Current Medical Research and Opinion

ISSN: 0300-7995 (Print) 1473-4877 (Online) Journal homepage: http://www.tandfonline.com/loi/icmo20

Co-occurrence and comorbidities in patients with immune-mediated inflammatory disorders: an exploration using US healthcare claims data, 2001–2002

Don Robinson Jr., Monica Hackett, John Wong, Alexa B. Kimball, Russell Cohen, Mohan Bala & for the IMID Study Group

To cite this article: Don Robinson Jr., Monica Hackett, John Wong, Alexa B. Kimball, Russell Cohen, Mohan Bala & for the IMID Study Group (2006) Co-occurrence and comorbidities in patients with immune-mediated inflammatory disorders: an exploration using US healthcare claims data, 2001–2002, Current Medical Research and Opinion, 22:5, 989-1000, DOI: 10.1185/030079906X104641

To link to this article: https://doi.org/10.1185/030079906X104641

Published online: 19 Apr 2006.

CURRENT MEDICAL RESEARCH AND OPINION®
VOL. 22, NO. 5, 2006, 989–1000

0300-7995
doi:10.1185/030079906X104641

ORIGINAL ARTICLE

Co-occurrence and comorbidities in patients with immune-mediated inflammatory disorders: an exploration using US healthcare claims data, 2001–2002

Don Robinson, Jr.[a], Monica Hackett[a], John Wong[b], Alexa B. Kimball[c], Russell Cohen[d] and Mohan Bala[a] for the IMID Study Group*

[a] *MSPH, Centocor, Inc., Malvern, PA, USA*

[b] *Tufts-New England Medical Center, Boston, MA, USA*

[c] *Massachusetts General and Brigham and Women's Hospitals, Boston, MA, USA*

[d] *University of Chicago Hospital, Chicago, IL, USA*

Address for correspondence: Don Robinson, Jr., Centocor, Inc., 200 Great Valley Parkway, Malvern, PA 19355, USA. Tel.: +1-610-240-8468; Fax: +1-610-651-6717; email: drobins4@cntus.jnj.com

Key words: Autoimmune – Diseases – Disorders – Epidemiology – Immunology – Inflammation

ABSTRACT

Objective: Research in immune-mediated inflammatory disorders (IMIDs) suggests that several diseases share disruptions in key cytokines. A common pathogenesis may present as similar patterns of disease co-occurrence and comorbidity, which could be observed through the analysis of healthcare claims datasets.

Methods: Adult patients continuously enrolled from 2001–2002 were identified in two US healthcare datasets containing medical and drug claims from health plans and self-insured employers. Patients with treatment records indicating an IMID were selected (e.g., rheumatoid arthritis, psoriasis, Crohn's disease); controls for each disorder were matched 3:1 based on age, gender, region, and previous insurance coverage. IMID cohorts and comorbidities were identified using International Classification of Diseases, 9th revision codes. Prevalence relative risk was used to assess co-occurrence and comorbidity rates in IMID cohorts and controls. Medical and drug utilization patterns were also explored.

Results: Findings were similar across the two datasets. IMID patients represented about 4% of the population; specific IMID prevalence matched the epidemiology literature. Patients with at least one IMID had a higher risk for another IMID when compared to controls. The risk for infectious, renal, liver, and ulcerative comorbidities was also elevated. Selected drug utilization patterns confirmed comorbidity findings. IMID patients used more healthcare resources compared to controls; findings were robust under sensitivity analyses.

Conclusions: IMID patients were generally more likely than controls to have another IMID, supporting the concept that the diseases are related. These patients also had higher comorbidity rates. Findings may be limited by the nature of claims datasets and the confounding effect of current treatments. Prospective studies are needed to determine whether IMIDs have a common pathogenesis.

* Members of the IMID Study Group are listed at the end of this paper

Introduction

In recent years, abnormal levels of multiple key cytokines have been found across autoimmune diseases that clinically present with inflammation[1–3]. These findings have led to speculation that there might be a shared pathological process involving cytokine dysregulation. The discovery of a common biological mechanism underlying these disorders would suggest that there may be a meta-disease category based on autoimmune, inflammatory criteria; a collection of diseases that can be referred to as immune-mediated inflammatory disorders (IMIDs)[1].

While the co-occurrence of some IMIDs has been observed, comprehensive epidemiology is lacking, aside from some common human leukocyte antigen (HLA) subtypes[4–11]. If IMIDs share a cytokine-based pathogenesis, then they may co-occur in afflicted patients[1,2,8,9,12]. The systemic nature of this dysfunctional mechanism may also produce common morbidities across the various IMIDs. The biological mechanisms and disease characteristics of various IMIDs are now being explored in laboratory and clinical studies. However, these studies require substantial time to conduct due to the complexity of biological pathways and the slow pace of the natural, chronic disease progression. The lack of adequate IMID epidemiology limits the understanding of co-occurrence patterns, comorbid illnesses, and overall disease burden. It also may obscure the diagnostic and treatment implications represented by this possible meta-disease.

In the US, patient-based electronic administrative claims data are available that can allow these issues to be investigated much more quickly than a longitudinal study. While originally developed for reimbursement purposes, these data also contain patient-level clinical information, such as diagnostic and medical procedure codes, as well as prescribed drug information. This study was conducted to examine the hypothesis of an IMID meta-disease and the attendant comorbidities that may result.

To evaluate the characteristics of the IMID population in detail, two national datasets were used, containing patient-level medical and prescribed drug claims, in a retrospective claims database analysis. Both datasets have often been used for exploring disease epidemiology, practice patterns, and cost of illness. In this study, the following research questions were explored:

- What is the US treatment prevalence of IMIDs during this 2-year period?
- How often does more than one IMID appear in the patient treatment record?
- Do IMID patients have higher occurrences of certain comorbidities than non-IMID patients?

Methods

Patients

The eligible population consisted of all adult patients who were continuously enrolled with medical and drug benefits from January 1, 2001 to December 31, 2002. Continuous medical and drug benefit enrollment permitted an equivalent observation period across the populations for medical and prescribed drug analyses. A cumulative monthly enrollment gap of up to 60 days was allowed. Patients could change their health plan during this period.

Most of the high-prevalent IMIDs in dermatology, gastroenterology, neurology, pulmonology, and rheumatology that have an abnormal T-cell response were included in this study (Appendix A)[1,2,13]. Patients who had at least one medical service claim with an appropriate International Classification of Diseases, 9th revision (ICD-9) code within any available diagnostic field were selected as part of the IMID cohort of interest. The ICD-9 codes used for IMID case identification (Appendix A) were based on literature, expert opinion, and exploratory analyses. To explore clinical patterns in psoriasis, and possibly contribute useful epidemiologic information, psoriasis patients were divided into two subgroups based on the presence of an arthritic component to the disease; psoriasis without arthritis (PsO) and psoriatic arthritis (PsA)[14,15].

