

Incidences and risk factors of serious infections among different rheumatic patients: a retrospective cohort study

Jin Ding (✉ dingjin1@foxmail.com)

Xijing Hospital <https://orcid.org/0000-0003-3585-6063>

Yang Xiang

University of Texas Health Science Center at Houston

Yujia Zhou

University of Texas Health Science Center at Houston

Cong Zhu

University of Texas Health Science Center at Houston

Xinyue Hu

University of Texas Health Science Center at Houston

Cui Tao

University of Texas Health Science Center at Houston

Hua Xu


University of Texas Health Science Center at Houston

Research article

Keywords: rheumatic disease, serious infection, risk factor

Posted Date: August 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-53263/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

To compare the incidence and risk factors of serious infections among patients of seven common rheumatic diseases including rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), polymyalgia rheumatica (PMR), Sjögren's syndrome (SS), systemic sclerosis (SSc), systemic vasculitis (VA), and other diffuse connective tissue diseases (oCTD).

Methods

In a retrospective cohort study using large Electronic Health Records (EHR) data, the infection rates of different rheumatic diseases in two years were calculated and variances on risk factors were analyzed using the multivariable Cox model.

Results

Among the overall 46,411 rheumatic patients, 8,308 presented serious hospitalized infections in the following two years (crude infection rates (IR) per 100 patient-years: 8.95, 8.76–9.14). SSc (10.89, 9.61–12.16), VA (10.82, 9.96–11.68), and SLE (10.13, 9.73–10.53) had relatively high IRs, while oCTD was with the lowest IR (7.71, 6.58–8.84). The Cox model identified previous infection (adjusted hazard ratio (HR):1.75, 1.62–1.88, $p < 0.001$), GCs usage (HR:1.64, 1.52–1.76, $p < 0.001$), nonbiologic DMARDs usage (HR:1.39, 1.32–1.45, $p < 0.001$) and congestive heart failure (CHF) (HR:1.42, 1.30–1.56, $p < 0.001$) as the top-ranked risk factors of infection in the overall population. While some of the variables were shared across subtypes of rheumatic diseases as highly associated risk factors, others varied significantly such as osteoporosis.

Conclusions

Infection rates and risk factors varied among cohorts of different rheumatic diseases. The analytical results may inform distinct strategies and influence clinical decision making to reduce the occurrence of infections.

Background

Systemic rheumatic diseases are associated with considerable morbidity and mortality, a significant proportion of which can be attributed to serious infections [1–3]. Possible causes of vulnerability of rheumatic patients towards infection include alterations of immunoregulation, disease severity, comorbid illnesses, and the use of immunosuppressive medications [4]. Therefore, to reduce serious infections, it is important to have a clear understanding of the disease progressions and potential risk factors.

The reported incidence rates of infection vary dramatically across different subtypes of rheumatic diseases, and even among different cohorts for a certain disease. For example, in two rheumatoid arthritis (RA) cohorts from Minnesota, the U.S., the incidences of serious infection were 7.2 and 9.1 per 100 patient-years respectively [5]. In contrast, the infection incidence rate per 100 person-years was higher (10.8) based upon a diverse, nationwide cohort of systemic lupus erythematosus (SLE) patients [6]. Moreover, investigations and understandings of population-level incidence rates of serious infections for some rheumatic diseases such as polymyalgia rheumatica (PMR), Sjögren's syndrome (SS), and vasculitis (VA) remained limited.

Differences in definitions of infections, drug exposures (types and dosages) and methodologies can also attribute to inconsistent estimations of infection risks across subtypes of rheumatic diseases, hindering the development of the post-diagnosis therapeutic regimens. A cohort study on 117 SSc patients from Thailand found that the incidence rate of infection was 20.3 per 100 patient-years (15.6–26.0) and the serious infection was 11.0 per 100 patient-years [7]. Results from five rheumatoid arthritis (RA) registries showed that age/sex-standardized rates of hospitalized infection were consistent across registries (range 1.14–1.62 per 100 patient-years) [8] but in another study in 31,801 RA patients who reported prior exposure to a biologic agent, the one-year incidence rates of hospitalized infection ranged from 13.1 to 18.7 per 100 patient-years [9] according to different types of medications.

Prior studies have found that older age, commodities including chronic obstructive pulmonary disease (COPD), chronic renal disease, congestive heart failure, cardiovascular disease (CVD) and diabetes mellitus were associated with an increased risk of infection in rheumatic diseases [10–13]. Some work also developed risk models to predict rheumatic disease-related infections [14–18]. However, the comparison across different disease subtypes of rheumatic drugs as well as other risk factors on affecting infection risks remains an underexplored topic. As well, there is a limited number of observational studies related to these questions using large Electronic Health Records (EHRs) data. We believed that by using large EHRs, different risk factors and disease subtypes could be compared in a more consistent and efficient manner. Thus, more comprehensive and informative messages could be obtained through this analysis.

The objective of this study is to address the aforementioned questions using a large de-identified EHRs database by Cerner Health Facts®. We targeted to investigate the risk of serious infections and their associated risk factors for seven rheumatic disease subtypes, including RA, SLE, PMR, SS, systemic sclerosis (SSc), VA, and other diffuse connective tissue diseases (oCTD), through comparing incidence rates and conducting survival analysis. We expected the comparisons on different risks could better inform decision makings of the therapeutic strategy of treating rheumatic diseases.

Materials And Methods

Dataset

We conducted a retrospective study of patients visited between January 2000 and 2017 using Electronic Health Records (EHRs) from the Cerner Health Facts® database. Cerner Health Facts® (version 2017) is a de-identified EHRs database that consists of over 600 hospitals and clinics in the United States, represents over 68 million unique patients, and includes longitudinal data from 2000–2017. The database consists of patient-level data including demographics, encounters, diagnoses, procedures, lab results, medication orders, medication administration, vital signs, microbiology, surgical cases, other clinical observations, and health systems attributes. The study was approved by the Institutional Review Board (IRB) of UTHealth. The IRB waived the requirement for informed consent.

Cohort Identification

The study population included individuals who were aged 18 years or older with at least 6 months of continuous enrollment in the database. The rheumatic diseases discussed in this study were all identified by International Classification of Disease, Clinical Modification, 9th or 10th revision (ICD-9 and ICD-10) codes. Details for all the ICD codes used in this study are listed in Supplements.

