis. Recruitment was steady with an average of 12 responses per day.

CPRD

Of 4 776 441 people in the CPRD, there were 20 594 (0.43%) patients with a diagnosis of RA on 1 June 2016.

Characteristics

Table 1 and figure 1 show that NRAS members had a reasonably similar age distribution to patients with RA from the CPRD up to age 55. After this age there were statistically significant differences in proportions, with an over-representation of people aged 55–75 years and an under-representation of people aged 75 years and over in NRAS members. HU users were a significantly younger population compared with the CPRD, with fewer responders aged 65 years or over. Both NRAS and HU were predominantly female, with significantly higher proportions (~85%) compared with CPRD (70%). HU users had shorter disease duration, with significantly more respondents diagnosed between 2010 and 2016 (62%) compared with CPRD participants (37%), while NRAS members has a longer disease duration, with significantly fewer people diagnosed between 2010 and 2016 (25%). HU responders had a significantly higher proportion of people from more deprived areas (most deprived Townsend quintile: HU: 22% vs CPRD 12%) and significantly less from affluent areas (least deprived Townsend quintile: HU: 18% vs CPRD 23%) (data not available for NRAS members). All DMARDs had significantly more ever use in both HU and NRAS compared with CPRD.

Participation in future research (HealthUnlocked only)

HU responders commonly reported they were definitely or probably willing to take part in future research, particularly questionnaires, with 89% reporting willingness to complete a questionnaire of 10 min. A lower proportion reported willingness to use an app (63%) compared with wearing an activity tracker (74%). Half of the respondents reported willingness to take part in a drug trial via the internet or with site visits. When stratified by age, overall those over 65 years of age reported less willingness to take part in all research types. When stratified by gender, men reported more willingness to use an app while women reported more willingness to wear an activity tracker (table 2). The most striking result from multivariate logistic regression showed that participants over 45 years of age were significantly less willing to use an app compared with those aged 18–34 years. Those aged 45–54 years had a 92% lower odds of using an app (OR: 0.08 (95% CI 0.01 to 0.58)), and those aged 75 years and over had 98% lower odds of using an app compared with those aged 18–34 years (OR: 0.02 (95% CI 0.002 to 0.23)). There were no statistically significant differences by gender (online supplementary table 1).

DISCUSSION

This study shows that people with RA who were NRAS members were reasonably representative of the general RA population, although fewer were aged 75 years or over. HU visitors were a younger RA population, with more recent disease and more deprivation than the general RA population. Most respondents from the OHC were willing to take part in studies with lower burden. More than half were willing to take part in any type of study including a drug trial via the internet. Younger participants were more willing to use an app. We have also demonstrated that over 600 responses to a short questionnaire can be collected over a short period of time using pop-ups within an OHC.

Recruiting online through HU was a straightforward and less labour-intensive method of recruiting a reasonably large sample of respondents in a short space of time. Once the survey had been designed and then implemented by HU, other than monitoring the numbers of surveys completed, it did not require further work by the study team as data were automatically captured. This contrasts to more traditional methods where a person is required to collect data for each survey throughout data collection. Ninety per cent of those aged 55–64 reported they recently used the internet in a UK national survey.[13] This is seen in this study with good representation of people aged 45–65 years in our sample. Those aged over 75 years were not well represented and may be expected given that one in four of those aged 75 or over are online,[13] and this may impact the generalisability of study results and would need to be considered by investigators designing studies. For example, if the disease of interest affects elderly people, studies using HU may wish to consider additional recruitment sources to ensure that elderly patients are represented. Conversely, if investigators are interested in deprivation or people recently diagnosed with RA, then HU may be a good source of participants.

Few studies have looked specifically at the representativeness of members of a patient organisation or internet users with RA. A study including a group of patients with RA found that internet users were younger, more educated and more commonly employed compared with those who did not use the internet for health.[14] In this study HU responders were younger, with a similar proportion employed compared with NRAS members, although we did not have CPRD as comparison. Although there was no CPRD comparison, the proportion who had ever

Table 1 Characteristics of patients with RA who are NRAS members, or who responded to a survey on HU and those identified from CPRD