Non-IMID patients were randomly selected as controls from the initial study population and frequency matched at a 3:1 ratio to each IMID patient based on age, gender, US Census region, and pre-baseline medical insurance coverage. Frequency matching of controls without replacement was carried out for the overall IMID population to ensure that no individual control patient was selected more than once. However, the same control patient could be included in the control groups constructed for more than one particular IMID. Additionally, patients with a particular IMID could be included in the control group for a different IMID. Given the clinical ambiguity and case recognition challenges for some IMIDs, a sensitivity analysis was conducted to examine the effect of a more stringent case identification criterion. In this analysis, data for patients with at least one inpatient, or at least two ambulatory, face-to-face treatment events were included, rather than having at least one medical service claim. Ambulatory face-to-face encounters were encounters resulting in outpatient-level medical service claims carrying office visit Current Procedural Terminology codes.

Data sources

Dataset A had an initial population of about six million patients during the observation period (Table 1). These patients were from self-insured employers, many being Fortune 500 firms offering wages and benefits that tend to promote stable workforces. However, employees who left their employers during the observation period, as well as their covered dependents, would be lost from the study. While missing Medicaid patients, dataset A does contain some retired and Medicare populations with medical and drug benefit coverage, which provided limited representation of the elderly.

Dataset B contained about 11 million unique patients for analysis between 2001 and 2002, drawn from health plans with predominantly commercial lines of business. As with dataset A, Medicaid, Medicare, and pension-plan patients with drug as well as medical coverage were few in number. Reflecting the healthcare plan source of the data, patients could disappear from the record in any given year if they changed health plans or lost healthcare benefits. While national in scope, both datasets drew predominantly from the Midwestern and Southern regions of the US (Table 1).

Data analyses

When examining the IMID co-occurrence data, unusually high-risk rates were found for some IMIDs across both datasets, particularly those with musculoskeletal manifestations. Diagnostic confusion, revisions, miscoding, or diagnostic differences across treating physicians could appear as a co-occurrence in the claims-based patient treatment record[4]. Some IMIDs lack established clinical criteria, while others overlap significantly in symptom manifestations. To address this issue, IMIDs were classified by medical experts into subgroups that had disparate organ-system involvement, were physiologically distinct, and presented with contrasting symptoms[4]. These IMIDs or IMID groupings were thought to have a low likelihood of diagnostic confusion with one another, even in primary care settings, and comprised the following (Figure 1):

- Musculoskeletal (MSK), including ankylosing spondylitis (AS), psoriatic arthritis (PsA), and rheumatoid arthritis (RA).
- Gastroenterological (inflammatory bowel disease[12] [IBD]), including Crohn's disease (CD) and ulcerative colitis (UC).
- Multiple sclerosis (MS).
- Scleroderma (SSc or systemic sclerosis).
- Sarcoidosis (Sar).

Comorbidities were selected based on the expected relationship with IMIDs, safety events known to be associated with IMID treatment, and the public health importance of the comorbidity either in prevalence or impact (Appendix A)[5–12,16–19]. As with IMID case identification, the ICD-9 codes used to define comorbidities were based on existing literature, expert opinion, and exploratory data analyses conducted in both datasets (Appendix A). Three conditions that were not expected to occur with increased frequency in the IMID population, i.e., head trauma, menopause, and migraine, were also included as comorbidity controls. Occurrence of these three conditions was thought to be independent of an immune-based, inflammatory mechanism of action.

To estimate prevalence rates, all patients who were continuously enrolled with medical benefits were included in the denominator, regardless of whether they sought health services during the period. Co-occurrence and comorbidity rates were compared to matched controls using the prevalence relative risk (RR) statistic[20,21]. The RR 95% confidence intervals (CIs) were estimated using the Mantel–Haenszel test. Two standard measures of overall chronic disease burden, the chronic disease score (CDS) and Elixhauser comorbidity index (ECI), were used to compare the cohorts. The CDS was calculated with the updated model now referred to as RxRisk[22], while the original method was used for the ECI[23–25]. The CDS is based on prescribed drug claims, while the ECI is derived from ambulatory and hospital medical claims. Both CDS and ECI are subject to coding issues; however, concordance in direction was considered indicative of a reliable trend[26,27]. One programmer (K.F.), familiar with both datasets, wrote computer programs for all study analyses, thus eliminating the influence of programming variation on the results. All analyses were conducted using the SAS system, version 8.2 (SAS Institute, Cary, NC, USA).

Results

Patient inclusion into the final study population was most affected by the requirement for continuous enrollment over the 2-year period. In dataset A, 59.0% of the initial population was lost from the study with this criterion, while in dataset B, 74.6% was lost. Another 7.4% and 1.0% of the initial population were lost in datasets A and B, respectively, by requiring continuous medical and drug benefit coverage, along with at least one medical claim. The most significant effect of these entry criteria was on the regional distribution of the study population when compared to the initial population. There was also an increase in

Table 1. *Study population selection and demographics**

	Dataset A				Dataset B			
	N (%)	Female, *n** (%)	Age*	Region* (%)	*N* (%)	Female, *n** (%)	Age*	Region* (%)
Mx enrolled adults Jan 1, 2001–Dec 31, 2002	6 389 321 (100)	3 432 967 (53.7)	48.4 (18.1) [48]	E: 12.8 MW: 31.0 S: 37.4 W: 18.6	10 897 605 (100)	5 787 775 (53.1)	40.2 (13.9) [40]	E: 15.2 MW: 30.0 S: 47.4 W: 7.4
Mx continuously-enrolled adults Jan 1, 2001–Dec 31, 2002	2 619 719 (41.0)	2 030 405 (77.5)	51.7 (17.4) [52]	E: 13.4 MW: 33.9 S: 35.9 W: 16.7	2 872 333 (26.4)	1 531 712 (53.3)	43.4 (13.9) [44]	E: 18.0 MW: 45.0 S: 27.3 W: 9.6
Mx and Rx continuously-enrolled adults, ≥ 1 medical claim	2 148 261 (33.6)	1 194 404 (55.6)	50.7 (17.1) [51]	E: 18.3 MW: 31.7 S: 42.7 W: 7.3	2 769 935 (25.4)	1 498 892 (54.1)	43.5 (13.9) [44]	E: 18.2 MW: 45.5 S: 26.8 W: 9.5
IMID population – high sensitivity†	114 430 (4.4)§	72 715 (63.5)	55.9 (16.1) [56]	E: 19.3 MW: 33.6 S: 40.1 W: 6.8	127 667 (4.4)§	81 258 (63.7)	47.6 (13.4) [48]	E: 24.1 MW: 40.2 S: 27.4 W: 8.0
IMID population – high specificity‡	49 213 (1.9)§	32 345 (65.7)	56.5 (15.3) [56]	E: 19.8 MW: 32.8 S: 40.5 W: 6.8	55 726 (1.9)§	36 499 (65.5)	48.6 (12.8) [49]	E: 22.8 MW: 41.4 S: 27.8 W: 7.8

*Based on 2001 data: mean, (SD), [median], and percentages reported as appropriate: percentages for the number of females are based on the row-specific patient *N*. Total percentage may not equal 100% due to rounding and missing data

†**At least** one medical claim over the observation period with an identifying ICD-9 within any diagnostic field

‡**At least one** inpatient claim, or **at least** two ambulatory face-to face encounters over the observation period, with an identifying ICD-9 within any diagnostic field

