

Extraction of Rheumatoid Arthritis Disease Activity Measures From Electronic Health Records Using Automated Processing Algorithms

Grant W. Cannon,¹  Jorge Rojas,² Andreas Reimold,³ Ted R. Mikuls,⁴  Debra Bergman,⁴ and Brian C. Sauer¹

Objective. The accurate and efficient collection and documentation of disease activity measures (DAMs) is critical to improve clinical care and outcomes research in rheumatoid arthritis (RA). This study evaluated the performance of an automated process to extract DAMs from medical notes in the electronic health record (EHR).

Methods. An automated text processing system was developed to extract the Disease Activity Score for 28 joints (DAS28) and its clinical and laboratory elements from the Veterans Affairs EHR for patients enrolled in the Veterans Affairs Rheumatoid Arthritis (VARA) registry. After automated text processing derivation, data accuracy was assessed by comparing the automated text processing system and manual extraction with gold standard chart review in a separate validation phase.

Results. In the validation phase, 1569 notes from 596 patients at 3 sites were evaluated, with 75 (6%) notes detected only by automated text processing, 85 (5%) detected only by manual extraction, and 1408 (90%) detected by both methods. The accuracy of automated text processing ranged from 90.7% to 96.7% and the accuracy of manual extraction ranged from 91.3% to 95.0% for the different clinical and laboratory elements. The accuracy of the two methods to calculate the DAS28 was 78.1% for automated text processing and 78.3% for manual extraction.

Conclusion. The automated text processing approach is highly efficient and performed as well as the manual extraction approach. This advance has the potential for significant improvements in the collection, documentation, and extraction of these data to support clinical practice and outcomes research relevant to RA as well as the potential for broader application to other health conditions.

INTRODUCTION

Guidelines proposed by the American College of Rheumatology (ACR) (1) and the European League Against Rheumatism (EULAR) (2) recommend the regular assessment of disease activity measures (DAMs) to direct a treat-to-target strategy for patients with rheumatoid arthritis (RA). Although these recommendations are evidence based, there are significant challenges with the practical implementation of these guidelines (3,4), particularly the systematic collection and documentation of DAMs during clinical practice (5). Issues identified as barriers to guideline implementation and DAM collection include patients' frequent preference to not implement change (3), providers' reluctance to initiate therapy in the context of comorbidity (4,5), providers' perceptions that

disease activity is insufficient to warrant treatment escalation despite elevated DAMs (4,5), and health systems issues, including inclusion of trainees in practice and provider education (3), insufficient time with patients (5), and racial disparities (5).

The Veterans Affairs Rheumatoid Arthritis (VARA) registry is an observational cohort registry that collects longitudinal data on US veterans with RA at 11 Department of Veterans Affairs (VA) medical centers across the United States (6). A key goal for the VARA registry is to collect core clinical data elements to compute DAMs such as the Disease Activity Score for 28 joints (DAS28). As with other groups, the collection of DAMs has been a challenge. One reported reason for poor adherence is the time and resources required to manually extract the core clinical components from the VA Computerized Patient Record System and upload these data

Supported by the Specialty Care Center of Innovation, the Veterans Health Administration, and the US Department of Veterans Affairs Health Services Research and Development Service.

¹Grant W. Cannon, MD, Brian C. Sauer, PhD: George E. Wahlen Department of Veterans Affairs Salt Lake City Health Care System and University of Utah, Salt Lake City; ²Jorge Rojas, MS: George E. Wahlen Department of Veterans Affairs Salt Lake City Health Care System, Salt Lake City, Utah; ³Andreas Reimold, MD: Dallas Department of Veterans Affairs Medical Center and University of Texas Southwestern, Dallas; ⁴Ted R.

Mikuls, MD, MSPH, Debra Bergman, MPH: Department of Veterans Affairs Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Grant W. Cannon, George E. Wahlen Department of Veterans Affairs Salt Lake City Health Care System, 500 Foothill Drive, Salt Lake City, UT 84148. E-mail: grant.cannon@va.gov.

Submitted for publication May 3, 2019; accepted in revised form September 6, 2019.

to the VARA registry software. Because the manual extraction of DAMs is time consuming and subject to human error, we explored the possibility of developing an automated DAM text-extraction process to improve efficiency, reduce human error, motivate collection and documentation of DAM elements, and support development of an automated audit and feedback approach to improve documentation. The goal of this study was to evaluate the performance of an automated DAM text-extraction process that we developed to support the VARA registry that could be leveraged for both research and clinical care.

