

RESEARCH ARTICLE

Open Access

Derivation and preliminary validation of an administrative claims-based algorithm for the effectiveness of medications for rheumatoid arthritis

Jeffrey R Curtis^{1*}, John W Baddley^{1,2}, Shuo Yang¹, Nivedita Patkar¹, Lang Chen¹, Elizabeth Delzell¹, Ted R Mikuls^{3,4}, Kenneth G Saag¹, Jasvinder Singh^{1,2}, Monika Safford¹ and Grant W Cannon^{5,6}

Abstract

Introduction: Administrative claims data have not commonly been used to study the clinical effectiveness of medications for rheumatoid arthritis (RA) because of the lack of a validated algorithm for this outcome. We created and tested a claims-based algorithm to serve as a proxy for the clinical effectiveness of RA medications.

Methods: We linked Veterans Health Administration (VHA) medical and pharmacy claims for RA patients participating in the longitudinal Department of Veterans Affairs (VA) RA registry (VARA). Among individuals for whom treatment with a new biologic agent or nonbiologic disease-modifying agent in rheumatic disease (DMARD) was being initiated and with registry follow-up at 1 year, VARA and administrative data were used to create a gold standard for the claims-based effectiveness algorithm. The gold standard outcome was low disease activity (LDA) (Disease Activity Score using 28 joint counts (DAS28) ≤ 3.2) or improvement in DAS28 by > 1.2 units at 12 ± 2 months, with high adherence to therapy. The claims-based effectiveness algorithm incorporated biologic dose escalation or switching, addition of new disease-modifying agents, increase in oral glucocorticoid use and dose as well as parenteral glucocorticoid injections.

Results: Among 1,397 patients, we identified 305 eligible biologic or DMARD treatment episodes in 269 unique individuals. The patients were primarily men (94%) with a mean (\pm SD) age of 62 ± 10 years. At 1 year, 27% of treatment episodes achieved the effectiveness gold standard. The performance characteristics of the effectiveness algorithm were as follows: positive predictive value, 76% (95% confidence interval (95% CI) = 71% to 81%); negative predictive value, 90% (95% CI = 88% to 92%); sensitivity, 72% (95% CI = 67% to 77%); and specificity, 91% (95% CI = 89% to 93%).

Conclusions: Administrative claims data may be useful in evaluating the effectiveness of medications for RA. Further validation of this effectiveness algorithm will be useful in assessing its generalizability and performance in other populations.

Keywords: rheumatoid arthritis, comparative effectiveness, administrative claims data, biologic

Introduction

Large administrative claims databases are commonly used to evaluate medication safety [1,2]. These data sources have a number of advantages, including large size, widespread availability, comprehensiveness and high generalizability to the population being studied.

These databases typically capture medical diagnoses, procedures, drug utilization, hospitalizations, costs and mortality. The diagnostic and procedural codes are submitted by healthcare providers in the course of clinical care and can be used alone or combined into a more complex algorithm to identify conditions of interest to researchers [3,4]. Algorithms are available to identify a number of safety-related conditions, including hospital-associated infections, myocardial infarction, stroke, gastrointestinal perforation, gastrointestinal bleeding and

* Correspondence: jcurtis@uab.edu

¹Department of Medicine, University of Alabama, 510 20th Street South, FOT 805D, Birmingham, AL 35294, USA

Full list of author information is available at the end of the article

fractures [5-14]. In validation studies, most of these algorithms have been shown to have high validity compared to a gold standard of medical record review.

Several studies have also confirmed the validity of various coding algorithms to identify arthritis-specific diagnoses and procedures in different medical settings [15-20]. However, the use of administrative data to study the clinical effectiveness of medications for inflammatory arthritis, such as rheumatoid arthritis (RA), has been limited by the lack of a validated algorithm to serve as a proxy for clinical improvement in RA disease activity. Our objective was to derive and test a claims-based algorithm to serve as a proxy for the effectiveness of medications for RA patients.

Materials and methods

Eligible patient population

After obtaining Institutional Review Board approval, we used data from a cohort of patients diagnosed with RA by a rheumatologist on the basis of the American College of Rheumatology 1987 criteria [21]. These patients were participants in the longitudinal Department of Veterans Affairs (VA) RA registry (VARA), which has been described elsewhere [22]. All VARA participants provided their written informed consent. VARA contains demographic, clinical and RA-specific information, including the Disease Activity Score using 28 joint counts (DAS28), as assessed by physicians using the DAS28 [23] and the Clinical Disease Activity Index (CDAI) [24], as well as a biorepository with banked DNA, serum and plasma. VARA data have been collected by rheumatologists at 11 VHA facilities throughout the United States since 2003. We linked VARA participants to the Veterans Health Administration's Medical SAS Datasets present in the VHA administrative databases from 2002 to 2010 to obtain medical and pharmacy claims.