§Percentage based on a denominator of 2 619 719 adults continuously enrolled from January 1, 2001 through December 31, 2002 for medical benefits in dataset A; 2 872 333 in dataset B

IMID = immune-mediated inflammatory disorder; E = East; MW = Midwest; S = South; W = West; Mx = medical benefits; Rx = prescribed drug benefits

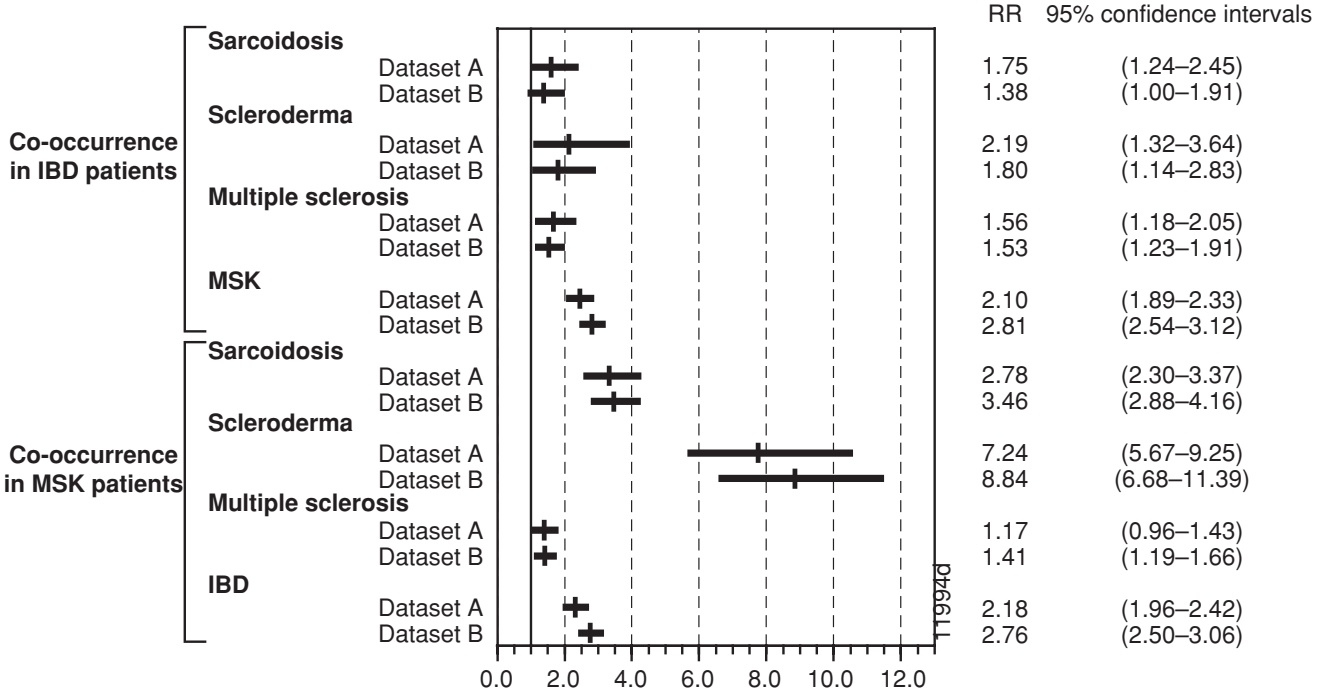

Figure 1. *IMID co-occurrence relative risk across datasets. IMID = immune-mediated inflammatory disorders; RR = relative risk; MSK = musculoskeletal, including ankylosing spondylitis, psoriatic arthritis, and rheumatoid arthritis; IBD = inflammatory bowel disease*

age and female representation when compared to the initial population (Table 1). Dataset A was an older population than dataset B.

Demographics and prevalence

Findings were similar across the datasets. IMID patients had demographic profiles consistent with the literature (Table 2). As a group, patients were older than the overall population and more likely to be female. The all-IMID treatment prevalence rate was about 4.4% of this working US population covered by private or self-administered health insurers. Specific IMID cohorts generally had prevalence rates similar to published reports (Table 2). Key exceptions were sarcoidosis, scleroderma, systemic lupus erythematosus, and giant cell arteritis, which had rates that were notably greater than the sparse literature would indicate. In contrast, the prevalence of ankylosing spondylitis was significantly lower than recent literature.

Co-occurrence

In the final study population, 9.1% and 9.8% of IMID patients had more than one IMID in datasets A and B, respectively. In general, there was a greater prevalence of a second clinically distinct IMID in patients with one IMID compared to controls (Figure 1). Among MSK IMIDs, AS patients had a particularly pronounced risk for inflammatory bowel disease (IBD). The RR for IBD among AS patients was 7.63 (4.95–11.74) and 8.62 (6.18–12.02) in datasets A and B, respectively. This risk relationship was bidirectional; among IBD patients, the RR for AS was 5.81 (3.92–8.61) and 7.75 (5.58–10.78), respectively, across the datasets.

Comorbidities

Consistent with previous reports, IMID patients were at a significantly elevated risk for infectious, renal and liver conditions as well as lymphoma. Infectious risk was the most consistent pattern across the datasets and population models (Figure 2). Lymphoma was the most elevated cancer risk among IMID patients as a whole. IMID patients across both datasets also had higher risks for the comorbidities functioning as controls: head trauma, menopause, and migraine. In dataset A the RRs with 95% CIs were: 1.47 (1.39–1.56), 1.14 (1.12–1.16) and 1.58 (1.52–1.63), respectively, while in dataset B the respective results were 1.60 (1.52–1.69), 1.25 (1.23–1.27) and 1.68 (1.64–1.73) for these conditions.

Healthcare utilization

Medical and drug utilization was higher among IMID patients than controls (Figure 3). Inpatient and outpatient surgical center use was notably greater, as were prescriptions for anxiolytics and anticoagulants. Overall, prescribed drug utilization agreed with the observed medical comorbidity patterns. The overall