MATERIALS AND METHODS

Overview. This study contained two phases: derivation (January 1, 2014, to December 31, 2014) and validation (January 1, 2015, to December 31, 2015). During the 12-month derivation phase, results were compared with those from manual extraction to improve performance of the electronic algorithms and improve structured note templates to facilitate data extraction. Disagreements between the automated and manual extraction processes during this phase primarily related to modifications made to electronic health record (EHR) note templates, either intentionally (eg, systematic change in a site template or in how it was applied) or unintentionally (eg, when “copy and paste” or deletions removed components of the template). The extraction algorithms were updated to address systematic deviation from templates and one-off type errors during the 12-month derivation phase. During the 12-month validation phase, the final algorithm developed during the derivation phase was implemented, and no changes or modifications to these algorithms were made during this latter period.

VARA registry and collection of core clinical elements of DAMs. The VARA registry (6–9) enrolls US veterans who fulfill the 1987 ACR classification criteria for RA (10), collects demographic and clinical information at baseline, and subsequently collects DAMs during routine follow-up clinic visits. A VARA registry database captures these baseline and follow-up data with the goal of collecting the DAS28 (11) at each follow-up visit. Three VARA sites (VA medical centers in Dallas, TX; Omaha, NE; and Salt Lake City, UT) participated in this effort to automate the collection of the core clinical components required for DAS28 calculation, which included the tender joint count (TJC) (0–28 joints), the swollen joint count (SJC) (0–28 joints), patient global assessment (PGA) (a 100-mm visual analogue scale), and the Westergren erythrocyte sedimentation rate (ESR). At the study visit, patients were provided a paper form for collection of clinical history items that included a 100-mm PGA visual analogue scale for completion prior to the rheumatologist encounter on which the patient marked their assessment of global disease activity. The provider subsequently measured the response and reported the PGA in the text of the medical record.

For this report, the TJC, the SJC, and PGA were defined as clinical DAS28 elements, and the ESR associated with each of these clinical visits was the laboratory data element collected. All clinic notes and associated ESRs in the laboratory data generated at these sites during the derivation and validation periods were evaluated.

Entry of DAMs by manual extraction into the VARA database. Prior to the derivation phase, each site’s enrollment center used a site-specific EHR note template that included the TJC, the SJC, and PGA as elements collected during rheumatology clinic visits. After clinic visits, site personnel identified registry participants and employed a manual extraction process to enter each clinical DAS28 element and ESR into the VARA database. This data entry process involved either a direct manual entry of each DAS28 element into the database or the copying and pasting of the EHR note text with clinical DAS28 elements into a note-text processor in the VARA database that would populate the clinical DAS28 elements into the VARA database. All ESR values required direct manual entry. The consistency and success of the collection and entry of these clinical and laboratory data into the VARA database was variable, subject to human error, and dependent on the successful implementation of these different processes at each site.

Development of an automated text processing extraction of clinical DAS28 elements. The processes for recording the TJC, the SJC, and PGA were identified for each of the three participating sites. We found that each site used an EHR note template that contained text strings for entry of the TJC, the SJC, and PGA (Supplemental Table S1), which were identified by site-specific note titles that contained variability within and between sites. It was noted that templates changed over time, note titles were not consistently used, and components of templates could vary depending on how providers applied the template during documentation. The automated text processing algorithms contained three primary functions: 1) retrieval of documents using templates, 2) extraction of DAM elements, and 3) storage of DAM elements to a structured data table with patient, document, and visit identifiers.

The automated text processing method was designed to leverage the recently established Corporate Data Warehouse (CDW), which centralizes data from Veterans Information Systems and Technology Architecture (VistA) systems, which contain regional EHR data (12). VA clinic notes are uploaded nightly to the national VA CDW as text integration utility (TIU) documents and contain specific identifiers for the patient, visit, standard note title, site of care, date, and time. We retrieved TIU notes from the CDW for VARA enrollees using unique patient identifiers and notes with rheumatology titles selected for visits that occurred between January 1, 2014, and December 31, 2014. We initially extracted all note titles with “Rheum” and all its variants to ensure high sensi-

tivity. We then selected notes for processing using regular expressions to identify the presence of site-specific templates.