	CPRD (N=20 594)	NRAS (N=2044)	Difference in proportion compared with CPRD (95% CI)	P values	HU (N=615)	Difference in proportion compared with CPRD (95% CI)	P values
	n (%)	n (%)			n (%)		
Age (years)							
18–34	499 (2.4)	37 (1.8)	0.6 (–0.002 to 1.2)	0.08	26 (4.2)	–1.8 (–3.4 to –0.02)	0.005
35–44	1249 (6.1)	129 (6.3)	–0.2 (–1.3 to 0.9)	0.66	79 (12.9)	–6.8 (–9.4 to –4.1)	<0.001
45–54	2936 (14.3)	311 (15.2)	–1 (–3 to 0.7)	0.24	195 (31.7)	–17.5 (–21.2 to –13.7)	<0.001
55–64	4484 (21.8)	568 (27.8)	–6 (–8 to –4)	<0.001	218 (35.5)	–13.7 (–17.5 to –9.9)	<0.001
65–74	5643 (27.4)	709 (34.7)	–7 (–9 to –5)	<0.001	81 (13.2)	14.2 (11.5 to 17.0)	<0.001
75 and over	5783 (28)	290 (14.2)	13.9 (12.3 to 15.5)	<0.001	16 (2.6)	25.4 (24.1 to 26.9)	<0.001
Gender							
Female	14 440 (70.1)	1728 (84.5)	–0.18 (–0.21 to –0.16)	<0.001	544 (88.5)	–18.3 (–20.9 to –15.7)	<0.001
Year of RA diagnosis							
<1990	34 (0.2)	274 (14.5)	–14.3 (–15.9 to –12.7)	<0.001	39 (6.3)	–6.2 (–8.1 to –4.2)	<0.001
1990–1994	869 (4.2)	122 (6.5)	–2.2 (–3.4 to –1.1)	<0.001	22 (3.6)	0.6 (–0.9 to 2.1)	0.43
1995–1999	1756 (8.5)	178 (9.4)	–0.9 (–2.3 to 0.5)	0.19	32 (5.2)	3.3 (1.5 to 5.1)	0.004
2000–2004	4336 (21.1)	304 (16.1)	5.0 (3.2 to 6.7)	<0.001	46 (7.5)	13.5 (11.4 to 15.7)	<0.001
2005–2009	5962 (29)	540 (28.6)	0.4 (–1.7 to 2.5)	0.72	95 (15.5)	13.5 (10.6 to 16.4)	<0.001
2010–2016	7637 (37.1)	473 (25)	12.1 (10 to 14.1)	<0.001	381 (62)	–24.9 (–28.8 to –21.0)	<0.001
Missing	0	153			0		
Employment status							
Full-time employed		543 (24.1)			387 (23.6)		
Part-time employed		418 (18.5)			288 (17.5)		
Unemployed		304 (13.5)			280 (17)		
Retired		698 (30.9)			577 (35.1)		
Retired due to arthritis		151 (6.7)			83 (5.1)		
Not working due to ill health		142 (6.3)			28 (1.7)		
Missing	NA	401			0		
Ethnicity							
White	8931 (93.9)	1572 (98.4)	–4.5 (–5.3 to –3.7)	<0.001	578 (94.6)	–0.73 (–2.6 to 1.1)	0.47
Mixed	44 (0.5)	5 (0.3)	0.1 (–0.2 to 0.5)	0.40	8 (1.3)	–0.8 (–1.8 to 0.06)	0.005
Asian	315 (3.3)	12 (0.8)	2.6 (2.0 to 3.1)	<0.001	11 (1.8)	1.5 (0.3 to 2.6)	0.04
Black	147 (1.6)	5 (0.3)	1.2 (0.9 to 1.6)	<0.001	13 (2.1)	–0.58 (–1.8 to 0.6)	0.26
Other	77 (0.8)	4 (0.3)	0.6 (0.3 to 0.9)	0.02	1 (0.2)	0.6 (0.3 to 1.0)	0.08
Missing	11 080	446			2		
Townsend Deprivation Index							
1 (least deprived)	2807 (23.1)				103 (18.3)	4.9 (1.6 to 8.1)	0.007
2	2949 (24.3)				112 (19.9)	4.4 (1 to 7.8)	0.016
3	2543 (20.9)				107 (19)	2.0 (–1.3 to 5.3)	0.26
4	2329 (19.2)				119 (21.1)	–1.9 (–5.4 to 1.5)	0.26
5 (most deprived)	1514 (12.5)				123 (21.8)	–9.3 (–12.8 to -5.9)	<0.001
Missing	8452	NA			49		
Ever taken methotrexate							
Yes	14 553 (70.7)	1783 (87.2)	–16.6 (–18.1 to –15.0)	<0.001	511 (84.9)	–14.2 (–17.1 to –11.3)	<0.001
Missing or unknown	0	0			13		