Table 2. *IMID population characteristics**

IMID	Dataset A				Dataset B				Published
	n	Age, years (SD)	Female, %	Prevalence, %	*n*	Age, years (SD)	Female, %	Prevalence, %	Prevalence, %
Ankylosing spondylitis	1370	54.4 (13.8)	37.9	0.05	1843	47.3 (12.0)	40.4	0.06	0.20–0.90[1,34]
Rheumatoid arthritis	28 255	59.6 (14.1)	72.5	1.08	28 208	51.9 (12.5)	72.6	0.98	0.50–1.00[1,13,45]
Psoriatic arthritis	2366	54.8 (12.8)	52.2	0.09	3066	49.7 (11.2)	50.9	0.11	0.10–1.00[1,15]
Psoriasis, no Arthritis	21 531	52.9 (16.7)	54.2	0.82	27 492	44.9 (13.5)	53.5	0.96	NA
Psoriasis, overall	23 897	53.1 (16.4)	54.0	0.91	30 558	45.4 (13.4)	53.2	1.06	0.54–2.00[1,14]
Atopic dermatitis	18 720	50.6 (17.4)	63.3	0.71	24 039	43.0 (13.9)	65.7	0.84	1.00–3.00[46]
Giant cell arteritis	1854	70.7 (12.7)	74.3	0.07	1028	62.6 (14.6)	73.3	0.04	0.01[1]
Multiple sclerosis	6826	51.4 (12.6)	75.3	0.26	9131	46.7 (10.3)	76.1	0.32	0.04–0.36[1,13,47]
Pulmonary fibrosis	12 678	66.8 (13.6)	54.3	0.48	8704	55.0 (12.7)	53.5	0.30	0.33[48]
Sarcoidosis	3792	53.4 (11.7)	65.8	0.14	4348	49.1 (10.2)	62.2	0.15	0.04–0.06[1,48]
Scleroderma	1309	57.4 (13.4)	82.9	0.05	1471	50.7 (11.8)	82.9	0.05	0.02[1,13,36]
Sjögren's syndrome	2799	59.8 (13.7)	85.9	0.11	2848	51.7 (10.7)	89.3	0.10	0.01–0.50[13,49]
Systemic lupus erythematosus	6236	52.1 (13.7)	88.1	0.24	7870	47.6 (12.5)	87.2	0.27	0.02–0.12[1,13,32]
Crohn's disease	7401	52.3 (15.6)	57.3	0.29	9267	45.6 (13.1)	57.0	0.32	0.12–0.20[1,50]
Ulcerative colitis	10 104	54.2 (15.6)	54.3	0.39	11 220	47.2 (13.0)	54.0	0.39	0.16–0.25[1,50]
Inflammatory bowel disease†	16 139	53.5 (15.6)	55.5	0.62	18 603	46.7 (13.0)	55.7	0.65	0.25–0.50[1,50]
All IMID§	114 430	55.9 (16.1)	63.5	4.37	127 667	47.6 (13.4)	63.7	4.44	3.00–5.00[1,13]

*Reported as means, (standard deviation) and percentages, as indicated. Two-year prevalence (%) based on a population of 2 619 719 adults continuously enrolled from January 1, 2001 through December 31, 2002 with medical benefits in dataset A; 2 872 333 in dataset B. Due to rounding, psoriasis clinical sub-type (PsA, PsO) prevalence does not equal that of the total psoriasis (Ps) population

†Combining ulcerative colitis and Crohn's disease into inflammatory bowel disease (IBD)

§Unique count since some patients have multiple IMIDs

IMID = immune-mediated inflammatory disorder; NA = not available or known

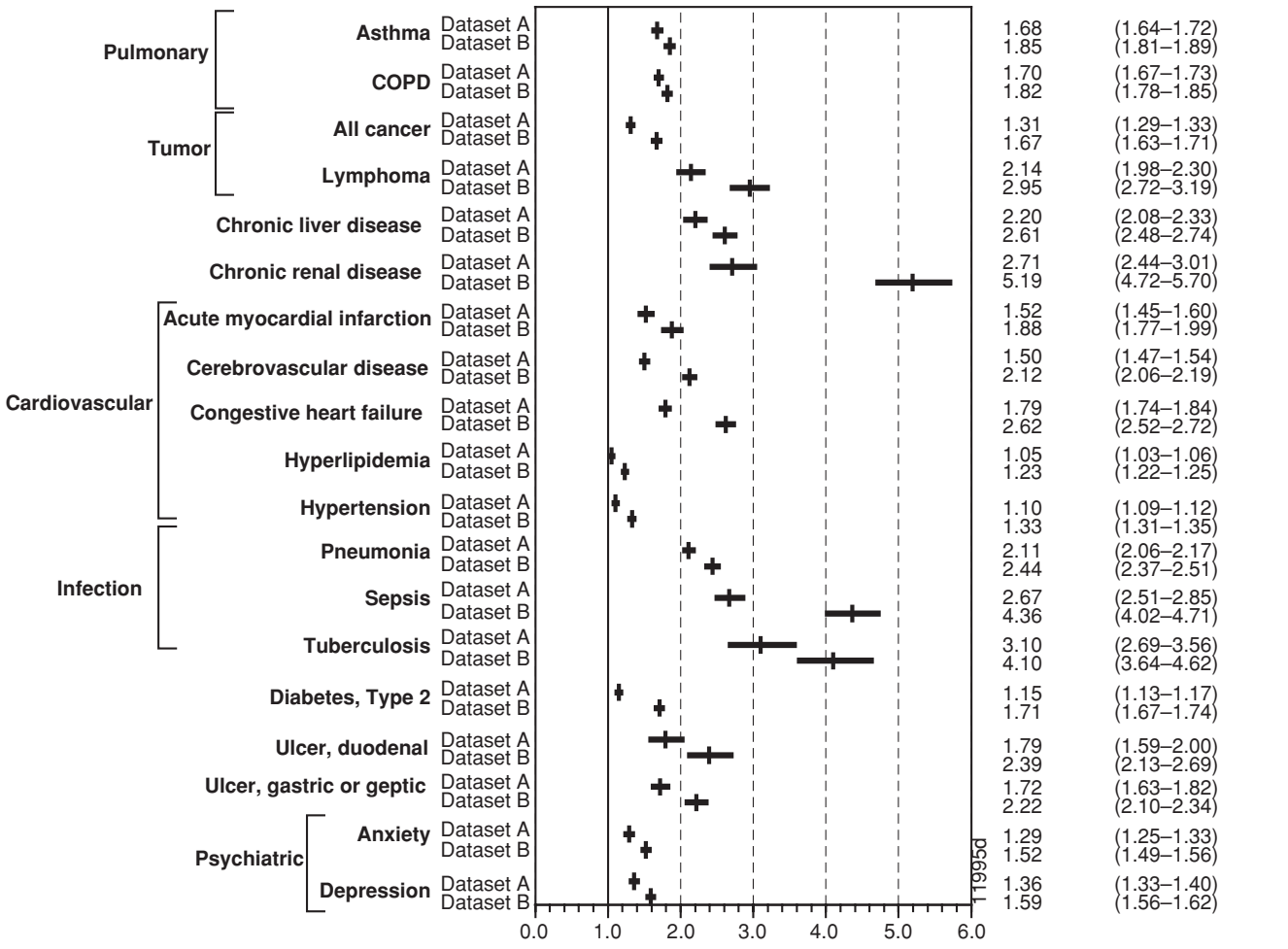

Figure 2. *IMID comorbidity relative risk across datasets*

chronic disease burden among IMID patients was higher than in matched controls, with CDS and ECI values agreeing in direction (Table 3).

Sensitivity analyses

Several analyses were conducted to examine data quality issues, as well as to test the inclusion and exclusion criteria used in the study[28]. The key sensitivity analysis involved using at least one inpatient or at least two ambulatory face-to-face encounters to identify IMID patients, rather than one medical service claim[29–31]. IMID co-occurrence rates were similar in this sensitivity analysis, but higher comorbidity RR rates were observed for more comorbidities. This change in case identification criteria lowered the population prevalence by about 50% (Table 1). It also produced results suggesting a demographically and clinically distinct population, with markedly higher CDS and ECI scores, as well as elevated medical utilization and prescribed drug use. Using at least one inpatient or at least two ambulatory face-to-face encounters for patient identification may select patients with greater disease severity or possibly those more likely to truly have the disease. With IMIDs, both datasets showed that this alternate method identified slightly older and more female patients who had higher chronic disease burden and used notably more healthcare resources.