We wrote a Java application, in conjunction with Transact-Structured Query Language, to analyze selected rheumatology notes using automated text processing to identify text unique to RA templates and text strings associated with the TJC, the SJC, and PGA. For example, the following text string indicates the use of an RA template: "Tender Joint Count =" During the development phase, we used an iterative process that included multiple runs and re-evaluations of the document retrieval and text-extraction algorithms to refine and optimize performance. During the iterative process of automated text processing development, 17 variants were identified, with 6 variants associated with TJC, 6 with SJC, and 7 with PGA (Supplemental Table S1). We used corresponding extraction algorithms by using regular expressions to identify these variants and extract their associated response values.

Development of automated systems to extract ESR values from the EHR. We extracted ESR values and the associated dates the ESR was collected (from December 1, 2013, to January 31, 2015) from the CDW and linked the ESRs with rheumatology clinic notes, allowing for up to 30 days between laboratory measurement and clinic visit. If multiple ESRs were obtained during this time window, the ESR collected closest to the clinic visit date was used. If there were two ESR values at equal time intervals before and after the clinical visit, the ESR before the clinic visit was used.

Establishing the reference/gold standard for evaluation. To establish a reference/gold standard, we first identified all clinic notes associated with a VARA registry database entry and all notes associated with at least one clinical DAS28 element identified by the automated text processing method. We conducted chart review when inconsistencies occurred in documents identified/retrieved or when inconsistencies occurred in extracted DAM elements.

Each note was classified as being identified by manual extraction only, automated text processing only, or both manual extraction and automated text processing. Clinical notes identified by only one method were evaluated by chart review, and reasons for the failure of either manual extraction or automated text processing to identify the note was determined, but no modifications were made to the automated text processing during the validation phase.

Each DAS28 element was assessed to determine whether the value for that element was reported by manual extraction only, automated text processing only, or both manual extraction and automated text processing. When a clinical DAS28 element or ESR was reported by both manual extraction and automated text processing, we compared the value for the data point by each method and determined whether the value was identical or discordant.

Notes that included at least one missing or nonidentical DAM value were subjected to manual chart review, and the DAM value

identified on the chart review was considered the reference standard. If a DAS28 element was identified by both manual extraction and automated text processing and the values were identical by both methods, the information for that element was assumed to be accurate and considered the reference standard. We evaluated the assumption that identical values were correct by extracting all DAM elements if the note was flagged for chart review because of any information being discordant, as described previously. In 679 (20.7%) notes subjected to review, the concordant values identified by both manual extraction and automated text processing were correct in 100% of the notes reviewed.

Accuracy of data retrieved by manual extraction and automated text processing. For each element, the value reported by manual extraction and automated text processing was compared with the reference standard during the derivation and validation periods. The accuracy for each DAS28 element was calculated as the number of notes in which that element was identical to the reference standard. In addition to evaluating the accuracy of each element individually, we reported the number of notes that had all four elements needed to calculate the DAS28 that were equivalent to the reference standard.

Human subjects review. Each site received institutional review board approval, and all participants provided written informed consent to participate in the VARA registry.

RESULTS

Identification of clinical notes containing DAMs.

There were 1699 notes with at least 1 value for the TJC, the SJC, or PGA in 633 unique patients identified during the derivation period, and there were 1569 notes with at least 1 value for the TJC, the SJC, or PGA in 596 unique patients during the validation period (Table 1). There were 1412 (83%) clinical notes detected by both manual extraction and automated text processing, 188 (11%) clinical notes identified only by manual extraction, and 99 (6%) clinical notes identified only by automated text processing during the derivation period. During the validation period, there were 1408 (90%) clinical notes detected by both manual extraction and automated text processing, 86 (5%) clinical notes detected only by manual extraction, and 75 (5%) clinical notes detected only by automated text processing.