Continued

Table 1 Continued

	CPRD (N=20 594)	NRAS (N=2044)	Difference in proportion compared with CPRD (95% CI)	P values	HU (N=615)	Difference in proportion compared with CPRD (95% CI)	P values
	n (%)	n (%)			n (%)		
Ever taken sulfasalazine							
Yes	9742 (47.3)	1064 (52.1)	−4.7 (−7 to −2.5)	<0.001	339 (56.9)	−9.6 (−13.6 to −5.5)	<0.001
Missing or unknown	0	0			19		
Ever taken hydroxychloroquine							
Yes	8255 (40.1)	883 (43.2)	−3.1 (−5.4 to −0.9)	0.006	359 (60.1)	−20 (−24 to −16)	<0.001
Missing or unknown	0	0			18		
Ever taken leflunomide							
Yes	2576 (12.5)	378 (18.5)	−6 (−7.7 to −4.2)	<0.001	113 (19)	−6.5 (−9.7 to −3.3)	<0.001
Missing or unknown	0	0			21		
Ever taken DMARDs*							
Yes	18 683 (90.7)	1956 (95.7)	−5 (−6 to −4)	<0.001	537 (93.4)	−2.7 (−4.7 to −0.6)	0.029
Missing or unknown	0	0			106		
Ever taken glucocorticoids							
Yes	11 889 (57.7)	351 (17.2)	40.6 (38.8 to 42.3)	<0.001	368 (61.7)	−4 (−8 to −0.05)	0.05
Missing or unknown	0	0			19		
Ever taken biologics							
Yes		897 (43.9)			196 (32.9)		
Missing or unknown	20 594	0			20		

*Ever taken DMARDs is based on the ever taken methotrexate, ever taken sulfasalazine, ever taken hydroxychloroquine and ever taken leflunomide data.
CPRD, Clinical PracticeResearch Datalink; DMARD, disease-modifyingantirheumatic drug; HU, HealthUnlocked.com; NA, not available; NRAS, National RheumatoidArthritis Society; RA, rheumatoid arthritis.

taken biologics was high in both HU and NRAS compared with the reported UK estimates of 11%–16%.[15 16] This may indicate that both NRAS and HU respondents have more severe disease requiring biologics, and this may be why they are using HU or NRAS. However as we do not have disease activity measures, we cannot be sure of this.

There were some limitations to this study. RA diagnosis in the CPRD relies on Read (diagnosis) codes and drug codes so there may be some misclassification. The prevalence rate of RA (0.43%) is lower than the 1% prevalence otherwise estimated,[17] which supports some misclassification. RA diagnosis relied on self-report for both the NRAS and HU, so there may have been some misclassification; however, as these groups are both specific for RA, it is likely that any misclassification would be small. NRAS characteristics relied on members providing personal details: there may be some selection bias if those who gave information were different from those who did not. Further to this, as NRAS membership requires payment, there may be some selection bias in that NRAS members may be less deprived than the general population; however, we were unable to capture the Townsend Deprivation Index for this group. There may be some HU respondents who were NRAS members also; 114 HU respondents indicated they were NRAS members. However, it was not possible to cross-reference the NRAS and HU data sets. The HU characteristics reflect the characteristics of those who completed the survey, so may not be representative of all HU users, but does provide insight into the characteristics of those willing to join studies via this route. Although we did not survey NRAS paid members about their willingness to participate in research, recent experience demonstrates that patients with RA, both members and non-members, are responsive to participating in research following outreach from the NRAS.