The *index date* (cohort entry) was defined as the date of the first disease-modifying antirheumatic drugs (DMARDs) (including glucocorticoids (GCs)) prescription and with a diagnosis of at least one rheumatic diseases. Patients who never used any of these drugs during the study period were excluded since they might represent a group of asymptomatic or mild cases that were not representative of the severe cases to whom these analyses were intended.

The event to assess is the first time of serious infection starting from the index date. A two-years' time window after the index date was defined as the *follow-up observation period*. *Serious infections* were defined as infections requiring either hospitalization or treatment with intravenous (IV) antibiotics. Patients were excluded if the first infection occurred within 30 days after the index date. To select the population, we identified patients who experienced a hospitalization with an infection discharge diagnosis in any position (primary or non-primary) on the hospital claim using diagnosis codes from ICD-9 and ICD-10. Follow-up extended from the entry date until the earliest of the following dates (the *end date*): 1) the first serious infection, 2) the last visit within 24 months. Each patient was then tagged as *Infection* or *Not* according to the existence of serious infection during the follow-up period. We identified the records of 12 months before the index date to gather the baseline information of each patient's comorbidities, including ischemic heart disease, cerebrovascular disease (CVD), chronic obstructive pulmonary disease (COPD), diabetes mellitus, and renal dysfunction, etc. (Table 1). Patients were excluded if they had a history of malignancy during the baseline period. Figure 1 shows the pipeline of cohort selection.

Table 1
Descriptive analysis of the overall cohort.

	Non-infection (n, %)	Infection (n, %)	Total (n, %)	Infection rate %	P-value
<i>Sex</i>					
Female	29,707 (77.96)	6,750 (81.25)	36,457 (78.55)	18.51	< 0.001
Male	8,396 (22.04)	1,558 (18.75)	9,954 (21.45)	15.65	
<i>Age group</i>					
Adults	7,889 (20.70)	1,750 (21.06)	9,639 (20.77)	18.16	< 0.001
Middle-aged	15,115 (39.67)	2,743 (33.02)	17,858 (38.48)	15.36	
Aged	15,099 (39.62)	3,815 (45.92)	18,914 (40.75)	20.17	
<i>Race</i>					
Caucasian	28,926 (75.92)	5,900 (71.02)	34,826 (75.04)	16.94	< 0.001
African American	5,870 (15.41)	1,689 (20.33)	7,559 (16.28)	22.34	
Asian	409 (1.07)	93 (1.12)	502 (1.08)	18.53	
Native American	745 (1.96)	222 (2.67)	967 (2.08)	22.96	
Hispanic	482 (1.26)	119 (1.43)	601 (1.29)	19.80	
Others	1,671 (4.39)	285 (3.43)	1,956 (4.21)	14.57	
<i>Comorbidities</i>					
COPD	2,012 (5.23)	720 (8.67)	2,732 (5.89)	26.35	< 0.001
asthma	1,762 (4.62)	521 (6.27)	2,283 (4.92)	22.82	< 0.001
chronic renal disease	1,902 (4.99)	761 (9.16)	2,663 (5.74)	28.58	< 0.001
diabetes mellitus	3,713 (9.74)	1,219 (14.67)	4,932 (10.63)	24.72	< 0.001
hypertension	9,738 (25.56)	2,753 (33.14)	12,491 (26.91)	22.04	< 0.001
congestive heart failure	1,394 (3.66)	620 (7.46)	2,263 (4.34)	27.40	< 0.001
CVD	3,366 (8.83)	1,057 (12.72)	4,423 (9.53)	23.90	< 0.001
osteoporosis	1,528 (4.01)	401 (4.83)	1,929 (4.16)	20.79	< 0.001
chronic anemia	3,297 (8.65)	1,101 (13.25)	4,398 (9.48)	25.03	< 0.001
<i>Infection before index</i>	1,730 (68.14)	809 (31.86)	2,539 (5.47)	31.86	< 0.001
<i>Medication</i>					
non-biologic DMARDs	13,189 (34.61)	3,430 (41.29)	16,619 (35.81)	20.63	< 0.001
hydroxychloroquine	8,113 (21.29)	1,957 (23.56)	10,070 (21.69)	19.43	< 0.001
methotrexate	3,309 (6.68)	943 (11.35)	4,252 (9.16)	22.18	< 0.001
sulfasalazine	914 (2.40)	210 (2.53)	1,124 (2.42)	18.68	< 0.001
leflunomide	983 (2.58)	303 (3.65)	1,286 (2.77)	23.56	< 0.001
cyclophosphamide	198 (0.52)	84 (1.01)	282 (0.61)	29.79	< 0.001
azathioprine	987 (2.59)	298 (3.59)	1,285 (2.77)	23.19	< 0.001
mycophenolate mofetil	1,065 (2.80)	425 (5.12)	1,490 (3.21)	28.52	< 0.001
tacrolimus	342 (0.90)	117 (1.41)	459 (0.99)	25.49	< 0.001
biologic DMARDs	1,314 (3.45)	202 (2.43)	1,516 (3.27)	13.32	< 0.001
GCs	32,578 (84.50)	7,417 (89.28)	39,995 (86.18)	18.54	< 0.001
% in column <i>Total</i> shows the percentage in the overall cohort, while % in Infection and Non-infection columns show the percentages in the infection or non-infection populations.					

Variables

We defined patients' demographic characteristics, comorbidities and drug usage (including biologic DMARDs, nonbiologic DMARDs and GCs) as the basic variables for statistical analysis. Also, we added *whether a patient has infections before the index* (during a 6-months' time window before index) as an important indicator since one would likely to be infected again if with a history of infection [19]. Since only a small proportion of patients had their BMI records, we excluded BMI from the primary analysis but discussed the inclusion of it in a sensitivity analysis.

Age was defined as the age in years on the index date and further categorized into three groups according to the World Health Organization (WHO)'s taxonomy (Adult: $18 \leq \text{age} \leq 44$, Middle-aged: $45 \leq \text{age} \leq 64$, Aged: $\text{age} \geq 65$). To address the inconsistent records of sex and race across encounters per patient, these variables were defined using the mode value (most frequently appeared) calculated based on all available records to reduce any inconsistencies. For example, if there are three records for a patient A and two of them indicate A's sex is male but one is female, we coded A as a male. Races were further categorized into *Caucasian*, *African American*, *Native American*, *Hispanic*, *Asian* and *Other* according to their proportions among the cohort. A patient's BMI was extracted as the average value of all recorded BMIs in the follow-up observation period. Included comorbidities and usage of medications were defined in supplement Table 3 according to the ICD codes or generic names of drugs and represented as binary variables.