Of the 1412 EHR notes that were detected during the derivation phase by both manual extraction and automated text processing and that included at least 1 DAM, 1059 (62%) contained all 4 elements of the DAS28 (TJC, SJC, PGA, and ESR). In the validation phase, 1167 (74%) notes identified using both methods contained all 4 elements. The notes above were a subset of all rheumatology notes on these subjects. Of all rheumatology notes (which could include procedure notes, telephone contacts, and other clinical encounters), there were 1699 of 2601 total rheuma-

Table 1. Number of electronic health record notes identified by manual extraction and automated text processing

	Identified by Both Methods	Manual Extraction Only	Automated Text Processing Only	Absent for Both Methods
	n, % (95% CI)	n, % (95% CI)	n, % (95% CI)	n, % (95% CI)
Derivation set (N = 1699)				
Notes with at least one value for TJC, SJC, or PGA recorded	1412, 83% (85%-81%)	188, 11% (13%-10%)	99, 6% (7%-5%)	...
Notes for individual components				
TJC	1199, 71% (73%-68%)	398, 23% (25%-21%)	71, 4% (5%-3%)	31, 2% (2%-1%)
SJC	1195, 70% (73%-68%)	400, 24% (26%-22%)	72, 4% (5%-3%)	32, 2% (3%-1%)
PGA	1322, 78% (80%-76%)	214, 13% (14%-11%)	85, 5% (6%-4%)	78, 5% (6%-4%)
All three clinical variables	1100, 65% (67%-62%)	431, 25% (27%-23%)	57, 3% (4%-2%)	111, 7% (8%-5%)
ESR	1388, 82% (84%-80%)	171, 10% (11%-9%)	100, 6% (7%-5%)	40, 2% (3%-2%)
All DAMs collected to allow calculation of DAS28	1059, 62% (65%-60%)	434, 26% (28%-23%)	61, 4% (4%-3%)	145, 9% (10%-7%)
Validation Set (N = 1569)				
Notes with at least one value for TJC, SJC, or PGA recorded	1408, 90% (91%-88%)	86, 5% (7%-4%)	75, 5% (6%-4%)	...
Notes for individual components				
TJC	1290, 82% (84%-80%)	166, 11% (12%-9%)	71, 5% (6%-3%)	42, 3% (3%-2%)
SJC	1287, 82% (84%-80%)	167, 11% (12%-9%)	72, 5% (6%-4%)	43, 3% (4%-2%)
PGA	1341, 85% (87%-84%)	101, 6% (8%-5%)	59, 4% (5%-3%)	68, 4% (5%-3%)
All three clinical variables	1211, 77% (79%-75%)	189, 12% (14%-10%)	55, 4% (4%-3%)	114, 7% (9%-6%)
ESR	1371, 87% (89%-86%)	77, 5% (6%-4%)	74, 5% (6%-4%)	47, 3% (4%-2%)
All DAMs collected to allow calculation of DAS28	1167, 74% (77%-72%)	189, 12% (14%-10%)	59, 4% (5%-3%)	154, 10% (11%-8%)

Abbreviation: CI, confidence interval; DAM, disease activity measure; DAS28, Disease Activity Score for 28 joints; ESR, erythrocyte sedimentation rate; PGA, patient global assessment of disease activity; SJC, swollen joint count; TJC, tender joint count.

tology notes (65%) in the derivation phase and 1569 of 2761 total notes (57%) in the validation phase that were used in this analysis, showing that the majority of rheumatology notes on these VARA subjects contained DAM elements.

Reasons for failure to detect notes reporting DAMs.

We identified the reasons that clinical notes were not detected by automated text processing, and only detected by manual extraction, in the derivation set. These reasons included the use of an incorrect note template (87; 46.2%), modification of the standard template (13; 6.9%), missing data in the CDW database (74; 39.4%), entry of data in nonnumerical format (7; 3.7%), and transfer of patients between VA facilities (7; 3.7%), with simi-

lar findings in the validation set (Table 2). The reasons notes were not detected by manual extraction, and only detected by automated text processing, in the derivation set included failure to identify the notes by chart abstractor or research assistant (94; 94.9%) and data entry errors (5; 5.1%), with similar findings in the validation set.

Accuracy of data retrieval by manual extraction and automated text processing.