Among VARA enrollees, we used claims data to identify eligible individuals in whom a biologic agent had been initiated. Biologics of interest included abatacept, adalimumab, etanercept, infliximab and rituximab. We defined "initiation" as no prior use of that biologic agent during the past 6 months. Eligible participants must have had a baseline VARA visit on the same day or within 1 month of biologic initiation. The date of initiation of the biologic (the index date) defined the start of a 1-year "treatment episode." To confirm that patients were receiving medications through the VA system, eligible individuals must have filled at least one prescription (of any duration) for any oral medication during the 6 to 12 months prior to the index date. Participants must also have had a follow-up VARA visit that occurred at 1 year \pm 2 months after the index date. If there was no VARA visit at 1 year, then these treatment

episodes were excluded, as there was no clinical gold standard with which to compare the algorithm's performance. VARA data were used only to capture the DAS28, the CDAI and other clinical characteristics measured at the baseline and outcome VARA visits. All other data used for the analysis were abstracted from the administrative claims data.

To test the performance of the effectiveness algorithm and to see whether it was similar for nonbiologic RA treatments, we performed a separate analysis of RA patients enrolled in VARA who were starting leflunomide (LEF), sulfasalazine (SSZ) or hydroxychloroquine (HCQ) and who also had any prior or current use of methotrexate (MTX). New MTX users were not represented in this analysis, because MTX is typically considered an "anchor" drug for RA patients and generally is continued even if the patient's therapeutic response is suboptimal, in contrast to other RA therapies, where the drugs are typically discontinued if they are not effective. Because of similarities in both the descriptive characteristics of the study populations of biologic and nonbiologic disease-modifying anti-rheumatic drug (DMARD) users and the performance characteristics of the effectiveness algorithm between biologic and DMARD treatment episodes, the data are shown throughout for the biologic users as a unique group and for a combined group of new biologic and nonbiologic DMARD users.

The clinical effectiveness outcome and the effectiveness algorithm

The gold standard for effectiveness was measured at the 1 year VARA visit following the index visit and was defined as DAS28 \leq 3.2 units (low disease activity (LDA)) or improvement in DAS28 $>$ 1.2 units [25,26]. The gold standard also required that the patient have high adherence to biologic treatment (for example, medication possession ratio for oral or injectable biologic therapy \geq 80%) (see Table 1 for further details). The purpose of the adherence requirement was to maximize confidence that observed changes in disease activity were more likely attributable to the treatment started on the index date rather than to natural variations in disease activity, switching to a different RA medication after the index date, or other factors.

The claims-based effectiveness algorithm described in Table 1 incorporates factors (selected *a priori* based upon content knowledge) that were expected to be associated with suboptimal clinical response and would be available within typical administrative claims data sources without laboratory results. The components of the effectiveness algorithm included increase in biologic dose compared to the starting dose; switch to a different biologic; addition of a new nonbiologic DMARD, including MTX, SSZ, LEF and HCQ; initiation of chronic

Table 1 Components of the effectiveness algorithm, assessed between the index date and the outcome visit date approximately one year later

Criteria*	Description and implementation
High adherence to index drug (required)	For etanercept, adalimumab and oral medications, must be $\geq 80\%$ adherent to therapy, calculated as a medication possession ratio [38] For infliximab, must have received at least the number of infusions expected between the index and outcome visit dates to conform to a schedule of 0, 2, 6 and 14 weeks and every 8 weeks thereafter For abatacept, must have received the number of infusions expected between the index and outcome visit dates to conform to a schedule of once-monthly dosing; missing one infusion is permissible For rituximab, criterion is not applicable
Biologic switch or add (prohibited)	Between the index and outcome visit dates, patient cannot initiate therapy with a new biologic agent
Addition of a new nonbiologic DMARD (prohibited)	Between the index and outcome visit dates, patient cannot initiate therapy with a new nonbiologic DMARD (methotrexate, sulfasalazine, leflunomide or hydroxychloroquine) that they were not already taking during the 6 months prior to the index date
Increase in biologic dose or frequency (prohibited)	For etanercept and adalimumab, dose escalation of etanercept to 50 mg twice weekly or adalimumab 40 mg once weekly is prohibited For infliximab, difference between ending and starting dose, with each dose rounded up to the nearest 100 mg cannot be ≥ 100 mg. The number of infusions must be within 120% of the number expected assuming a 0-, 2-, or 6-week load and an 8-week infusion schedule For abatacept, difference between ending and starting dose cannot be ≥ 100 mg For rituximab, criterion is not applicable
More than one glucocorticoid joint injection (prohibited)	Cannot receive glucocorticoid injection† on more than one unique calendar day between the index date + 90 days and the outcome visit date, inclusive
Increase in dose of oral glucocorticoid (prohibited)	For patients who received no prescriptions for oral glucocorticoids during the 6 months prior to the index date, cannot have received more than 30 days of oral glucocorticoids between the index date + 90 days and the outcome visit date, inclusive For patients who received prescriptions for oral glucocorticoids in the 6 months prior to the index date, the cumulative glucocorticoid dose in the 6 months prior to the outcome visit date must be similar (that is, within 120%) to the cumulative dose in the 6 months prior to the index visit date

DMARD: disease-modifying agent in rheumatic disease. †Glucocorticoid injection CPT codes: 20600, 20605, 20610. *All criteria must be satisfied to have met the effectiveness algorithm.

glucocorticoids (for those with no oral glucocorticoid prescriptions during the 6 months prior to the index date); increase in glucocorticoid dose during months 6 to 12 (for those who received any oral glucocorticoid prescriptions in the 6 months prior to the index date); and more than one parenteral or intraarticular injection on unique days after the patient had been receiving the new treatment for more than 3 months. Each of these factors was included in the algorithm as a series of dichotomous conditions that were either satisfied or not. Patients must have satisfied all conditions to have met the effectiveness rule.