Table 3
Adjusted hazard ratios (HR) of different variables on each sub-cohort.

	RA	p	SLE	p	PMR	p	SS	p	SSc	p	VA	p	oCTD	p	Overall
<i>gender-male</i>	0.88 (0.82–0.95)	<.001	0.76 (0.66–0.88)	<.001	0.72 (0.60–0.86)	<.001	0.73 (0.43–1.23)	.234	1.07 (0.75–1.52)	.713	1.02 (0.85–1.22)	.713	0.88 (0.82–0.95)	<.001	0.85 (0.81–0.90)
<i>African American</i>	1.12 (0.95–1.32)	.183	1.29 (1.05–1.58)	.014	0.78 (0.41–1.48)	.443	1.33 (0.52–3.40)	.553	1.50 (0.79–2.87)	.219	1.38 (0.82–2.32)	.219	1.12 (0.95–1.32)	.183	1.26 (1.11–1.42)
<i>Asian</i>	0.80 (0.57–1.12)	.201	1.17 (0.80–1.70)	.424	1.08 (0.46–2.54)	.858	1.14 (0.22–5.92)	.878	0.00 (-)	.989	1.64 (0.73–3.69)	.989	0.80 (0.57–1.12)	.201	1.07 (0.85–1.35)
<i>Caucasian</i>	0.96 (0.82–1.12)	.585	0.94 (0.76–1.14)	.514	0.72 (0.41–1.25)	.247	1.11 (0.46–2.71)	.816	1.13 (0.60–2.11)	.706	1.17 (0.72–1.92)	.706	0.96 (0.82–1.12)	.585	0.96 (0.85–1.08)
<i>Hispanic</i>	0.97 (0.72–1.31)	.854	1.17 (0.83–1.64)	.379	0.51 (0.14–1.78)	.289	1.53 (0.44–5.40)	.506	2.13 (0.85–5.34)	.106	2.29 (1.06–4.92)	.106	0.97 (0.72–1.31)	.854	1.14 (0.92–1.41)
<i>Native American</i>	1.37 (1.11–1.69)	.004	1.04 (0.70–1.53)	.847	0.48 (0.11–2.12)	.331	1.49 (0.48–4.57)	.488	1.31 (0.47–3.63)	.604	2.29 (1.17–4.48)	.604	1.37 (1.11–1.69)	.004	1.33 (1.12–1.59)
<i>Mid-aged</i>	0.90 (0.82–0.98)	.018	0.84 (0.77–0.92)	<.001	1.07 (0.51–2.25)	.859	0.89 (0.64–1.23)	.473	1.09 (0.80–1.50)	.571	1.03 (0.80–1.33)	.571	0.90 (0.82–0.98)	.018	0.84 (0.79–0.90)
<i>Aged</i>	1.33 (1.21–1.45)	<.001	0.98 (0.87–1.11)	.749	1.88 (0.93–3.80)	.080	1.18 (0.83–1.67)	.350	1.14 (0.80–1.63)	.467	1.47 (1.15–1.87)	.467	1.33 (1.21–1.45)	<.001	1.21 (1.14–1.28)
<i>COPD*</i>	1.25 (1.13–1.38)	<.001	1.17 (0.98–1.41)	.089	1.14 (0.87–1.49)	.347	1.50 (0.97–2.32)	.068	1.61 (1.06–2.46)	.026	1.41 (1.04–1.93)	.026	1.25 (1.13–1.38)	<.001	1.24 (1.14–1.34)
<i>asthma</i>	1.05 (0.93–1.18)	.431	1.05 (0.90–1.23)	.551	1.05 (0.72–1.54)	.784	1.10 (0.73–1.64)	.650	0.95 (0.58–1.56)	.836	1.53 (1.09–2.16)	.836	1.05 (0.93–1.18)	.431	1.07 (0.98–1.17)
<i>chronic renal disease</i>	1.29 (1.14–1.46)	<.001	1.07 (0.93–1.24)	.324	1.30 (0.99–1.69)	.055	1.00 (0.58–1.70)	.988	0.90 (0.60–1.37)	.630	1.29 (0.99–1.67)	.630	1.29 (1.14–1.46)	<.001	1.23 (1.13–1.34)
<i>diabetes mellitus</i>	1.21 (1.11–1.31)	<.001	1.28 (1.11–1.46)	<.001	1.14 (0.92–1.41)	.220	0.87 (0.58–1.31)	.511	0.95 (0.61–1.46)	.801	1.34 (1.05–1.70)	.801	1.21 (1.11–1.31)	<.001	1.19 (1.11–1.27)
<i>hypertension</i>	0.99 (0.92–1.07)	.838	0.98 (0.88–1.09)	.681	0.91 (0.76–1.09)	.322	1.31 (0.99–1.74)	.060	1.25 (0.92–1.69)	.150	0.85 (0.69–1.05)	.150	0.99 (0.92–1.07)	.838	0.99 (0.93–1.04)
<i>congestive heart failure*</i>	1.49 (1.32–1.67)	<.001	1.20 (0.99–1.46)	.068	1.62 (1.25–2.12)	<.001	1.64 (0.99–2.72)	.056	1.04 (0.68–1.58)	.871	1.25 (0.90–1.74)	.871	1.49 (1.32–1.67)	<.001	1.42 (1.30–1.56)
<i>CVD</i>	1.04 (0.95–1.14)	.418	1.09 (0.94–1.28)	.261	1.16 (0.94–1.42)	.157	0.99 (0.66–1.50)	.965	1.24 (0.85–1.81)	.264	1.05 (0.82–1.34)	.264	1.04 (0.95–1.14)	.418	1.08 (1.00–1.16)
<i>osteoporosis</i>	0.91 (0.80–1.03)	.134	1.15 (0.91–1.45)	.240	0.77 (0.57–1.03)	.081	1.32 (0.82–2.12)	.248	1.17 (0.63–2.17)	.621	0.91 (0.61–1.36)	.621	0.91 (0.80–1.03)	.134	0.91 (0.82–1.01)
<i>chronic anemia*</i>	1.14 (1.03–1.25)	.007	1.14 (1.01–1.30)	.039	1.23 (0.98–1.53)	.072	1.32 (0.92–1.91)	.132	1.20 (0.86–1.68)	.277	1.21 (0.96–1.52)	.277	1.14 (1.03–1.25)	.007	1.14 (1.06–1.22)
<i>GCs*</i>	1.65 (1.50–1.81)	<.001	1.48 (1.30–1.68)	<.001	1.57 (0.78–3.15)	.206	1.69 (1.19–2.40)	.003	1.35 (0.92–1.99)	.126	1.53 (0.87–2.69)	.126	1.65 (1.50–1.81)	<.001	1.64 (1.52–1.76)
<i>biologic DMARD</i>	0.83 (0.71–0.97)	.019	0.64 (0.41–0.99)	.044	0.21 (0.03–1.53)	.124	0.78 (0.32–1.91)	.586	1.20 (0.52–2.79)	.668	1.32 (0.84–2.07)	.668	0.83 (0.71–0.97)	.019	0.82 (0.71–0.95)
<i>nonbiologic DMARD*</i>	1.50 (1.41–1.59)	<.001	1.14 (1.05–1.24)	.003	1.11 (0.85–1.46)	.431	1.50 (1.16–1.93)	.002	1.45 (1.11–1.88)	.006	1.29 (1.06–1.58)	.006	1.50 (1.41–1.59)	<.001	1.39 (1.32–1.45)