TJC and SJC. Because the TJC and SJC were almost always recorded concurrently, the issues with data collection of these elements were highly correlated. The accuracy of the TJC was 95.2% and 95.9% by manual extraction during the derivation and validation periods,

Table 2. Reasons for failure to detect computerized patient record system notes or disease activity measures by manual extraction or automated text processing

	Derivation Set	Validation Set
	n, % (95% CI)	n, % (95% CI)
N = 188		
Notes only detected by manual extraction		
Standard template not used	87, 46.2% (53.4%-39.1%)	47, 54.7% (65.2%-44.1%)
Correct template modified	13, 6.9% (10.5%-6.4%)	5, 5.8% (10.8%-0.8%)
Data failed to capture note	74, 39.4% (46.3%-32.4%)	22, 25.6% (34.9%-16.3%)
Data entry not in numerical format	7, 3.7% (6.4%-1%)	11, 12.8% (19.9%-5.7%)
Move between VARA registry sites	7, 3.7% (6.4%-1%)	1, 1.1% (3.4%-0%)
N = 99		
N = 75		
Notes only detected by automated text processing		
Investigators did not identify note for manual extraction	94, 94.9% (99.3%-90.6%)	73, 97.3% (100%-93.6%)
Data entry errors by manual extraction	5, 5.1% (9.3%-0.7%)	2, 2.7% (6.3%-0%)

Abbreviation: CI, confidence interval; VARA, Veterans Affairs Rheumatoid Arthritis.

respectively, compared with 76.2% and 89.2% by automated text processing during the same respective periods (Table 3). The accuracy of the SJC was 95.4% and 94.6% by manual extraction during the derivation and validation periods, respectively, compared with 76.4% and 89.2% by automated text processing during the same respective periods. A chart review of these discrepancies showed that other than six episodes of data capture failure from the CDW, the reasons for failure of the automated text processing method to detect the TJC and SJC was related to incorrect use of the note template through either modification, deletion, or entering text in place of numerical data. For example, in the derivation set, there were 97 reports of “no tenderness” and 113 reports of “no swelling” in the EHR note. The entries by the investigator doing the manual extraction and putting data into the VARA database for these episodes were “0” for the TJC and “0” for the SJC.

There were 321 cases in which the TJC and SJC were detected correctly by automated text processing but not entered by manual extraction in the VARA database. In 287 (89.4%) of these 321 episodes, the investigator entering data into the VARA database by manual extraction did not make any entry into the VARA database when these data were present on the EHR chart review. In 34 (10.6%) of these 321 episodes the investigator entered incorrect data by the manual extraction process.

PGA. The accuracy of PGA was 91.5% by manual extraction during the derivation period, compared with 91.3% in the validation period (Table 3). The accuracy of PGA was 76.2% by automated text processing during the derivation period, compared with 93.3% in the validation period (Table 3). A chart review of these discrepancies showed that the reason for failure of the automated text processing method to detect PGA was related to incorrect use of the note templates through either modification or deletion in the vast majority of notes (212 [95.1%] of 223 notes during the derivation period and 102 [97.1%] of 105 during the validation period, with only 11 [4.9%] and 3 [2.9%]

notes with PGA data entered as text in the derivation and validation periods, respectively).

There were 143 notes in the derivation period and 135 notes in the validation period in which PGA was detected correctly by automated text processing but not entered by manual extraction. During the derivation period, 85 (59.4%) of these notes were cases in which the investigator entering data into the VARA database by manual extraction did not make any PGA entry into the VARA database when data were present on the chart review, and 58 (40.5%) notes were cases in which incorrect data were entered by the manual extraction process. During the validation period, 59 (43.7%) of these notes were cases in which the investigator entering data into the VARA database by manual extraction did not make an entry of PGA into the VARA database when data were present on the chart review, and 76 (56.3%) notes were cases in which the investigator entered incorrect data by the manual extraction process.

ESR. The accuracy of the ESR by manual extraction was 88.6% during the derivation period and 92.3% during the validation period (Table 3). The accuracy of ESR capture by automated text processing was 90.9% during the derivation period, compared with 95.7% during the validation period (Table 3). Our extraction program for the ESR by automated text processing required that a note have at least one of the three clinical data elements (TJC, SJC, or PGA). When a note with one of these clinical elements was identified by automated text processing, the automated text processing method was triggered to search for ESR values within 30 days of that date. When automated text processing identified a clinical element, it successfully reported ESRs for 1510 (99.9%) of 1511 notes during the derivation period and 1483 (100%) of 1483 notes during the validation period. For two cases, the ESR value was not identified because of a data capture error in the CDW.