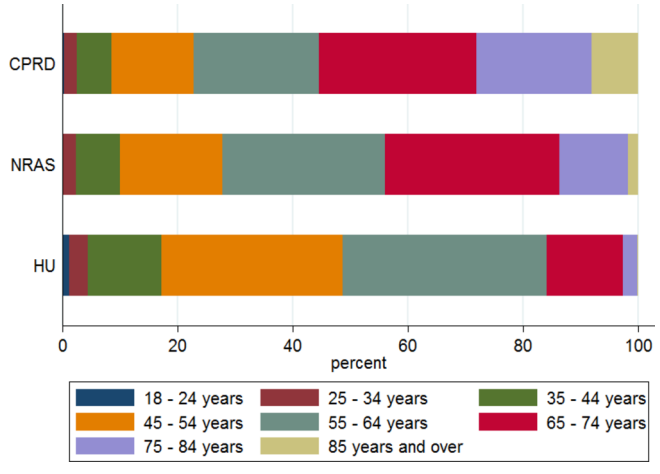


Figure 1 Proportions in each age group by data set. CPRD, Clinical PracticeResearch Datalink; HU, HealthUnlocked.com; NRAS, National RheumatoidArthritis Society.

Table 2 Proportion of HealthUnlocked users who would definitely or probably take part in types of research

	Type of research						
	Complete questionnaires				Participate in trials		
Characteristics	Single 10 min	Multiple questionnaires over months?	Using an app	Wearing an activity tracker	Non-drug treatment	Drug treatment via internet alone	Drug treatment with site visits
Total	546 (89.1)	503 (82.1)	387 (63.1)	453 (73.9)	401 (65.4)	305 (49.8)	327 (53.3)
Age category (years)							
18–34	23 (88.5)	19 (73.1)	25 (96.2)	20 (76.9)	18 (69.2)	12 (46.2)	15 (57.7)
35–44	73 (92.4)	69 (87.3)	60 (76)	65 (82.3)	55 (69.6)	48 (60.8)	51 (64.6)
45–54	172 (88.2)	155 (79.5)	123 (63.1)	145 (74.4)	134 (68.7)	108 (55.4)	108 (55.4)
55–64	201 (92.2)	184 (84.4)	133 (61)	157 (72)	138 (63.3)	100 (45.9)	111 (50.9)
65–74	63 (77.8)	65 (80.3)	41 (50.6)	58 (71.6)	48 (59.3)	32 (39.5)	38 (46.9)
75 and over	16 (100)	13 (81.3)	6 (37.5)	9 (56.3)	9 (56.3)	5 (31.3)	5 (31.3)
Gender							
Female	483 (89.1)	446 (82.3)	337 (62.2)	408 (75.3)	353 (65.1)	267 (49.3)	288 (53.1)
Employment status							
Full-time employment	140 (89.7)	122 (78.2)	102 (65.4)	119 (76.3)	105 (67.3)	78 (50)	72 (46.2)
Part-time employment	115 (88.5)	102 (78.5)	75 (57.7)	90 (69.2)	78 (60)	64 (49.2)	71 (54.6)
Unemployed	21 (87.5)	20 (83.3)	16 (66.7)	18 (75)	16 (66.7)	12 (50)	16 (66.7)
Retired	101 (83.5)	98 (81)	67 (55.4)	84 (69.4)	69 (57)	47 (38.8)	53 (43.8)
Retired due to arthritis	59 (86.8)	59 (86.8)	44 (64.7)	50 (73.5)	43 (63.2)	34 (50)	40 (58.8)
Not working due to ill health	110 (96.5)	102 (89.5)	83 (72.8)	92 (80.7)	90 (79)	70 (61.4)	75 (65.8)
Missing	2	2	2	2	2	2	2
Ethnicity							
White	518 (89.6)	481 (83.2)	362 (62.6)	426 (73.7)	373 (64.5)	291 (50.4)	311 (53.8)
Mixed	8 (100)	7 (87.5)	6 (75)	7 (87.5)	6 (75)	5 (62.5)	5 (62.5)
Asian	7 (63.6)	3 (27.3)	6 (54.6)	7 (63.6)	7 (63.6)	3 (27.3)	3 (27.3)
Black	10 (76.9)	9 (69.2)	10 (76.9)	11 (84.6)	13 (100)	5 (38.5)	6 (46.2)
Other	1 (100)	1 (100)	1 (100)	1 (100)	0 (0)	1 (100)	1 (100)
Missing	4	4	4	4	4	4	4
Townsend Deprivation Index							
1	96 (93.2)	86 (83.5)	60 (58.3)	73 (70.9)	58 (56.3)	52 (50.5)	50 (48.5)
2	103 (92)	96 (85.7)	72 (64.3)	92 (82.1)	84 (75)	58 (51.8)	58 (51.8)
3	100 (93.5)	86 (80.4)	71 (66.4)	84 (78.5)	78 (72.9)	64 (59.8)	66 (61.7)
4	107 (89.9)	102 (85.7)	73 (61.3)	77 (64.7)	68 (57.1)	54 (45.4)	63 (52.9)
5	104 (84.6)	101 (82.1)	87 (70.7)	97 (78.9)	91 (74)	59 (48)	71 (57.7)
Missing	51	51	51	51	51	51	51