Discussion

As a group, IMIDs occurred in about 4.4% of the continuously enrolled, predominantly working populations used in this study. The prevalence of specific IMIDs generally matched those published, except for ankylosing spondylitis, sarcoidosis, scleroderma, systemic lupus erythematosus, and giant cell arteritis. However, treatment prevalence for these conditions has not been well established, with prevalence rates changing over time as diagnostic criteria are revised and clinical epidemiology advances[1,32–35]. Compared with the entire enrolled population, IMID patients were 5 years older across both datasets and population models, supporting the 'ageing immune system' thesis[3]. In general, having one IMID increased the risk of having

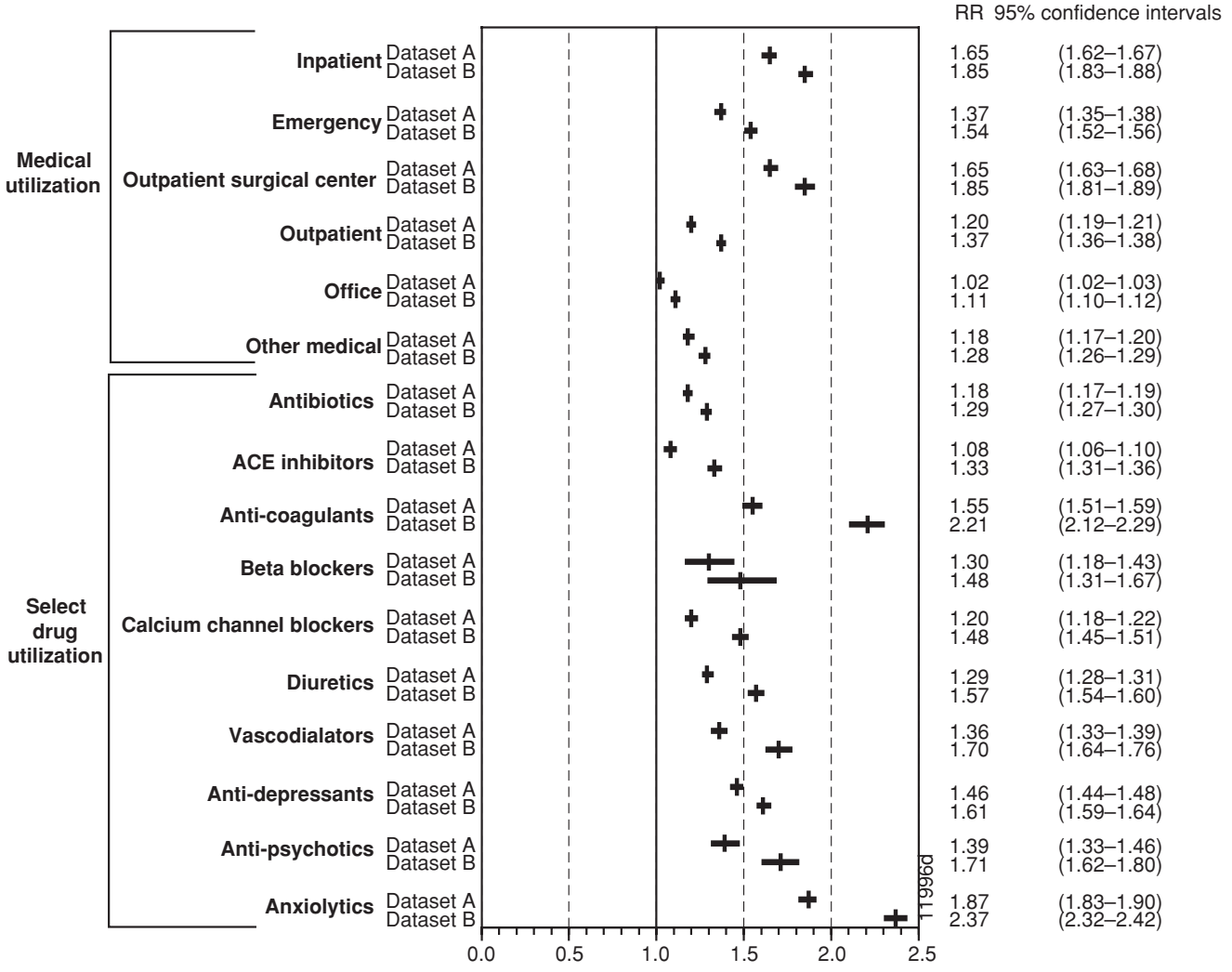

Figure 3. *IMID healthcare utilization relative risk across datasets. Percent of patients with at least one healthcare service use: medical visit or prescription. Other medical services include diagnostic site or home healthcare visits. Prescribed drug categories based on the 'Red Book' (Murray L, Kelly GL (Eds.).* Red Book. *Thomson PDR: Montvale, NJ, 2004) definitions*

another IMID (data not shown). This finding persisted even when certain IMIDs were aggregated into groups with distinct clinical presentations, and when a more rigorous case-identification method was employed to reduce misdiagnosis or diagnostic revision over time.

The IMID population carried a chronic disease burden that was significantly higher than control populations. Patients with the less prevalent IMIDs have especially elevated chronic disease scores. As expected, greater healthcare utilization accompanied this higher chronicity burden. The comorbidity findings across datasets and population models were generally consistent with the literature[6,8,9,16,19,36,37]. Prolonged exposure to systemic agents by some IMID patients may have affected the comorbidity rates because of associated medication toxicity (e.g., hepatotoxicity, nephrotoxicity, carcinogenicity or teratogenicity) or mechanism of action (e.g., immunosuppression). The relationship between illness, treatment, and comorbidity is difficult to untangle within a cross-sectional, observational study design.

Taken together, these data support the possibility of an IMID meta-disease category, based on a shared, systemic, immune-mediated malfunction. Understanding this potential disease category from a pathologic, diagnostic, and treatment perspective is important for effective clinical care. Prospective clinical and epidemiological studies are needed to explore a potentially common pathogenesis across some of the IMIDs, such as those related to the musculoskeletal and gastroenterological systems. Such longitudinal research could clarify primary and secondary IMID onset over the natural disease course. Exploratory genetic associations among IMIDs may also lead to the identification of common mechanisms.