In 192 cases, the ESR was correctly identified by automated text processing, but was not identified by manual extraction, during the derivation phase. For 100 (52.1%) of these notes, the investigator failed to enter the ESR value, and for the other 92

Table 3. Accuracy of manual extraction and automated text processing at correctly identifying clinical and laboratory elements in comparison with chart review gold standard

	Manual Extraction	Automated Text Processing
	n, % (95% CI)	n, % (95% CI)
Derivation set (N = 1699)		
Tender joint count	1618, 95.2% (96.2%-94.2%)	1295, 76.2% (78.2%-74.1%)
Swollen joint count	1621, 95.4% (96.4%-94.4%)	1299, 76.4% (78.4%-74.4%)
Patient global assessment	1556, 91.5% (92.9%-90.2%)	1476, 86.8% (88.4%-85.2%)
Westergren erythrocyte sedimentation rate	1507, 88.6% (90.2%-87.1%)	1544, 90.9% (92.2%-89.5%)
All DAMs collected to allow calculation of DAS28	1345, 79.1% (81.0%-77.2%)	1117, 65.7% (68.0%-63.5%)
Validation set (n = 1569)		
Tender joint count	1491, 95.0% (96.1%-93.9%)	1401, 89.2% (90.8%-87.7%)
Swollen joint count	1485, 94.6% (95.7%-93.5%)	1400, 89.2% (90.7%-87.6%)
Patient global assessment	1434, 91.3% (92.7%-90.0%)	1464, 93.3% (94.5%-92.0%)
Westergren erythrocyte sedimentation rate	1449, 92.3% (93.6%-91.0%)	1502, 95.7% (96.7%-94.7%)
All DAMs collected to allow calculation of DAS28	1230, 78.3% (80.4%-76.3%)	1226, 78.1% (80.2%-76.1%)

Abbreviation: CI, confidence interval; DAM, disease activity measure; DAS28, Disease Activity Score for 28 joints.

(47.9%) notes, ESR values were entered incorrectly. In 120 cases correctly identified by automated text processing but missed by manual extraction during the validation phase, 73 (60.8%) were cases in which the investigator failed to enter the ESR, and 47 (39.2%) were cases in which an incorrect value for the ESR was entered. Incorrect values were usually entered because the ESR closest to the visit was not selected or because an ESR outside the 30-day range was entered.

Clinical notes with full data reported to allow for calculation of the DAS28. Complete clinical and laboratory data sufficient to calculate the DAS28 were present by manual extraction in 79.1% of clinical notes during the derivation phase and in 78.3% of clinical notes during the validation phase. Complete clinical and laboratory data sufficient to calculate the DAS28 were present by automated text processing in 65.7% of clinical notes during derivation phase and in 78.0% of clinical notes during the validation phase.

DISCUSSION

Our work has demonstrated that our automated text processing approach can successfully extract DAMs directly from the EHR in the VA health care system. This study provided insight about the novel extraction process through comparison against the established manual approach as well as a chart review–adjudicated reference standard. The inaccuracies of both systems are predominately a result of different human errors. Manual extraction has the advantage of overcoming human errors when templates are modified or when data are entered as text rather than as numerical values. However, manual extraction has the potential for failure when clinical notes are not properly identified and retrieved for extraction and as a result of data entry errors.

In comparison, performance of automated text processing is impacted by human error related to initial data entry into the EHR clinic note at the point of care. Examples of these errors include incomplete data entry, using text instead of numerical values, or modifying templates. Automated text processing is also subject to failure of electronic data capture, which, although rare, can occur. However, automated text processing has distinct advantages over manual extraction, including efficiency, transportability across sites, and improved accuracy.

Our experience after the implementation phase has resulted in a significant reduction of investigator time (estimated at 5 minutes per patient for each DAM entry), which has greatly improved our efficiency in collection of DAMs. This new methodology provides a critically needed system if we are going to expand the VARA registry to other VA sites where clinical research support is not available for the manual extraction of DAMs. Our system used a structured text and was not a traditional natural language processing method. We required an exact match for data extraction. This method has the potential for a high degree of accuracy in any system that can consistently use standard templates and is not

limited to VA-based clinical care. As noted previously, template modification is a significant challenge, and we are currently developing educational efforts, standard template use, and an audit and feedback program to promote consistent data collection by using standard templates in the VARA registry. These practices could also be implemented at non-VA sites.