*: Variables with HR > 1 across all disease types. The maximum value in each row is tagged as **bold**.

	RA	p	SLE	p	PMR	p	SS	p	SSc	p	VA	p	oCTD	p	Overall
<i>Infection before index*</i>	1.75 (1.59–1.94)	<.001	1.82 (1.60–2.08)	<.001	1.64 (1.24–2.18)	<.001	1.68 (1.13–2.49)	.010	1.25 (0.80–1.97)	.324	1.08 (0.81–1.46)	.324	1.75 (1.59–1.94)	<.001	1.75 (1.62–1.88)

*: Variables with HR > 1 across all disease types. The maximum value in each row is tagged as **bold**.

To facilitate the survival analysis, we defined the disease duration (the “survival” time) as the number of days between the index date and the end date (defined above).

Statistical Analysis

Fisher exact test was applied in the univariate analyses to explore the distribution of infections across subtypes of rheumatic diseases including RA, SLE, PMR, SS, SSc, VA and oCTD. The statistical significance was defined as a p-value of less than 0.05. The crude incidence rate and the cumulated incidence curve were calculated on the overall cohort as well as each sub-cohort. The crude incidence rate was calculated as the number of total new cases developed over the follow-up time divided by the total number of patients at risk at baseline. Cumulated incidence curve was computed using Kaplan–Meier estimator, which is a non-parametric method representing conditional probability of survival or risk at specific time interval. In the multivariable analyses, Cox proportional hazard models were employed on each individual cohort to evaluate the risk of serious infection adjusted for age, race, sex and other predictors. Sensitivity analyses were further conducted using the cox proportional hazard model on a sub-cohort with record BMI values as well as separating different types of nonbiologic DMARDs. The analyses were performed using R 3.4.1.

Results

Table 1 is the descriptive analysis of the cohort. Female patients accounted for a larger proportion of patients than males (78.55% vs 21.45%). They also presented a higher percentage of infections than males (18.51% vs 15.65%). Aged patients showed the highest infection rate (20.17%), while adult patients were the second highest (18.16%). Native American had a much larger infection rate (22.96%) than the others, while Caucasian who represented the largest population (75.04%), reported the second lowest infection rate (16.94%). The prevalence of infection varied by comorbidities (all $p < 0.001$) where chronic renal disease (28.58%) and congestive heart failure (27.40%) groups were with the highest proportion of infection than others. Patients with previous infections had a very high rate of infection again (31.86%). For medication, GCs were still the most widely used for rheumatic patients (86.18%), and in nonbiologic DMARDs (35.81%), hydroxychloroquine was the most frequently used (21.69%). Among them, cyclophosphamide (29.79%) and mycophenolate mofetil (28.52%) were with the highest infection rates. In comparison, biologic DMARDs were seldomly used in this cohort (3.27%), but the corresponding medication group shows a much lower infection rate (13.32%) than the other medications.

Table 2 displays the crude incidence rate (IR) of serious infections among the overall cohort as well as seven types of rheumatic diseases. Among the 46,411 rheumatic patients, 8,308 (17.90%) experienced serious hospitalized infections during the observed follow-up period. The overall IR was 8.95 per 100 patient-years. For individual rheumatic disease categories, those with SSc, VA, SLE observed the highest IRs (10.89, 10.82, 10.13 cases per 100 patient-years) at the end of follow-up time. RA patients who represented the majority (62.3%) of the rheumatic population in this study, observed IR at 8.6 cases per 100 patient-years. oCTD and SS also had a lower IR than the overall IR which were 7.71 and 8.32 cases per 100 patient-years respectively. Note that patients might have overlapped diseases (4,852 patients).

Table 2
The numbers of patients and the infection rates per 100 patient-years on different cohorts.

Cohort	No. of infections	No. of patients	Crude IR per 100 patient-years% (95% CI)
SSc	280	1,286	10.89 (9.61–12.16)
VA	605	2,796	10.82 (9.96–11.68)
SLE	2,465	12,167	10.13 (9.73–10.53)
PMR	688	3,740	9.20 (8.51–9.89)
RA	4,976	28,925	8.60 (8.36–8.84)
SS	296	1,779	8.32 (7.37–9.27)
oCTD	179	1,161	7.71 (6.58–8.84)
Overall	8,308	46,411	8.95 (8.76–9.14)

Figure 2 compares the cumulative incidences across seven rheumatic diseases and significant differences were observed between their incidence of infections ($p < 0.001$). We observed distinct distributions on the infection rate across these cohorts. Similar to Table 2, patients with SSc, VA and SLE demonstrated the highest overall cumulative incidences compared to other groups. In terms of growth trend, patients of VA, SSc, and SLE seemed to infection more rapidly than those of PMR, RA, oCTD and SS. SS tends to have a lower overall incidence of infection.

Table 3 shows the subgroup analyses on sub-cohorts of rheumatic diseases using the Cox proportional hazard model with the same set of covariates. Of note that several variables such as GCs, COPD, congestive heart failure, chronic anemia and infection before index were observed a similar level of correlation with increased risks in infections across all the seven types of diseases as well as the overall population. However, we observed variations in the risks of some variables across sub-cohorts. For example, osteoporosis was with increased risks (HR > 1) among SLE, SS, and SSc but decreased risks among others (HR < 1). In addition, the associations of some predictors fluctuated heavily in different diseases, e.g. the usage of biological DMARDs had an HR of 1.32 in VA, but only 0.21 in PMR. Patients with some specific diseases such as VA and SS tended to be more vulnerable to infection as more factors were associated with increased risks than other diseases (according to the number of bold values). The risks for most racial categories were not statistically significant.