Statistical analysis and additional sensitivity analyses

We calculated the performance characteristics, including positive predictive value (PPV), negative predictive value (NPV), sensitivity (Se) and specificity (Sp), to compare the effectiveness algorithm to the effectiveness gold standard, and we used the binomial distribution to calculate 95% confidence intervals. Because patients were allowed to contribute multiple treatment episodes, we performed an additional analysis where all patients were permitted to contribute only one treatment episode each. This approach was felt to be more conservative than alternate strategies, such as using generalized

estimating equations that account for the within-person variance by widening the confidence intervals of the PPV, NPV, Se and Sp, but leave the point estimates unchanged.

For all treatment episodes where there was discordance between the administrative data-based effectiveness rule and the gold standard for clinical effectiveness, we abstracted additional data from the medical records using a structured case report form developed to descriptively inform the reason for discordance.

Although not explicitly part of the effectiveness rule, we also identified comorbidities (posttraumatic stress disorder, low-back pain, fibromyalgia, hepatitis C and depression) that were hypothesized to be associated with worse patient global scores independently of RA disease activity. As part of a sensitivity analysis, we restricted the cohort to patients without any of these ICD-9 codes. As part of two additional sensitivity analyses, we dropped the requirement that patients have a baseline VARA visit. This allowed for inclusion of a modest number of additional VARA treatment episodes where only an outcome VARA visit (but not a baseline VARA visit) was available. In these sensitivity analyses, clinical effectiveness was defined by low disease activity as (1) DAS28 ≤ 3.2 with high adherence or (2) CDAI < 11

with high adherence. All analyses were performed using SAS 9.2 software (SAS Institute, Cary, NC, USA).

Results

The characteristics of the VARA participants were measured at the start of each treatment episode. Because the characteristics of VARA patients at the start of non-biologic DMARD treatment episodes were similar to those of the biologic treatment episodes, these data were pooled and are shown in Table 2 as biologic treatment episodes (left column) and a combined group of biologic or nonbiologic DMARD treatment episodes (right column). As shown, and consistent with expectations for this RA population of US veterans [27], 94% were male, the majority were Caucasian and there was a high prevalence of current or past smoking. The most commonly initiated biologic was adalimumab (38%). For all eligible biologic treatment episodes ($n = 197$), patients had high starting disease activity as evidenced by a mean DAS28 of 5.0, a mean tender joint count of 9.6 and a mean swollen joint count of 7.9. After combining the biologic

treatment episodes with the DMARD treatment episodes ($n = 305$ total), the descriptive characteristics of the eligible cohort remained similar (right column in Table 2).

The primary results of the study are shown in Tables 3 and 4. Among patients treated with biologics (Table 3), a total of 28% of treatment episodes were deemed effective based upon the patients' remaining on therapy and achieving either low disease activity (mean DAS28 ≤ 3.2) and/or a 1.2 unit improvement in DAS28. The PPV and NPV of the administrative data-based effectiveness algorithm were 75% and 90%, respectively. The sensitivity of the effectiveness algorithm was 75%, and its specificity was 90%. If patients were restricted to contributing only one treatment episode ($n = 161$ unique patients), the PPV was 76% and the NPV was 91%.

Among the biologic users in Table 3, the most common reasons why patients failed to meet the effectiveness algorithm criteria were suboptimal adherence, discontinuation and/or switching to a different biologic agent ($n = 118$, 60%); glucocorticoid dose increase ($n = 30$, 15%); addition of new nonbiologic DMARDs ($n =$

Table 2 Baseline characteristics of VARA participants at the start of each biologic treatment episode

Characteristics	Biologics only (N = 197)	Biologics or DMARDs* (N = 305)
Patient demographics		
Age, years	60.9 \pm 10.3	62.3 \pm 10.4
Males	185 (94%)	287 (94%)
Race/ethnicity		
Caucasian, non-Hispanic	159 (81%)	248 (81%)
Non-Caucasian, Hispanic	7 (4%)	8 (3%)
Black, non-Hispanic	27 (14%)	45 (15%)
American Indian or Pacific Islander	4 (2%)	4 (1%)
RA drug initiated		
Abatacept	9 (5%)	9 (3%)
Adalimumab	74 (38%)	74 (24%)
Etanercept	60 (31%)	60 (20%)
Infliximab	34 (17%)	34 (11%)
Rituximab	20 (10%)	20 (7%)
Hydroxychloroquine	n/a	63 (21%)
Leflunomide	n/a	20 (7%)
Sulfasalazine	n/a	25 (8%)
RA-related characteristics		
DAS28	5.0 \pm 1.5	4.9 \pm 1.6
CDAI (0-76)	30.2 \pm 16.3	27.5 \pm 15.2
Physician global (0 to 100)	51.0 \pm 22.1	50.3 \pm 22.6
Patient global (0 to 100)	57.4 \pm 25.2	54.8 \pm 24.2
Tender joint count (0 to 28)	9.6 \pm 8.6	8.5 \pm 7.9
Swollen joint count (0 to 28)	7.9 \pm 7.2	7.8 \pm 6.6
MDHAQ (0 to 3)	1.2 \pm 0.6	1.2 \pm 0.6
ESR, mm/hour	27.9 \pm 23.3	29.9 \pm 24.6
CRP, mg/dL	1.9 \pm 2.4	2.1 \pm 2.5