This study gives an indication of the representativeness of groups that investigators may consider using to recruit people with RA to studies, while also demonstrating the feasibility of recruitment from OHCs. People in OHCs are willing to take part in many types of research, with the proportion declining as the burden of the research increases.

Author affiliations
[1]Arthritis Research UK Centre for Epidemiology, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, The University of Manchester, Manchester, UK
[2]National Rheumatoid Arthritis Society, Berkshire, UK
[3]HealthUnlocked (Everything Unlocked), London, UK
[4]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[5]Health eResearch Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK

Acknowledgements We would like to thank the following colleagues for their advice in developing the survey: Rebecca Joseph, Meghna Jani and Kimme Hyrich. We would also like to thank the Research User Group for their input during survey development. We would like to thank Kamilla Kopec-Harding for her help converting postcodes to the Townsend Deprivation Index.

Contributors WGD conceived the idea. WGD, JM and RC were responsible for the design of the study, MJE and CJ contributed towards the design of the study. RC conducted the analysis and drafted the manuscript. All authors interpreted the results, critically revised the manuscript for important intellectual content and approved the final manuscript.

Funding The work was supported by Arthritis Research UK Centre for Epidemiology (grant number: 20380).

Competing interests RC, JM and WGD have no competing interests. MJE is the Chief Medical Officer and cofounder of HealthUnlocked.com. CJ is employed by the NRAS.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Sully BG, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. *Trials* 2013;14:166.
2. Susannah Fox MD. Health Online 2013. Pew Research Center's Internet & American Life Project. 2013.
3. Heywood J, Evangelou M, Goymer D, *et al*. Effective recruitment of participants to a phase I study using the internet and publicity releases through charities and patient organisations: analysis of the adaptive study of IL-2 dose on regulatory T cells in type 1 diabetes (DILT1D). *Trials* 2015;16:86.
4. Hatch EE, Hahn KA, Wise LA, *et al*. Evaluation of selection bias in an internet-based study of pregnancy planners. *Epidemiology* 2016;27:98–104.
5. Hercberg S, Castetbon K, Czernichow S, *et al*. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 2010;10:242.
6. Mikkelsen EM, Hatch EE, Wise LA, *et al*. Cohort profile: the Danish web-based pregnancy planning study--'Snart-Gravid'. *Int J Epidemiol* 2009;38:938–43.
7. Mishra GD, Hockey R, Powers J, *et al*. Recruitment via the Internet and social networking sites: the 1989-1995 cohort of the Australian longitudinal study on women's health. *J Med Internet Res* 2014;16:e279.
8. McAlindon T, Formica M, Kabbara K, *et al*. Conducting clinical trials over the internet: feasibility study. *BMJ* 2003;327:484–7.
9. Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
10. HealthUnlocked annual user survey 2017. 2017.
11. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality in the north. 1988.
12. Thomas SL, Edwards CJ, Smeeth L, *et al*. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis Rheum* 2008;59:1314–21.
13. Office for National Statistics. Internet users in the UK: 2017. 2017.
14. van Uden-Kraan CF, Drossaert CH, Taal E, *et al*. Health-related Internet use by patients with somatic diseases: frequency of use and characteristics of users. *Inform Health Soc Care* 2009;34:18–29.
15. Gwinnutt JM, Symmons DPM, MacGregor AJ, *et al*. Have the 10-year outcomes of patients with early inflammatory arthritis improved in the new millennium compared with the decade before? Results from the Norfolk arthritis register. *Ann Rheum Dis* 2018;77.
16. Sokka T, Kautiainen H, Toloza S, *et al*. QUEST-RA: quantitative clinical assessment of patients with rheumatoid arthritis seen in standard rheumatology care in 15 countries. *Ann Rheum Dis* 2007;66:1491–6.
17. Symmons D, Turner G, Webb R, *et al*. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology* 2002;41:793–800.