The results on the overall cohort may inform us of some general senses about the covariates in the general rheumatic population, where most of them were associated with increased risks in infections. Specifically, infection before index ranked top. For gender, females were more likely to be infected. For races, Caucasian had a decreasing risk but was not statistically significant ($p = 0.510$). Aged patients were observed with a larger likelihood of infection than the adult and middle-aged groups, but the middle-aged seems to be less vulnerable than the adult group. While GCs and non-biologic DMARDs were highly associated with serious infections, biological DMARDs delivered a protected effect towards infection. Most of the comorbidities as covariates showed increased, e.g. COPD (HR:1.24), or marginally increased risks, e.g. hypertension (HR:0.99).

Sensitivity Analyses

To evaluate the robustness of the findings, we conducted two sensitivity analyses: on a sub-cohort with recorded BMIs as well as the overall cohort but separating different types of non-biological DMARDs.

Figure 3 demonstrates that most variables show similar polarities with those on the overall cohort (Table 3). The only factors with reversed risks are the race of Hispanic (race 4) and Asian (race 5) but their associations were not statistically significant. Those changes might be due to the differences in the distributions of patients. Interestingly, the average BMI value had a very mild association with infection (HR:0.99, 0.99-1.00).

Some researchers discovered that therapy with nonbiologic DMARDs represented an additional factor of increased risk of infections and the risk varied depending on the different nonbiologic DMARDs used [4]. For explorative purposes, the second analysis was conducted on the overall cohort by separating each nonbiologic DMARD to see their individual effect (Table 4). While other factors didn't demonstrate much changes, the HRs revealed that most nonbiologic DMARDs including cyclophosphamide (HR:1.91, 1.16–3.14), tacrolimus (HR:1.59, 1.14–2.23), leflunomide (HR:1.51, 1.34–1.71), azathioprine (HR:1.48, 1.24–1.76), mycophenolate mofetil (HR:1.46, 1.14–1.85), methotrexate (HR:1.46, 1.36–1.58), hydroxychloroquine (HR:1.19, 1.11–1.28), and sulfasalazine (HR:1.08, 0.93–1.25), and were associated with serious infections (mostly significant with $p < 0.05$ except for sulfasalazine with $p = 0.3$). Cyclosporine A and thalidomide were not associated with enough patient samples (0 and 1 respectively) to generate meaningful statistical results.

Table 4
Adjusted hazard ratios (HR) of different variables when separating each nonbiologic DMARDs.

Variable	HR (CI)	p
gender-male	1.14 (1.06–1.22)	< .001
African American	1.12 (0.95–1.32)	.182
Asian	0.80 (0.57–1.13)	.207
Caucasian	0.97 (0.83–1.13)	.656
Hispanic	0.98 (0.73–1.31)	.888
Native American	1.37 (1.11–1.69)	.004
Mid-aged	0.90 (0.82–0.98)	.017
Aged	1.34 (1.22–1.46)	< .001
COPD	1.26 (1.14–1.39)	< .001
asthma	1.06 (0.94–1.20)	.315
chronic renal disease	1.27 (1.13–1.44)	< .001
diabetes mellitus	1.19 (1.09–1.30)	< .001
hypertension	0.99 (0.93–1.07)	.869
congestive heart failure	1.51 (1.34–1.69)	< .001
CVD	1.03 (0.94–1.14)	.500
osteoporosis	0.90 (0.80–1.02)	.100
chronic anemia	1.14 (1.03–1.25)	.007
GCs	1.55 (1.42–1.70)	< .001
biologic DMARD	0.80 (0.68–0.93)	.004
hydroxychloroquine	1.19 (1.11–1.28)	< .001
methotrexate	1.46 (1.36–1.58)	< .001
sulfasalazine	1.08 (0.93–1.25)	.300
leflunomide	1.51 (1.34–1.71)	< .001
cyclophosphamide	1.91 (1.16–3.14)	.010
azathioprine	1.48 (1.24–1.76)	< .001
mycophenolate mofetil	1.46 (1.14–1.85)	.002
tacrolimus	1.59 (1.14–2.23)	.007
infection before index	1.75 (1.62–1.88)	< .001

Discussion

The primary finding of the study was different types of rheumatic diseases were with different infection rates and distributions of risk factors.

Analyses on the real-world large U.S. cohort offered us a comprehensive understanding of infection rates of rheumatic diseases which mostly need glucocorticoids and DMARDs treatment. In general, the infection rates per 100 patient-years was relatively high (8.95) during the first two years of treatment. While for some particular disease types, such as SLE, VA, and SSc, their IRs are higher than others. It is not surprising as these three diseases tend to affect the critical organs, have more severe impaired immune functions, and need more immunosuppressive treatment regimens. For example, it was found that the impaired immune function in SLE including an impaired acute inflammatory response, the decreased number of T lymphocytes and the T-helper cell activity, and complement dysfunction contributed to the increase of the infection risk [16].

Our study validated as well as provided important complements to the current literature on the risk of serious infections associated with rheumatic diseases. Our analyses found a 20.4% prevalence of severe infection among SLE patients, which is in the range of 12–40% reported by different prior works [20–22]. Our results were a little higher than one another study which was based on a cohort of 33,565 SLE patients aged 18–64 years old [6]. The discrepancies might be owing to that we also included elder patients. In our infection group, 45.92% of patients were > 65 years, and the elder people have an increased risk of infection. Recently, a large England cohort study showed that the cumulative incidence of infection over 1 year's follow-up was 18.3% (17.9–18.7) in patients of PMR or giant cell arteritis [17]. Our results seemed much lower in PMR patients according to Table 3 (9.20, 8.51–9.89) but we included two years' data and

the target was calculating the serious infection rate rather than the all-cause infection rate. For SS, there are few studies concentrated on the infection rate [2] but we reported the rate is a bit lower (8.32, 7.37–9.27) than RA (8.60, 8.36–8.84).

A study from the European League against Rheumatism (EULAR) Scleroderma Trials and Research (EUSTAR) database reported among the non-SSc-related causes of death, infections accounted for 33% [23]. Our results also revealed a very high infection rate among SSc patients (10.89 per 100 patient-years), which reminded us to concern more about this point. Further studies about the related risk factors and how to decrease the infection should be implemented. We also discovered for the first time that certain types of rheumatic diseases such as VA and SSc tend to have infection more rapidly than others. In anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis, it was reported that severe infections developed almost in one out of four patients (23%), most of them during the first year [24]. A high dose of glucocorticoids is perhaps a significant factor in infection development, while neutropenia due to cyclophosphamide is also a contributing factor [25, 26].