Data are n (%) or means \pm SD. DMARD: disease-modifying agent; RA: rheumatoid arthritis; CDAI: Clinical Disease Activity Index; CRP: C-reactive protein; DAS28: Disease Activity Score in 28 joints; MDHAQ: Multi-Dimensional Health Assessment Questionnaire; ESR: sedimentation rate; n/a: not applicable; SD: standard deviation. *Includes hydroxychloroquine, leflunomide and sulfasalazine.

Table 3 Comparison of effectiveness algorithm versus effectiveness gold standard for biologic users

Met effectiveness algorithm**	Effectiveness gold standard*		Total	PPV (95% CI)	NPV (95% CI)
	Yes	No			
Yes	42	14	56 (28%)	75% (62 to 86)	
No	14	127	141 (72%)		90% (84 to 94)
Total	56 (28%)	141 (72%)	197 (100%)		
	Se 75% (95% CI = 62 to 86)	Sp 90% (95% CI = 84 to 94)			

CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value; Se: sensitivity; Sp: specificity. *DAS28 \leq 3.2 or DAS28 improvement by $>$ 1.2 units and high adherence (for example, \geq 80%) to the biologic started on the index date. **The components of the effectiveness algorithm are shown in Table 1.

23, 12%); biologic agent dose increase ($n = 15$, 8%); glucocorticoid initiation ($n = 10$, 6%); and more than one joint injection ($n = 11$, 6%). The results of the sensitivity analysis that excluded biologic treatment episodes for patients with any of the several comorbidities of interest (33%, $n = 131$ treatment episodes remaining) yielded a slightly higher PPV (81%) and a similar NPV (89%) compared to the main analysis.

The performance characteristics of the combined cohort that included both biologic and nonbiologic treatment episodes are shown in Table 4 and were generally quite similar to the PPV and NPV shown for the biologic treatment episodes in Table 3. Further details obtained from medical record review were available for the patients in the off-diagonal (discordant) cells given in Table 4 and are shown in Table 5. For the 19 treatment episodes where the effectiveness algorithm criteria were satisfied but the gold standard criteria were not, the most common reasons found were either that an inadequate clinical response was recognized but medication changes were precluded because of new or worsened comorbidities, or the physician and/or the patient was satisfied with the level of disease activity, even though the patient did not meet the DAS28 criteria for low disease activity or improvement. For the 23 treatment episodes in which the effectiveness algorithm criteria were not satisfied but the gold standard criteria were, the most common reasons were an increase in the dose of oral glucocorticoids and the addition of new nonbiologic DMARDs.

The extent of bias resulting from misclassification of our algorithm is described in Table 6. After varying a hypothetical response rate as measured by the algorithm from 30% and 60%, the amount of bias compared to the true response rate ranged from 1% to 21%.

The results of the second sensitivity analysis that had no baseline VARA visit (and thus could not include change in disease activity as part of the effectiveness gold standard) but included all patients, regardless of comorbidities, are shown in Additional file 1. Many more treatment episodes were available ($n = 380$ for biologic treatment episodes and $n = 699$ for biologic or DMARD treatment episodes). Approximately 20% of patients achieved the effectiveness gold standard, which in this analysis was low disease activity (DAS28 \leq 3.2). The NPV of the effectiveness algorithm was high (92%), but the PPV was substantially lower (49%). After substituting CDAI $<$ 11 for DAS28 \leq 3.2 as the gold standard for clinical effectiveness in the third sensitivity analysis, the results were nearly identical (data not shown).

Discussion

We developed a novel, administrative data-based clinical effectiveness algorithm for use in future studies as a proxy for the clinical effectiveness of RA medications. In this preliminary assessment of its performance, we showed that it has acceptable sensitivity, specificity, PPV and NPV. Our sensitivity, specificity, PPV and NPV that were in the 75% to 90% range reflect good, although not perfect, performance of our effectiveness algorithm

Table 4 Comparison of effectiveness algorithm versus effectiveness gold standard for biologic and nonbiologic disease-modifying agent in rheumatic disease treatments**

Met effectiveness algorithm*	Effectiveness gold standard***		Total	PPV (95% CI)	NPV (95% CI)
	Yes	No			
Yes	60	19	79 (26%)	76% (71 to 81)	
No	23	203	226 (74%)		90% (88 to 92)
Total	83 (27%)	222 (73%)	305 (100%)		
	Se = 72% (95% CI = 67 to 77)	Sp = 91% (95% CI = 89 to 93)			

CI: confidence interval; DMARD: disease-modifying; NPV: negative predictive value; PPV: positive predictive value; Se: sensitivity; Sp: specificity. *The components of the effectiveness algorithm are given in Table 1. **DMARDs include hydroxychloroquine, leflunomide and sulfasalazine. ***Defined as DAS28 \leq 3.2 or DAS28 improvement by $>$ 1.2 units and high adherence (for example, \geq 80%) to the biologic and/or DMARD started on the index date.