The limits of claims databases for clinical research purposes are well known[28,38–43]. Despite their limits, claims data have been used in the US for over a decade to measure and improve the quality of clinical care in health plans and hospitals, as well as to explore treatment epidemiology and drug safety[5,9,17,22,26–28,37,39,44]. The case-

Table 3. *Population chronic disease burden for IMID and control patients**

IMID and *control*	Dataset A			Dataset B		
	n	ECI, 2001	CDS, 2001	*n*	ECI, 2001	CDS, 2001
Ankylosing spondylitis	1370	1.7 (1.3)	37.8 (36.2)	1843	1.7 (1.5)	28.4 (31.9)
Ankylosing spondylitis control	5480	0.8 (1.1)	24.4 (27.3)	7372	0.7 (1.1)	16.0 (20.8)
Rheumatoid arthritis	28 255	1.9 (1.5)	49.0 (40.0)	28 208	1.9 (1.6)	38.1 (35.4)
Rheumatoid arthritis control	113 016	0.9 (1.2)	29.3 (29.0)	112 832	0.8 (1.2)	20.1 (22.5)
Psoriatic arthritis	2366	1.3 (1.3)	45.9 (38.3)	3066	1.4 (1.5)	37.2 (34.2)
Psoriatic arthritis control	9464	0.8 (1.1)	23.8 (25.8)	12 264	0.8 (1.1)	17.4 (20.4)
Psoriasis, non-arthritic	21 531	0.9 (1.2)	28.1 (32.1)	27 492	0.8 (1.2)	18.4 (22.2)
Psoriasis, non-arthritic control	86 124	0.8 (1.1)	24.7 (27.6)	109 968	0.7 (1.1)	15.9 (19.4)
Psoriasis	23 897	0.9 (1.2)	29.9 (33.2)	30 558	0.8 (1.2)	20.3 (24.4)
Psoriasis control	95 588	0.8 (1.1)	24.7 (27.4)	121 469	0.7 (1.1)	16.1 (19.5)
Atopic dermatitis	18 720	0.9 (1.2)	27.0 (30.1)	24 039	0.8 (1.2)	18.2 (20.8)
Atopic dermatitis control	74 880	0.7 (1.1)	23.6 (26.3)	96 154	0.6 (1.0)	15.7 (18.9)
Giant cell arteritis	1854	2.0 (1.8)	61.5 (47.7)	1028	2.1 (1.9)	45.8 (43.7)
Giant cell arteritis control	7416	1.2 (1.3)	41.4 (34.2)	4112	1.2 (1.4)	29.9 (30.9)
Multiple sclerosis	6826	1.7 (1.4)	33.9 (35.5)	9131	1.8 (1.5)	28.0 (29.0)
Multiple sclerosis control	27 304	0.7 (1.0)	22.2 (24.3)	36 524	0.7 (1.1)	17.0 (19.1)
Pulmonary fibrosis	12 678	2.0 (1.7)	54.0 (44.7)	8704	2.0 (2.0)	37.9 (43.5)
Pulmonary fibrosis control	50 712	1.1 (1.3)	37.7 (33.2)	34 816	0.9 (1.3)	22.3 (25.0)
Sarcoidosis	3792	1.4 (1.5)	36.0 (36.6)	4348	1.5 (1.6)	27.4 (28.8)
Sarcoidosis control	15 168	0.8 (1.1)	22.7 (24.5)	17 392	0.8 (1.1)	17.5 (20.7)
Scleroderma	1309	2.2 (1.6)	47.1 (45.6)	1471	2.3 (1.8)	39.8 (43.7)
Scleroderma control	5236	0.9 (1.2)	26.8 (26.7)	5884	0.8 (1.2)	19.8 (21.1)
Sjögren's syndrome	2799	2.0 (1.5)	43.8 (38.9)	2848	2.0 (1.7)	34.3 (36.3)
Sjögren's syndrome control	11 196	1.0 (1.2)	29.7 (29.3)	11 392	0.9 (1.2)	19.7 (22.3)
Systemic lupus erythematosus	6236	2.1 (1.6)	45.9 (47.8)	7870	2.1 (1.8)	37.7 (41.1)
Systemic lupus erythematosus control	24 944	0.8 (1.1)	23.6 (24.3)	31 480	0.8 (1.1)	19.0 (21.4)
Crohn's disease	7401	1.1 (1.4)	40.9 (36.9)	9267	1.1 (1.5)	35.0 (37.5)
Crohn's disease control	29 604	0.7 (1.1)	23.6 (26.6)	37 068	0.7 (1.1)	16.4 (19.9)
Ulcerative colitis	10 104	1.1 (1.4)	38.7 (37.1)	11 220	1.0 (1.4)	30.5 (32.2)
Ulcerative colitis control	40 416	0.8 (1.1)	25.3 (28.9)	44 880	0.7 (1.1)	16.9 (21.0)
Inflammatory bowel disease	16 139	1.1 (1.4)	39.0 (36.7)	18 603	1.1 (1.4)	31.8 (34.6)
Inflammatory bowel disease control	64 556	0.8 (1.1)	24.5 (26.9)	66 969	0.7 (1.1)	16.7 (20.5)
All IMID	114 430	1.4 (1.5)	38.6 (38.2)	127 667	1.3 (1.5)	27.6 (31.3)
All IMID control	457 711	0.8 (1.1)	26.5 (27.8)	501 590	0.6 (1.0)	15.8 (17.9)

*Reported as means and (standard deviation)
ECI and CDS values are rounded; CDS values are divided by 100. ECI and CDS were similar across both years, with an ECI trend for higher 2002 values
IMID = immune-mediated inflammatory disorder; ECI = Elixhauser comorbidity index; CDS = chronic disease score

identification method for IMIDs and comorbidities depended on the reliability of medical claims coding and administrative processing within naturalistic treatment environments. Furthermore, requiring only one ICD-9 code to identify the study population could result in the inclusion of patients with initial diagnoses ultimately determined to be not present, or the inclusion of patients where the claim was for a 'rule-out' test. A sensitivity analysis was conducted to examine the effect of a more stringent case-identification criterion. In this analysis, data for patients with at least one inpatient or at least two ambulatory encounters were included, rather than having at least one medical service claim. These analyses showed that the overall results of the study were consistent, regardless of the case-identification criterion.

Some IMID co-occurrence and comorbidity associations could have resulted from diagnostic reassessment over time (e.g., psoriatic arthritis may have been initially diagnosed as rheumatoid arthritis). To address diagnostic ambiguity in the study, IMIDs were classified by medical experts into subgroups that had disparate organ-system involvement, were physiologically distinct, and presented with contrasting symptoms. Nonetheless, the higher associations continued to be observed even when co-occurrence for IMIDs with distinct presentations was examined. Both logistic regression and propensity scoring methods were considered as additional confirmatory analyses. The incremental value of these approaches was not clear, given the limited number of clinical and demographic information that could be used in these models. Propensity scoring and logistic regression are best suited for research which requires the adjustment of imbalance in one or more factors between groups (including selection bias issues). The robust frequency

matching of the IMID and control populations in this study adjusted for all available demographic factors, as well as an important pre-baseline health insurance factor associated with the likelihood of chronic disease. Hence the incremental value of further multivariate modeling was minimal. We did not control for comorbidity index scores since comorbidity patterns were one of the outcomes of interest.

A long observation period is required to capture health events in chronic, progressive diseases with low mortality. The 2-year observation period used in this study attempted to capture co-occurrence or comorbidity onset while minimizing the sizeable population attrition known to occur over time in these databases. Of note, the 74.6% reduction in the patient population of dataset B, caused by the continuous enrollment criteria, is an obvious source of concern regarding the internal validity of these findings. A RAND study found continuous enrollment criteria in a healthcare claim dataset tends to select older patients with a greater likelihood for chronic conditions and comorbidities[31]. The selection bias for age was observed in this study, along with gender in dataset A, but the magnitude was not substantial. The external validity of these findings is also constrained by the lack of a Medicaid population in both datasets; sizeable numbers of Medicare patients and retirees were available only in dataset A.