Several other RA registries have been established, and many of these registries collect DAMs (13–19). These registries use electronic data sources as well as paper questionnaires (16–18). Although these registries may validate key data elements, such as rheumatic disease diagnosis (20), serious treatment-related adverse events (16), and other relevant health outcomes (21), they generally report DAMs without validating their accuracy. Outside of rheumatic disease registries, EHR systems have been used in rheumatology for case finding (22,23) but have not routinely been used to collect DAM assessments as part of routine care. Efforts have also been undertaken to have DAMs (including patient-reported outcomes) entered online (24) and to drive changes in clinical practice (25,26). Our group and others have previously reported on efforts to use EHR administrative data to identify rheumatoid disease activity, but this effort has not performed with sufficient accuracy to be used in clinical practice (27–29). This previous work emphasizes the need for systems to accurately collect DAMs in clinical practice.

We recognize that our system is only one of many efforts to employ information technology to improve clinical care and research. Computer programs and handheld device applications are also available for calculation of DAMs (25,30–32). These applications can be available in clinics for real-time calculations and, if desired, can be copied into the medical record. These resources are certainly available to assist in clinical RA management. Our automated text processing approach has the distinct advantage of being fully integrated with the VA EHR computer system; another advantage is that the information may be recorded directly from the clinic note without needing to be copied into another application. We are currently developing and doing initial testing of the display of these data elements in a longitudinal dashboard that provides DAMs over time and in correlation with disease modifying anti-rheumatic drug therapy. We hope our efforts can contribute to this knowledge in the expansion of this field and as these systems are more fully deployed in medical practice to improve the care of patients with RA.

This study has several strengths, including the availability of both a gold standard and manual extraction method with which the automated text processing system can be compared, involvement of multiple sites, a uniform EHR that is used across the VA system, and a group of committed investigators necessary for the implementation and evaluation of the automated text processing method. Our study has the limitation of having a single EHR with limited use outside the VA; however, we feel that the principles used in this work could be applied to other health care systems.

In conclusion, the development of this validated electronic system for DAM extraction provides not only an efficient system to collect these data for use in outcomes research but also the potential to provide an accurate and reliable system for the presentation of these data to clinicians and/or patients in the clinical practice setting. Having immediate and facile access to these data provides a significant opportunity to enhance clinical decision-making in disease management and improve the quality of care received by patients with RA.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Cannon, Reimold, Mikuls, Sauer.

Acquisition of data. Cannon, Rojas, Reimold, Mikuls, Bergman, Sauer.

Analysis and interpretation of data. Cannon, Rojas, Reimold, Mikuls, Sauer.