Table 5 Reasons for discordance between the effectiveness algorithm and the effectiveness gold standard

Reasons for discordance	Satisfied effectiveness algorithm, did not meet effectiveness gold standard (false-positives) (n = 19)	Did not satisfy effectiveness algorithm, met effectiveness gold standard (false-negatives) (n = 23)
Presumed reasons for not meeting gold standard, obtained from medical record review		
Biologic change deferred in light of concerns for new/worsened comorbidity	10	-
Clinically stable or improved and patient/physician satisfied, but DAS and DAS change did not meet gold standard effectiveness criteria	4	-
Physician recognized inadequate response, but chose to retreat with rituximab only after 1 year	2	-
Receiving some medications (for example, glucocorticoids) outside of the VHA system	1	-
Biologic change deferred in light of surgery or procedure	1	-
Physician recommended biologic change or dose change, but patient declined	1	-
Noncompliance with nonbiologic RA medications	1	-
Components of the effectiveness algorithm that were not met despite having met the effectiveness gold standard		
Glucocorticoid dose increase or initiation	-	15
Added new DMARD(s)	-	6
Increase in biologic dose and/or frequency	-	2

VHA: Veterans Health Administration; RA: rheumatoid arthritis; DMARD: disease-modifying. Data shown are the number of treatment episodes in the off-diagonal cells given in Table 4. Column totals may sum to > 100% because there may be multiple reasons why patients did not meet the effectiveness gold standard or the effectiveness algorithm. -, criterion is not applicable.

applied to administrative claims data. By way of comparison, the corresponding performance characteristics of administrative data for a number of rheumatology conditions, including diagnoses for RA, spondyloarthropathies, systemic lupus erythematosus, fibromyalgia, osteoarthritis, joint injection and joint replacement procedures [15-20] were similar and ranged from approximately 80% to 95%. Besides a new or worsened comorbidity, the most common reason why patients met the effectiveness algorithm criteria but failed to meet the gold standard criteria was that the physician and patient were satisfied with the level of disease activity, despite not having achieved low disease activity or an improvement in the DAS28 by ≥ 1.2 units. In this

circumstance, providers may feel that the patient is getting at least some benefit from the drug and that the clinical response is adequate to continue its use. It is also possible that quantitative disease activity measures such as the DAS28 may not adequately capture underlying RA disease activity for some patients (for example, those with concomitant fibromyalgia). Moreover, patients may fear that their condition will worsen after switching to a new therapy or may have trepidation regarding new side effects [28], and therefore they may be reluctant to change medications. Further studies are needed to validate the effectiveness algorithm in other data sets and RA patient populations. However, these results are encouraging and suggest that administrative data can be used to estimate medication effectiveness for RA patients.

Table 6 Extent of bias associated with misclassification* of the effectiveness algorithm according to observed response rate

Observed response rate	True response rate**	Bias (observed-true)/true (%)
30%	30%	< 1%
40%	36%	10%
50%	43%	16%
60%	49%	21%

*From Table 4, where the positive predictive value (PPV) was 76% and the negative predictive value (NPV) was 90%. **Computed as True rate = Observed rate \times (PPV + NPV-1)/NPV + 1 [39]. Numbers are rounded to the nearest whole integer but the actual values were used to calculate the bias.

As our gold standard for medication effectiveness, we selected low disease activity (DAS28 ≤ 3.2) or improvement in DAS28 by > 1.2 units. It might be argued that these criteria are not stringent enough, although they are broadly consistent with (albeit not identical to) the European League Against Rheumatoid Arthritis (EULAR) responder definition [26]. Consistent with our focus on the DAS28, results from a preference analysis found that RA disease activity score (also measured using the DAS28) was the most important factor in rheumatologists' decisions to escalate care [29]. The

results from the Consortium of Rheumatology Researchers of North America (CORRONA) registry showed that low disease activity or a DAS28 improvement > 1.2 units was sufficient for the majority of patients to continue treatment with biologic therapy [30]. As part of a sensitivity analysis, we modified our gold standard to require patients to achieve only LDA (DAS28 ≤ 3.2) and did not include patients who achieved only some improvement (change in DAS28 ≥ 1.2) in the absence of LDA. This lowered the PPV, indicating that many patients had clinical improvement but did not achieve LDA. Many of these patients were continued on therapy, suggesting that both the patients and physicians were in many cases satisfied enough with the response. We also note that the DAS28 response rate (approximately 30%) (Table 3) observed for our clinical effectiveness gold standard was relatively low. However, given the comorbidity profile and other characteristics of the RA patients enrolled in VARA [31], response rates are typically lower than those reported in clinical trials of more selectively included RA patients with fewer comorbidities [32].