Given structured EHRs only, although it is difficult to interpret all these results, we can still get some clues and pay more attention to these particular types of rheumatic diseases in their early stages of treatment in clinical practice.

Apart from common risk factors like older age, previous serious infection, and some commodities, we also found that chronic anemia was strongly associated with an increased risk of infection, i.e. HR = 1.14 (1.06–1.22) in the overall cohort. Chronic anemia is often seen in many chronic diseases, including RA and chronic renal disease. In RA, anemia is the most common comorbidity with an estimated prevalence of 33.3–59.1% [27]. Although we didn't observe a such a high prevalence according to a rough statistic using the ICD codes only (8.65% and 13.25% in the whole group and infection group, which might underestimate the numbers), chronic anemia afflicted a substantial proportion of rheumatic patients and treating anemia is strongly encouraged regardless of the underlying disease process responsible for anemia.

The drugs used in rheumatic diseases can be divided into glucocorticoids, traditional (non-biological) DMARDs, and biological DMARDs. Voluminous previous studies have validated that the usage of certain dosage of GCs would increase the risk of infections, in both RA [28, 29] and other types of rheumatic diseases such as SLE and lupus nephritis [22, 30, 31]. In line with these previous studies, our data revealed the increased risk of serious infection associated with systemic GCs (RR:1.64, 1.52–1.76) among the cohorts although there are subtle differences between different disease types. Compared with many previous studies, our study used a relatively larger database and offered a more comprehensive perspective. Withdrawal of GCs should therefore be carried out systematically for all patients receiving GCs therapies long-term, and if this is hard to implement, a change in treatment should be considered [32].

Previous discussions on the risk of infection by using nonbiologic DMARDs were not converging. In RA, the initiation of leflunomide, sulfasalazine or hydroxychloroquine may do not boost serious infections compared with methotrexate [33]. A systematic review and meta-analysis of randomized controlled trials (RCTs) also showed that methotrexate was associated with an increased risk of infection in RA (RR: 1.25; 1.01–1.56), but not in other non-RA inflammatory rheumatic diseases populations [34]. However, another population-based RA cohort in British Columbia, Canada indicated that the use of nonbiologic DMARDs, including methotrexate, did not increase the risk of infection in RA [35]. Our results confirmed that the use of most nonbiologic DMARDs, including cyclophosphamide, mycophenolate mofetil, leflunomide, methotrexate, azathioprine, hydroxychloroquine, sulfasalazine and tacrolimus increased the risk of serious infection on the overall cohort. And on larger sub-cohorts such as RA and SLE, the associations were also kept with only slight differences in the hazard ratios. Compared with results based solely on RCTs, our results are derived from the real-world data from EHRs, which were not limited by the underrepresentation of specified populations, e.g. elderly and high-risk patients, and thus might have higher generalizability.

In registry or observational studies, biologics were associated with a higher risk of serious infections, compared both to non-use of biologics and to the use of nonbiologic DMARDs [36]. A meta-analysis reported a 31% increased risk of serious infections in standard dose biologic-treated RA patients compared to nonbiologic DMARDs (OR:1.31, 1.09–1.58) [1]. In our cohort, the proportion of biological DMARDs users was relatively low (3.27%) which might have led to limited statistical power, but we observed a decreased risk (HR:0.82, 0.71–0.95), reflecting the protected effect.

This study used structured EHRs and was thus subject to some potential biases and limitations. It lacks linkage to other data sources, so cares provided by non-participating physicians were missed. With respect to the disease definition and outcome measurement, we cannot exclude some misclassifications since data were based on diagnosis codes and not validated through medical record reviews. However, the diagnoses were based on hospitalization with infection as the primary diagnosis, thus limiting potential misclassification. In addition, no direct measures of disease activity and disease severity exist within the administrative database, therefore, the impact of disease status on the DMARDs initiation could hardly be determined from this study. Also, there might exist some selection biases for GCs and DMARDs use as physicians tended to treat more severe patients with (higher dosages of) GCs and more powerful DMARDs but we didn't include the dosage and duration of drug usages. Finally, we cannot rule out the potential for residual confounding, since we selected variables based on experience and reports and it is possible that the results remained affected by unmeasured confounders. The strengths of our data were that they were the real-world data and the sample size is large enough to allow an adequate number of events. We plan to include the dosage and duration of drugs for in-depth analyses in the future.

Conclusion

These comprehensive analyses presented novel findings while confirmed previous data from observational studies using a large administrative EHR database. Patients with rheumatic diseases have a significant risk of serious infections and require enhanced vigilance in the management of their pharmacotherapy and comorbidities. While patients of different rheumatic diseases were observed variances in infection rates and risk factors, the therapeutic strategies should be adjusted accordingly.

Abbreviations

CHF
congestive heart failure
COPD
chronic obstructive pulmonary disease
CVD
cerebrovascular disease
DMARDs
disease-modifying antirheumatic drugs
EHR
Electronic Health Records
GC
glucocorticoid
HR
hazard ratio
oCTD
other diffuse connective tissue diseases
PMR
polymyalgia rheumatica
RA
rheumatoid arthritis
SLE
systemic lupus erythematosus
SS
Sjögren's syndrome
SSc
systemic sclerosis
VA
systemic vasculitis

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board (IRB) of UTHealth. The IRB waived the requirement for informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from UTHealth and Cerner but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of UTHealth and Cerner.

Competing interests

The authors declare that they have no competing interests.

Funding

Supported by Cerner and the School of Biomedical Informatics Data Service, UTHealth, which is supported in part by the CPRIT grant, RP170668.