Prevalence-based observation studies mix new and old cases, making exposure to disease or treatment factors variable across the population. While these data are suggestive, observational claims datasets cannot untangle the complex relationships between patient characteristics, treatment patterns, disease severity, and healthcare benefit design. This study found evidence supporting prospective studies to further explore pathogenic linkages among several IMIDs. Longitudinal studies using inception cohorts would illuminate the onset of primary and secondary IMIDs, as well as comorbidity occurrence. Properly designed, these studies could separate disease from therapeutic factors and other background influences. Various biomarkers could be analyzed to better understand physiological processes over the disease course. Finally, the incorporation of patient quality of life questionnaires and experimental clinical response measures in such studies would add much to our knowledge of IMID patient health status and treatment effectiveness over time.

Acknowledgements

Declaration of interest: This study was funded by Centocor, Inc, Malvern, PA, USA. Dataset A was analyzed at Centocor while dataset B was analyzed at United BioSource Corporation, Bethesda, MD, USA under Centocor contract. Quality assurance on all study analyses was conducted by United BioSource Corporation.

The authors appreciate the insightful comments of Alexander Walker, MD, DrPH and William Crown, PhD of i3Magnifi; Charles Ruetsch, PhD and Dennis Revicki, PhD of United BioSource Corporation; Arthur Kavanaugh, MD of the University of California; Leigh Callahan, PhD of the University of North Carolina; James Murray, PhD of Merck and Co.; and those of Daniel Baker, MD, Elliot Barnathan, MD, Michael Elliot, MD, and Gregory Keenan, MD of Centocor, Inc. The medical writing assistance of James Barrett and Mary Whitman, PhD of Centocor, Inc., also deserves mention.

The IMID study group

The IMID Study Group, listed alphabetically, comprises:

Mohan Bala, PhD, Centocor, Inc, Malvern, PA; Russell Cohen, MD, University of Chicago Hospital, Chicago, IL; Kathy Fraeman, SM, AK Analytic Programming, Olney, MD; Mark Genovese, MD, Stanford University, Palo Alto, CA; Cynthia Guzzo, MD, Centocor, Inc, Malvern, PA; Monica Hackett, MHS, Centocor, Inc, Malvern, PA; Chenglong Han, PhD, Centocor, Inc, Malvern, PA; Alexa Kimball, MD, MPH, Massachusetts General and Brigham and Women's Hospitals, Boston, MA; Paul J. Nietert, PhD, Medical University of South Carolina, Charleston, SC; Clark Paramore, MSPH, United BioSource Corp, Bethesda, MD; Kevin Renahan, MBA, Centocor, Inc, Malvern, PA; Don Robinson Jr., Centocor, Inc, Malvern, John Wong, MD, Tufts-New England Medical Center, Boston MA; Newman Yielding, MD, Centocor, Inc, Malvern, PA; Ning Zhao, PhD, Centocor, Inc, Malvern, PA.

References

1. Williams JP, Meyers JA. Immune-mediated inflammatory disorders (I.M.I.Ds): the economic and clinical costs. Am J Manag Care 2002;8:S664-81
2. Davidson A, Diamond B. Autoimmune diseases. N Engl J Med 2001;345:340-50
3. Boren E, Gershwin ME. Inflamm-aging: autoimmunity and the immune-risk phenotype. Autoimmun Rev 2004;3:401-6
4. Sloka S. Observations on recent studies showing increased co-occurrence of autoimmune diseases. J Autoimmun 2002;18: 251-7
5. Bernstein CN, Blanchard JF, Rawsthrone RN, et al. The prevalence of extraintestinal diseases in inflammatory bowel disease: A population based study. Am J Gastroenterol 2001;96:1116-22
6. Kaplan MJ, Ike RW. The liver is a common non-exocrine target in primary Sjögren's syndrome: a retrospective review. BMC Gastroenterol 2002;2:21-30