Another component of our gold standard is that we required that patients have high (that is, $\geq 80\%$) adherence to their medication regimen. We recognize that any threshold for adherence is arbitrary. Requiring $\geq 80\%$ compliance is conventional and has been used when studying other conditions, such as osteoporosis and cardiovascular disease [33-36]. The main purpose of the adherence requirement was to focus on medication effectiveness. Medications that the patient does not continue, whether for reasons of inefficacy, safety, tolerability or something else, are not effective. Adherence has been required in other observational analyses of comparative effectiveness in RA [37]. Also, we wanted to maximize confidence in the patient's disease activity's being attributable to the RA treatment started on the index date rather than on a medication that was later substituted because the previous medication begun on the index date had failed. Finally, the requirement of continued adherence to the RA therapy is consistent with clinical trial methodology in which patients who do not adhere to the study protocol, including continuing to take the medication, are generally excluded from the trial. These patients' outcomes are often imputed as nonresponse, which is the same classification to which they were assigned in our effectiveness algorithm.

Although many of the elements of our effectiveness algorithm are intuitive, a few deserve special mention. The requirement that patients not initiate or escalate the dose of oral glucocorticoids assumes that the dominant prescribing indication for glucocorticoids is RA. For patients who may have another indication for glucocorticoids (for example, chronic obstructive pulmonary

disease, which is very common in VHA patients), this criterion may not perform optimally. As described in Table 5, this issue was the most common reason why patients failed the effectiveness algorithm. Our algorithm might be expected to perform better in other RA populations that have been shown to have a lower prevalence of comorbidities for which systemic glucocorticoids are used [31]. We also limited the number of intraarticular injections allowable to no more than 1 unique day on which the patient received such injections. VA physicians are not directly compensated for these injections and other procedures and therefore are likely to under-report them. For this reason, our effectiveness algorithm may perform better when there is a financial incentive to code these procedures more accurately. We also found certain comorbidities (for example, fibromyalgia and depression) were common, and we hypothesized that they might be associated with high patient global scores even if the patient's RA is under good control. This is not a unique feature of the VARA cohort or our study, but is potentially problematic for the measurement of patient-reported outcomes in all RA studies that include patients with these conditions. Restricting the population to individuals without these comorbidities improved the PPV of our effectiveness algorithm by 6%, but limits our study's generalizability as it excluded one-third of our data.

The strengths of our study include evaluation of a large number of patients participating in a RA registry at 11 VA medical centers. All patients had rheumatologist-confirmed RA and well-characterized measures of RA disease activity. The novel linkage between the registry and the national VHA administrative data made developing and testing of our effectiveness algorithm possible. Additionally, there are strong financial incentives for RA patients to fill their biologic medications within the VHA system, and it is likely that most if not all RA medications were captured in the VHA administrative data. Despite these strengths, we acknowledge the potentially limited generalizability of patterns of care in the VHA system, and the possible dissimilarity in the RA patients who receive treatment in that system, compared to other RA populations. However, sensitivity and specificity, unlike PPV and NPV, should be less dependent on the prevalence in the population, and more reflective of the test itself, thereby decreasing the impact of any unique features of the VA population. Moreover, we might expect that the PPV and NPV of the algorithm might perform better in other RA cohorts, given the higher prevalence of comorbidities in this VARA population compared to other RA cohorts [31]. We also acknowledge that while the effectiveness algorithm, which was based upon factors selected from content knowledge, appeared to perform well and have

good face validity in VARA, further validation in more recently recruited VARA participants who were not included in our sample, and in different RA cohorts where there is a link to administrative data, is needed to confirm our algorithm's robustness. We also recognize that using more empirical approaches to let the data guide optimization of the algorithm would be desirable, but substantially more data would be required for this approach and for validation. Finally, as an additional opportunity to extend the algorithm in the future, we note that our effectiveness outcome was measured at 1 year, and assessing effectiveness at other time points (for example, at 6 and 24 months) is important. Although we expect similar performance of the algorithm at these different time points, this hypothesis remains to be confirmed.

Conclusions

In conclusion, the results of this work provide a preliminary mechanism with which to evaluate the effectiveness of RA medications on the basis of administrative claims and pharmacy data. While clinical disease activity measures remain the gold standard for assessing effectiveness in RA, the many large administrative data sources in the United States and internationally are an as yet untapped resource that might be used to assess effectiveness in large real-world populations of RA patients.

Additional material

Additional file 1: Sensitivity analysis comparing the effectiveness algorithm to an alternate definition of the effectiveness gold standard. Table S1 Sensitivity analysis comparing the effectiveness algorithm to an alternate definition of the effectiveness gold standard for biologic users. Table S2 Sensitivity analysis comparing the effectiveness algorithm to an alternate definition of the effectiveness gold standard for biologic and nonbiologic disease-modifying agent in rheumatic disease treatments.

Abbreviations

CDAI: Clinical Disease Activity Index; CORRONA: Consortium of Rheumatology Researchers of North America; DAS28: Disease Activity Score using 28 joint counts; DMARD: disease-modifying anti-rheumatic drug; DSS: Decision Support System; EULAR: European League Against rheumatism; GEE: generalized estimating equation; LDA: low disease activity; MPR: medication possession ratio; NPV: negative predictive value; PPV: positive predictive value; RA: rheumatoid arthritis; Se: sensitivity; Sp: specificity; VARA: Veterans Affairs Rheumatoid Arthritis; VHA: Veterans Health Administration.