Authors' contributions

HX and JD conceived the research project. JD, YX, and YZ designed the analysis. JD led the writing with substantial inputs from YX. YX conducted the statistical analysis with substantial inputs from CZ, YZ and JD. YZ extracted data from Cerner and cleaned the data. XH did the preliminary analysis and helped collect the required definitions. HX and CT proofread the paper and provided valuable suggestions. All the authors have read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Singh JA, Cameron C, Noorbaloochi S, Cullis T, Tucker M, Christensen R, et al. Risk of serious infection in biological treatment of patients with rheumatoid arthritis: a systematic review and meta-analysis. *The Lancet*. 2015;386(9990):258–65.
2. Juarez M, Misischia R, Alarcon GS. Infections in systemic connective tissue diseases: systemic lupus erythematosus, scleroderma, and polymyositis/dermatomyositis. *Rheum Dis Clin North Am*. 2003;29(1):163–84.
3. Falagas ME, Manta KG, Betsi GI, Pappas G. Infection-related morbidity and mortality in patients with connective tissue diseases: a systematic review. *Clin Rheumatol*. 2007;26(5):663–70.
4. Di Franco M, Lucchino B, Spaziante M, Iannuccelli C, Valesini G, Iaiani G. Lung Infections in Systemic Rheumatic Disease: Focus on Opportunistic Infections. *International journal of molecular sciences*. 2017;18(2).
5. Crowson CS, Hoganson DD, Fitz-Gibbon PD, Matteson EL. Development and validation of a risk score for serious infection in patients with rheumatoid arthritis. *Arthritis Rheum*. 2012;64(9):2847–55.
6. Feldman CH, Hiraki LT, Winkelmayr WC, Marty FM, Franklin JM, Kim SC, et al. Serious infections among adult Medicaid beneficiaries with systemic lupus erythematosus and lupus nephritis. *Arthritis Rheumatol*. 2015;67(6):1577–85.
7. Foocharoen C, Siriphannon Y, Mahakkanukrauh A, Suwannaroj S, Nanagara R. Incidence rate and causes of infection in Thai systemic sclerosis patients. *International journal of rheumatic diseases*. 2012;15(3):277–83.
8. Yamanaka H, Asklung J, Berglund N, Franzen S, Frisell T, Garwood C, et al. Infection rates in patients from five rheumatoid arthritis (RA) registries: contextualising an RA clinical trial programme. *RMD Open*. 2017;3(2):e000498.
9. Yun H, Xie F, Delzell E, Levitan EB, Chen L, Lewis JD, et al. Comparative Risk of Hospitalized Infection Associated With Biologic Agents in Rheumatoid Arthritis Patients Enrolled in Medicare. *Arthritis Rheumatol*. 2016;68(1):56–66.
10. Schenfeld J, Iles J, Trivedi M, Accortt NA. Dose relationship between oral glucocorticoids and tumor necrosis factor inhibitors and the risk of hospitalized infectious events among patients with rheumatoid arthritis. *Rheumatology international*. 2017;37(7):1075–82.
11. Curtis JR, Winthrop K, O'Brien C, Ndlovu MN, de Longueville M, Haraoui B. Use of a baseline risk score to identify the risk of serious infectious events in patients with rheumatoid arthritis during certolizumab pegol treatment. *Arthritis research therapy*. 2017;19(1):276.
12. Zink A, Manger B, Kaufmann J, Eisterhues C, Krause A, Listing J, et al. Evaluation of the RABBIT Risk Score for serious infections. *Annals of the rheumatic diseases*. 2014;73(9):1673–6.
13. van Dartel SA, Fransen J, Kievit W, Dutmer EA, Brus HL, Houtman NM, et al. Predictors for the 5-year risk of serious infections in patients with rheumatoid arthritis treated with anti-tumour necrosis factor therapy: a cohort study in the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry. *Rheumatology*. 2013;52(6):1052–7.
14. Listing J, Gerhold K, Zink A. The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment. *Rheumatology (Oxford)*. 2013;52(1):53–61.
15. Curtis JR, Xie F, Chen L, Muntner P, Grijalva CG, Spettell C, et al. Use of a disease risk score to compare serious infections associated with anti-tumor necrosis factor therapy among high- versus lower-risk rheumatoid arthritis patients. *Arthritis Care Res (Hoboken)*. 2012;64(10):1480–9.
16. Danza A, Ruiz-Irastorza G. Infection risk in systemic lupus erythematosus patients: susceptibility factors and preventive strategies. *Lupus*. 2013;22(12):1286–94.
17. Wu J, Keeley A, Mallen C, Morgan AW, Pujades-Rodriguez M. Incidence of infections associated with oral glucocorticoid dose in people diagnosed with polymyalgia rheumatica or giant cell arteritis: a cohort study in England. *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2019;191(25):E680–8.
18. Fardet L, Petersen I, Nazareth I. Common Infections in Patients Prescribed Systemic Glucocorticoids in Primary Care: A Population-Based Cohort Study. *PLoS Med*. 2016;13(5):e1002024.
19. Widdifield J, Bernatsky S, Paterson JM, Gunraj N, Thorne JC, Pope J, et al. Serious infections in a population-based cohort of 86,039 seniors with rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2013;65(3):353–61.
20. Goldblatt F, Chambers S, Rahman A, Isenberg DA. Serious infections in British patients with systemic lupus erythematosus: hospitalisations and mortality. *Lupus*. 2009;18(8):682–9.
21. Gonzalez-Echavarri C, Capdevila O, Espinosa G, Suarez S, Marin-Ballve A, Gonzalez-Leon R, et al. Infections in newly diagnosed Spanish patients with systemic lupus erythematosus: data from the RELES cohort. *Lupus*. 2018;27(14):2253–61.
22. Rua-Figueroa I, Lopez-Longo J, Galindo-Izquierdo M, Calvo-Alen J, Del Campo V, Olive-Marques A, et al. Incidence, associated factors and clinical impact of severe infections in a large, multicentric cohort of patients with systemic lupus erythematosus. *Seminars in arthritis rheumatism*. 2017;47(1):38–45.
23. Tyndall AJ, Bannert B, Vonk M, Airo P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Annals of the rheumatic diseases*. 2010;69(10):1809–15.
24. McGregor JG, Negrete-Lopez R, Poulton CJ, Kidd JM, Katsanos SL, Goetz L, et al. Adverse events and infectious burden, microbes and temporal outline from immunosuppressive therapy in antineutrophil cytoplasmic antibody-associated vasculitis with native renal function. *Nephrology dialysis transplantation: official publication of the European Dialysis Transplant Association - European Renal Association*. 2015;30(Suppl 1):i171–81.