7. Cervera R, Abarca-Costalago M, Abramovicz D, et al. Lessons from the 'Euro-lupus cohort'. Ann Med Interne 2002;153: 530-6
8. Edwards LJ, Constantinescu CS. A prospective study of conditions associated with multiple sclerosis in a cohort of 658 consecutive outpatients attending a multiple sclerosis clinic. Mult Scler 2004;10:575-81
9. Simpson CR, Anderson WJ, Helms PJ, et al. Coincidence of immune-mediated diseases driven by Th1 and Th2 subsets suggests a common aetiology. A population-based study using computerized general practice data. Clin Exp Allergy 2002;32:37-42
10. Salvarani C, Vlachonikolis IG, van der Heijde DM, et al. Musculoskeletal manifestations in a population-based cohort of inflammatory bowel disease patients. Scand J Gastroenterol 2001;36:1307-13
11. Palm O, Moum B, Ongre A, et al. Prevalence of ankylosing spondylitis and other spondylarthropathies among patients with inflammatory bowel disease: A population study (the IBSEN study). J Rheumatol 2002;29:511-5
12. Martins NB, Peppercorn MA. Inflammatory bowel disease. Am J Manag Care 2004;10:544-52
13. Cooper GS, Stroehla BC. The epidemiology of autoimmune diseases. Autoimmun Rev 2003;2:119-25
14. Langley RGB, Krueger GG, Griffiths CEM. Psoriasis: epidemiology, clinical features, and quality of life. Ann Rheum Dis 2005;64:ii18-23
15. Gladman DD, Antoni C, Mease P, et al. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis 2005;64:ii14-17
16. Margolis D, Bilker W, Hennessy S, et al. The risk of malignancy associated with psoriasis. Arch Dermatol 2001;137:778-83
17. Feltelius N, Ekbom A, Blomqvist P. Cancer incidence among patients with ankylosing spondylitis in Sweden 1965–95: a population based cohort study. Ann Rheum Dis 2003;62: 1185-8
18. Lewis JD, Bilker WB, Brensinger C, et al. Inflammatory bowel disease is not associated with an increased risk of lymphoma. Gastroenterology 2001;121:1080-7
19. Cucino C, Sonnenberg A. The comorbid occurrence of other diagnoses in patients with ulcerative colitis and Crohn's disease. Am J Gastroenterol 2001;96:2107-12
20. Zhang JMB, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690-1
21. Woodward M. Epidemiology: study design and data analysis (1st edn). Boca Raton, Florida: Chapman and Hall, 1999. pp.113-7
22. Fishman PA, Goodman MJ, Hornbrook MC, et al. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. Med Care 2003;41:84-99
23. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. Med Care 1998;36:8-27
24. Stukenborg GJ, Wagner DP, Connors AF Jr. Comparison of the performance of two comorbitity measures, with and without information from prior hospitalizations. Med Care 2001;39:727-39
25. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Med Care 2004;42:355-60
26. Parker JP, McCombs JS, Graddy EA. Can pharmacy data improve prediction of hospital outcomes? Comparisons with a diagnosis-based comorbidity measure. Med Care 2003;41:407-19
27. Schneeweiss S, Seeger JD, Maclure M, et al. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol 2001;154:854-64
28. Hennessy S, Bilker WB, Weber A, et al. Descriptive analyses of the integrity of a US Medicaid claims database. Pharmacoepidemiol Drug Saf 2003;12:103-11
29. Zhang JX, Rathouz PJ, Chin MH. Comorbidity and the concentration of healthcare expenditures in older patients with heart failure. J Am Geriatr Soc 2003;51:476-82
30. Gupta G, Gelfand JM, Lewis JD. Increased risk for demyelinating disease in patients with inflammatory bowel disease. Gastroenterology 2005;129:819-26
31. Peterson E, Van Vorst K, Wickstrom S, et al. In: McGlynn EA, editor. Exploring issues in manage care: six illustrative case studies. RAND Draft Report DRU-1560-HGl 1997. pp.11,16-20, Appendix 1A
32. Wallace DJ, Hahn BH, editor. Dubois' Lupus Erythematosus (6th edn). Philadelphia: Lippincott, Williams and Wilkins, 2002. pp.66-7
33. Newman LS, Rose CS, Maier LA. Sarcoidosis. New Engl J Med 1997;336:1224-34
34. Sieper J, Braun J, Rudwalait M, et al. Ankylosing spondylitis: an overview. Ann Rheum Dis 2002:61(Suppl III):iii8-18
35. Mayes MD. Scleroderma epidemiology. Rheum Dis Clin North Am 2003;29:239-54
36. Chatterjee S, Dombi GW, Severson RK, et al. Risk of malignancy in scleroderma. A population-based cohort study. Arthritis Rheum 2005;52:2415-24
37. Bernstein CN, Wajda A, Blanchard JF. The clustering of other chronic inflammatory diseases in inflammatory bowel disease: a population-based study. Gastroenterology 2005;129:827-36
38. Kolodner K, Lipton RB, Lafata JE, et al. Pharmacy and medical claims data identified migraine sufferers with high specificity but modest sensitivity. J Clin Epidemiol 2004;57:962-72
39. Kiyota Y, Schneeweiss S, Glynn RJ, et al. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. Am Heart J 2004;148:99-104
40. Goff DC, Pandey DK, Chan FA, et al. Congestive heart failure in the United States: is there more than meets the I(CD Code)? The Corpus Christi heart project. Arch Intern Med 2000;160:197-202
41. Birnbaum HG, Cremieux PY, Greenberg PE, et al. Using health-care claims data for outcomes research and pharmacoeconomic analyses. Pharmacoeconomics 1999;16:1-8
42. Strom BL. Data validity issues in using claims data. Pharmaco-epidemiol Drug Saf 2001;10:389-92
43. Platt R, Kleinman K, Thompson K, et al. Using automated health plan data to assess infection risk from coronary artery bypass surgery. Emerg Infect Dis 2002;8:1433-41
44. Hennessy S, Bilker WB, Knauss JS, et al. Comparative cardiac safety of low-dose thioridazine and low-dose haloperidol. Br J Clin Pharmacol 2004;58:81-7
45. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27:269-81
46. Schultz-Larsen F, Hanifin JM. Epidemiology of atopic dermatitis. Immunol Allergy Clin North Am 2002:22:1-24
47. Pope GC, Urato CJ, Kulas ED, et al. Prevalence, expenditures, utilization, and payment for persons with MS in insured populations. Neurology 2002;58:37-43
48. Coultas DB, Zumwalt RE, Black WC, et al. The epidemiology of interstitial lung diseases. Am J Respir Crit Care Med 1994;150:967-72
49. Thomas E, Hay EM, Hajeer A, et al. Sjögren's syndrome: a community-based study of the prevalence and impact. Br J Rheumatol 1998;37:1069-76
50. Loftus EV Jr. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence and environmental influences. Gastroenterology 2004;126:1504-17

CrossRef links are available in the online published version of this paper:
http://www.cmrojournal.com
Paper CMRO-3328_4, *Accepted for publication*: 16 March 2006
Published Online: 19 April 2006
doi:10.1185/030079906X104641

Appendix A

ICD-9 codes for identifying IMIDs and comorbidities

IMIDs	ICD-9 codes
Ankylosing spondylitis*	720.0
Atopic dermatitis	691.x
Crohn's disease	555.x
Giant cell arteritis	446.5
Inflammatory bowel disease*	555.x, 556.x
Multiple sclerosis*	340.x
Psoriasis	696.x
non-arthritic	696.1–8
Psoriatic arthritis*	696.0
Pulmonary fibrosis	515.x
Rheumatoid arthritis*	714.x
Sarcoidosis*	135.x
Scleroderma*	710.1
Sjögren's syndrome	710.2
Systemic lupus erythematosus	710.0
Ulcerative colitis	556.x
Comorbidities	
Asthma	493.x
Chronic obstructive pulmonary disease	490.x–492.x, 494.x, 496.x
All cancers	140.x–208.x, 172.x, 173.x
lymphoma	200.x–202.x
skin	172.x, 173.x
Chronic liver disease and cirrhosis	571.x, 572.8
Liver fibrosis	571.2, 571.5, 571.6
Chronic renal disease	582.x, 583.x, 585.0, 586.0
Acute myocardial infarction	410.x
Cerebrovascular disease	430.x–438.x
Congestive heart failure	428.x
Hyperlipidemia	272.0–4
Hypertension	401.x
Peripheral vascular disease	440.x, 441.x, 443.x, 447.1, 557.1, 557.9, V43.4
atherosclerosis	440.x
Abscess and cellulitis	006.3–5, 324.x, 513.x, 528.0–5, 530.1, 566.x, 572.x, 597.x, 616.3–4, 681.x, 682.x, 730.x
Hepatitis, viral A, B, and C	070.x
Joint infection	711.x
Otitis	380.x–382.x
Pneumonia	480.x–486.x, 516.8–9
Sepsis	038.x, 785.52, 995.90–2
Sinusitis	461.x, 473.x
Tuberculosis	10.x–18.x
Diabetes mellitus, type 2	250.x0, 250.x2
Grave's disease	242.0x
Hemochromatosis	275.0
Ulcer, duodenal	532.x
Ulcer; peptic or gastric	531.x, 533.x
Amnesia	437.7, 780.93
Anxiety disorder	300.0x, 300.21–23, 300.3
Bipolar disorder	296.0x–1x, 296.4x–9x
Delirium	293.x
Dementia	290.x
Depression	296.2, 296.3, 298.0, 300.4, 309.1, 311.x
Schizophrenia	295.x
Other psychosis	291.xx, 292.xx, 294.xx, 297.xx, 298.xx
Head trauma	850.x–854.x, 959.01
Menopause	627.0–2, 627.8–9, 716.30
Migraine	346.xx

*IMIDs considered clinically distinct from each other, with ankylosing spondylitis, psoriatic arthritis, and rheumatoid arthritis combined as musculoskeletal IMIDs (MSK); Crohn's disease and ulcerative colitis combined as gastroenterological IMIDs (inflammatory bowel disease)

IMID = immune-mediated inflammatory disorder; ICD-9 = *International Classification of Diseases*, 9th revision