Acknowledgements

We thank Mike Connor and Sheryl Berryman at the Birmingham VA Medical Center for their assistance in working with the Decision Support System data. This work was supported by the Agency for Healthcare Research and Quality (AHRQ) (U18 HS106956). Dr. Curtis receives support from the National Institutes of Health (grant AR-053351) and the AHRQ (grant R01 HS018517). Dr. Cannon receives funding from a VA Health Services Research & Development (HSR&D) grant (SHP 08-172). Dr. Mikuls receives research

support from the VHA (VA Merit). The VARA Registry has received research support from the HSR&D Program of the VHA in addition to unrestricted research funds from Abbott Laboratories and Bristol-Myers Squibb.

Author details

¹Department of Medicine, University of Alabama, 510 20th Street South, FOT 805D, Birmingham, AL 35294, USA. ²Department of Medicine, Birmingham VA Medical Center, 700 19th Street South, Birmingham, AL 35233, USA. ³Division of Rheumatology, Omaha VA Medical Center, 4101 Woolworth Avenue, Omaha, NE 68105, USA. ⁴Division of Rheumatology, University of Nebraska Medical Center, 42nd and Emile, Omaha, NE 68198, USA. ⁵George E Wahlen VA Medical Center, 500 Foothill Drive, Salt Lake City, UT 84148, USA. ⁶Division of Rheumatology, University of Utah, 30 North 1900 East, SOM4B200, Salt Lake City, UT 84132, USA.

Authors' contributions

All authors made substantial contributions to the study's conception and design and to the analysis and interpretation of the data. TRM and GWC handled the acquisition of data. All authors contributed to the manuscript revision process and addressed important intellectual content. All authors read and approved the final manuscript for publication.

Competing interests

JRC has performed research and consulting for Roche, Genentech, UCB, Abbott, Amgen, CORRONA, Centocor and Bristol-Myers Squibb. All other coauthors have nothing to disclose.