25. Charlier C, Henegar C, Launay O, Pagnoux C, Berezne A, Bienvenu B, et al. Risk factors for major infections in Wegener granulomatosis: analysis of 113 patients. *Annals of the rheumatic diseases*. 2009;68(5):658–63.
26. David Morgan M, Richter A, Al-Ali S, Flint J, Yiannakis C, Drayson M, et al. Association of Low B Cell Count and IgG Levels With Infection, and Poor Vaccine Response With All-Cause Mortality in an Immunosuppressed Vasculitis Population. *Arthritis care research*. 2016;68(6):853–60.
27. Wilson A, Yu HT, Goodnough LT, Nissenson AR. Prevalence and outcomes of anemia in rheumatoid arthritis: a systematic review of the literature. *The American journal of medicine*. 2004;116(Suppl 7A):50S–57S.
28. 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 research therapy*. 2011;13(4):R139.
29. 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;71(7):1128–33.
30. Singh JA, Hossain A, Kotb A, Wells G. Risk of serious infections with immunosuppressive drugs and glucocorticoids for lupus nephritis: a systematic review and network meta-analysis. *BMC medicine*. 2016;14(1):137.
31. Herrinton LJ, Liu L, Goldfen R, Michaels MA, Tran TN. Risk of Serious Infection for Patients with Systemic Lupus Erythematosus Starting Glucocorticoids with or without Antimalarials. *The Journal of rheumatology*. 2016;43(8):1503–9.
32. Ruysen-Witrand A, Constantin A. Controversies in rheumatoid arthritis glucocorticoid therapy. *Joint Bone Spine*. 2018;85(4):417–22.
33. Grijalva CG, Kaltenbach L, Arbogast PG, Mitchel EF Jr, Griffin MR. Initiation of rheumatoid arthritis treatments and the risk of serious infections. *Rheumatology*. 2010;49(1):82–90.
34. Ibrahim A, Ahmed M, Conway R, Carey JJ. Risk of Infection with Methotrexate Therapy in Inflammatory Diseases: A Systematic Review and Meta-Analysis. *Journal of clinical medicine*. 2018;8(1).
35. Lacaille D, Guh DP, Abrahamowicz M, Anis AH, Esdaile JM. Use of nonbiologic disease-modifying antirheumatic drugs and risk of infection in patients with rheumatoid arthritis. *Arthritis rheumatism*. 2008;59(8):1074–81.
36. Singh JA. Infections With Biologics in Rheumatoid Arthritis and Related Conditions: a Scoping Review of Serious or Hospitalized Infections in Observational Studies. *Current rheumatology reports*. 2016;18(10):61.

Figures

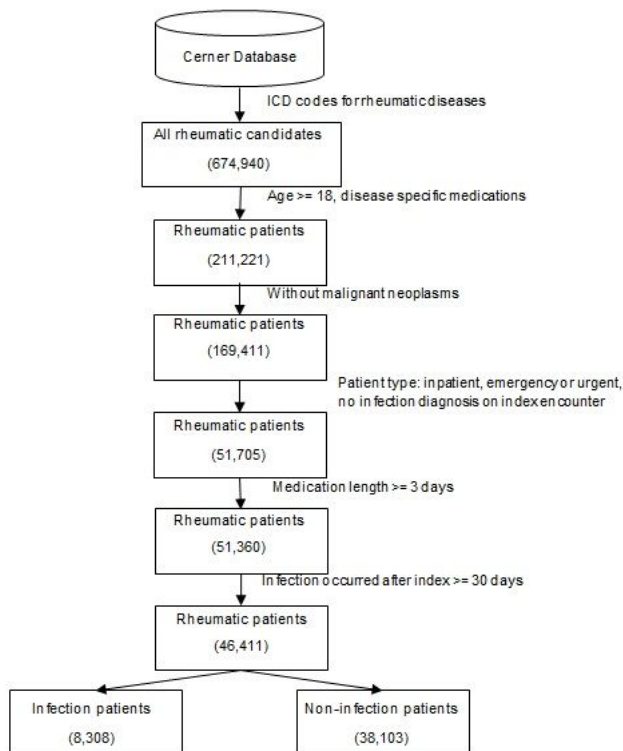


Figure 1

The cohort selection process

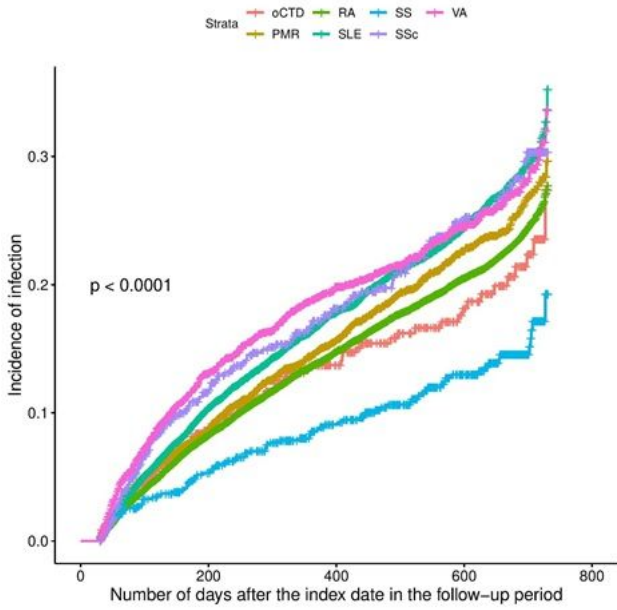


Figure 2
 Cumulative incidence of serious infection during the follow-up period. We excluded 4,852 patients with overlapped rheumatic diseases to facilitate the comparison. There might be some discrepancies with results in Table 3 due to censoring data and differences in computation.

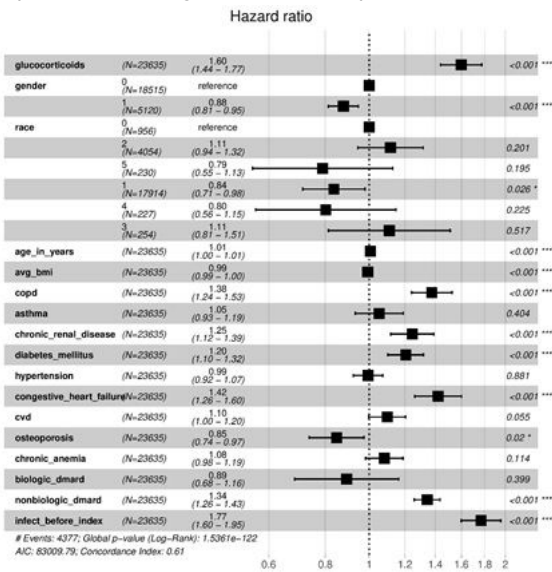


Figure 3
 The results of the Cox model on the sub-cohort with recorded BMIs from the overall population. We factorized the genders as 0: Female, 1: Male, ages as 0: Adult, 1: Middle-aged, 2: Aged and races as 0: Other, 1: Caucasian, 2: African American, 3: Native American, 4: Hispanic, 5: Asian.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement.docx](#)