Received: 22 March 2011 Revised: 11 May 2011

Accepted: 20 September 2011 Published: 20 September 2011

References

- Schneeweiss S, Avorn J: A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005, **58**:323-337.
- West SL, Strom B, Poole C: **Validity of pharmacoepidemiology drug and diagnosis data.** *Pharmacoepidemiology West Sussex*, UK: John Wiley and Sons; 2000.
- Miller DR, Safford MM, Pogach LM: **Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data.** *Diabetes Care* 2004, **27**(Suppl 2):B10-21.
- Singh JA: **Accuracy of Veterans Affairs databases for diagnoses of chronic diseases.** *Prev Chronic Dis* 2009, **6**:A126.
- Curtis JR, Mudano AS, Solomon DH, Xi J, Melton ME, Saag KG: **Identification and validation of vertebral compression fractures using administrative claims data.** *Med Care* 2009, **47**:69-72.
- Patkar NM, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, Saag KG: **Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients.** *J Clin Epidemiol* 2008.
- Roumie CL, Mitchel E, Gideon PS, Varas-Lorenzo C, Castellsague J, Griffin MR: **Validation of ICD-9 codes with a high positive predictive value for incident strokes resulting in hospitalization using Medicaid health data.** *Pharmacoepidemiol Drug Saf* 2008, **17**:20-26.
- Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF: **Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors.** *Med Care* 2005, **43**:480-485.
- Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A: **Stroke: who's counting what?** *J Rehabil Res Dev* 2001, **38**:281-289.
- Liu L, Reeder B, Shuaib A, Mazagri R: **Validity of Stroke Diagnosis on Hospital Discharge Records in Saskatchewan, Canada: Implications for Stroke Surveillance.** *Cerebrovasc Dis* 1999, **9**:224-230.
- Goldstein LB: **Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: effect of modifier codes.** *Stroke* 1998, **29**:1602-1604.
- Wahl PM, Rodgers K, Schneeweiss S, Gage BF, Butler J, Wilmer C, Nash M, Esper G, Gitlin N, Osborn N, Short LJ, Bohn RL: **Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population.** *Pharmacoepidemiol Drug Saf* 19:596-603.
- Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH: **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.
14. Curtis JR, Chen S, Werther W, John A, Johnson D: Validation of ICD-9-CM to Identify GI Perforation in Administrative Claims Data Among Rheumatoid Arthritis Patients. *Pharmacoepidemiol Drug Saf* 2011, **20**:1150-8.
 15. Singh JA, Holmgren AR, Noorbaloochi S: Accuracy of Veterans Administration databases for a diagnosis of rheumatoid arthritis. *Arthritis Rheum* 2004, **51**:952-957.
 16. Losina E, Barrett J, Baron JA, Katz JN: Accuracy of Medicare claims data for rheumatologic diagnoses in total hip replacement recipients. *J Clin Epidemiol* 2003, **56**:515-519.
 17. MacLean CH, Louie R, Leake B, McCaffrey DF, Paulus HE, Brook RH, Shekelle PG: Quality of care for patients with rheumatoid arthritis. *JAMA* 2000, **284**:984-992.
 18. Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kievit RI, Lindsey SM, Roberts WN, Sheff DM, Spencer RT, Weaver AL, Baron JA: Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. *Arthritis Rheum* 1997, **40**:1594-1600.
 19. Singh JA, Ayub S: Accuracy of VA databases for diagnoses of knee replacement and hip replacement. *Osteoarthritis Cartilage* 2010, **18**:1639-1642.
 20. Singh JA, Holmgren AR, Krug H, Noorbaloochi S: Accuracy of the diagnoses of spondylarthritides in veterans affairs medical center databases. *Arthritis Rheum* 2007, **57**:648-655.
 21. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, et al: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988, **31**:315-324.
 22. Mikuls TR, Kazi S, Cipher D, Hooker R, Kerr GS, Richards JS, Cannon GW: The association of race and ethnicity with disease expression in male US veterans with rheumatoid arthritis. *J Rheumatol* 2007, **34**:1480-1484.
 23. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995, **38**:44-48.
 24. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, Smolen JS: Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005, **7**:R796-806.
 25. van Gestel AM, Haagsma CJ, van Riel PL: Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998, **41**:1845-1850.
 26. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL: Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996, **39**:34-40.
 27. 2001 National Survey of Veterans. [http://www.virec.research.va.gov/DataSourcesName/NationalSurvey/Veterans/2001NationalSurveyofVeterans.htm], Accessed 3/21/2011.
 28. Wolfe F, Michaud K: Resistance of rheumatoid arthritis patients to changing therapy: discordance between disease activity and patients' treatment choices. *Arthritis Rheum* 2007, **56**:2135-2142.
 29. Kievit W, van Hulst L, van Riel P, Fraenkel L: Factors that influence rheumatologists' decisions to escalate care in rheumatoid arthritis: results from a choice-based conjoint analysis. *Arthritis Care Res (Hoboken)* 2010, **62**:842-847.
 30. Zhang J, Shan Y, Reed G, Kremer J, Greenberg J, Baumbartner S, Curtis JR: Thresholds in Disease Activity for Switching Biologics in RA Patients: Experience from a Large United States Cohort. *Arthritis Care Res (Hoboken)* 2011, **63**:1672-9.
 31. Curtis JR, Jain A, Asking J, Bridges SL Jr, Carmona L, Dixon W, Finckh A, Hyrich K, Greenberg JD, Kremer J, Listing J, Michaud K, Mikuls T, Shadick N, Solomon DH, Weinblatt ME, Wolfe F, Zink A: A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid arthritis registries. *Semin Arthritis Rheum* 2010, **40**:2-14, e11.
 32. Greenberg JD, Kishimoto M, Strand V, Cohen SB, Oleginski TP, Harrington T, Kafka SP, Reed G, Kremer JM, Consortium of Rheumatology Researchers of North America I: Tumor necrosis factor antagonist responsiveness in a United States rheumatoid arthritis cohort. *Am J Med* 2008, **121**:532-538.
 33. Siris ES, Selby PL, Saag KG, Borgstrom F, Herings RM, Silverman SL: Impact of osteoporosis treatment adherence on fracture rates in North America and Europe. *Am J Med* 2009, **122**:S3-13.
 34. Brunner R, Dunbar-Jacob J, Leboff MS, Granek I, Bowen D, Snetselaar LG, Shumaker SA, Ockene J, Rosal M, Wactawski-Wende J, Cauley J, Cochrane B, Tinker L, Jackson R, Wang CY, Wu L: Predictors of adherence in the Women's Health Initiative Calcium and Vitamin D Trial. *Behav Med* 2009, **34**:145-155.
 35. Wei L, Fahey T, MacDonald TM: Adherence to statin or aspirin or both in patients with established cardiovascular disease: exploring healthy behaviour vs. drug effects and 10-year follow-up of outcome. *Br J Clin Pharmacol* 2008, **66**:110-116.
 36. Shmussen JN, Chong A, Alter DA: Relationship between adherence to evidence-based pharmacotherapy and long-term mortality after acute myocardial infarction. *JAMA* 2007, **297**:177-186.
 37. Hetland ML, Christensen IJ, Tarp U, Dreyer L, Hansen A, Hansen IT, Kollerup G, Linde L, Lindegaard HM, Poulsen UE, Schlemmer A, Jensen DV, Jensen S, Hostenkamp G, Østergaard M, All Departments of Rheumatology in Denmark: Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. *Arthritis Rheum* 2010, **62**:22-32.
 38. Sikka R, Xia F, Aubert RE: Estimating medication persistency using administrative claims data. *Am J Manag Care* 2005, **11**:449-457.
 39. Brenner H, Gefeller O: Use of the positive predictive value to correct for disease misclassification in epidemiologic studies. *Am J Epidemiol* 1993, **138**:1007-1015.

doi:10.1186/ar3471

Cite this article as: Curtis et al.: Derivation and preliminary validation of an administrative claims-based algorithm for the effectiveness of medications for rheumatoid arthritis. *Arthritis Research & Therapy* 2011 **13**:R155.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