REFERENCES

- Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol* 2016;68:1–26.
- Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
- Yu Z, Lu B, Agosti J, Bitton A, Corrigan C, Fraenkel L, et al. Implementation of treat-to-target for rheumatoid arthritis in the US: analysis of baseline data from a randomized controlled trial. *Arthritis Care Res (Hoboken)* 2018;70:801–6.
- Zak A, Corrigan C, Yu Z, Bitton A, Fraenkel L, Harrold L, et al. Barriers to treatment adjustment within a treat to target strategy in rheumatoid arthritis: a secondary analysis of the TRACTION trial. *Rheumatology (Oxford)* 2018;57:1933–7.
- Curtis JR, Sharma P, Arora T, Bharat A, Barnes I, Morrissy MA, et al. Physicians' explanations for apparent gaps in the quality of rheumatology care: results from the US Medicare Physician Quality Reporting System. *Arthritis Care Res (Hoboken)* 2013;65:235–43.
- Mikuls TR, Reimold A, Kerr GS, Cannon GW. Insights and implications of the VA Rheumatoid Arthritis registry. *Fed Pract* 2015;32:24–9.
- Cannon GW, Mikuls TR, Hayden CL, Ying J, Curtis JR, Reimold AM, et al. Merging Veterans Affairs rheumatoid arthritis registry and pharmacy data to assess methotrexate adherence and disease activity in clinical practice. *Arthritis Care Res (Hoboken)* 2011;63:1680–90.
- Miriovsky BJ, Michaud K, Thiele GM, O'Dell JR, Cannon GW, Kerr G, et al. Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis. *Ann Rheum Dis* 2010;69:1292–7.
- Mikuls TR, Fay BT, Michaud K, Sayles H, Thiele GM, Caplan L, et al. Associations of disease activity and treatments with mortality in men with rheumatoid arthritis: results from the VARA registry. *Rheumatology (Oxford)* 2011;50:101–9.
- Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- 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;38:44–8.
- Fihn SD, Francis J, Clancy C, Nielson C, Nelson K, Rumsfeld J, et al. Insights from advanced analytics at the Veterans Health Administration. *Health Aff (Millwood)* 2014;33:1203–11.
- Vermeer M, Kuper HH, Hoekstra M, Haagsma CJ, Posthumus MD, Brus HL, et al. Implementation of a treat-to-target strategy in very early rheumatoid arthritis: results of the Dutch Rheumatoid Arthritis Monitoring remission induction cohort study. *Arthritis Rheum* 2011;63:2865–72.
- Francisco M, Johansson T, Kazi S. Overview of the American College of Rheumatology's electronic health record-enabled registry: the Rheumatology Informatics System for Effectiveness. *Clin Exp Rheumatol* 2016;34 Suppl 101:S102–4.
- Michaud K. The National Data Bank for Rheumatic Diseases (NDB). *Clin Exp Rheumatol* 2016;34 Suppl 101:S100–1.
- Curtis JR, Jain A, Askling J, Bridges SL Jr, Carmona L, Dixon W, et al. A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid arthritis registries. *Semin Arthritis Rheum* 2010;40:2–14.
- Yazdany J, Bansback N, Clowse M, Collier D, Law K, Liao KP, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. *Arthritis Care Res (Hoboken)* 2016;68:1866–73.
- Kremer JM. The Corona US registry of rheumatic and autoimmune diseases. *Clin Exp Rheumatol* 2016;34 Suppl 101:S96–9.
- Tonner C, Schmajuk G, Yazdany J. A new era of quality measurement in rheumatology: electronic clinical quality measures and national registries. *Curr Opin Rheumatol* 2017;29:131–7.
- Curtis JR, Baddley JW, Yang S, Patkar N, Chen L, Delzell E, et al. Derivation and preliminary validation of an administrative claims-based algorithm for the effectiveness of medications for rheumatoid arthritis. *Arthritis Res Ther* 2011;13:R155.
- Fisher MC, Furer V, Hochberg MC, Greenberg JD, Kremer JM, Curtis JR, et al. Malignancy validation in a United States registry of rheumatoid arthritis patients. *BMC Musculoskelet Disord* 2012;13:85.
- Liao KP, Cai T, Gainer V, Goryachev S, Zeng-treitler Q, Raychaudhuri S, et al. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2010;62:1120–7.
- Carroll RJ, Thompson WK, Eyer AE, Mandelin AM, Cai T, Zink RM, et al. Portability of an algorithm to identify rheumatoid arthritis in electronic health records. *J Am Med Inform Assoc* 2012;19:e162–9.
- Williams CA, Templin T, Mosley-Williams AD. Usability of a computer-assisted interview system for the unaided self-entry of patient data in an urban rheumatology clinic. *J Am Med Inform Assoc* 2004;11:249–59.
- Jacobs JW, Utrecht Rheumatoid Arthritis Cohort study group. The Computer Assisted Management in Early Rheumatoid Arthritis programme tool used in the CAMERA-I and CAMERA-II studies. *Clin Exp Rheumatol* 2016;34 Suppl 101:S69–72.
- Hetland ML, Krogh NS, Horslev-Petersen K, Schiottz-Christensen B, Sorensen IJ, Dorte Vendelbo J. Using an electronic platform interactively to improve treatment outcome in patients with rheumatoid arthritis: new developments from the DANBIO registry. *Clin Exp Rheumatol* 2016;34 Suppl 101:S75–8.
- Sauer BC, Teng CC, Accortt NA, Burningham Z, Collier D, Trivedi M, et al. Models solely using claims-based administrative data are poor predictors of rheumatoid arthritis disease activity. *Arthritis Res Ther* 2017;19:86.

28. Ting G, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Young M, et al. Development of a health care utilisation data-based index for rheumatoid arthritis severity: a preliminary study. *Arthritis Res Ther* 2008;10:R95.
29. Sato M, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Avorn J, et al. The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28. *Arthritis Res Ther* 2006;8:R57.
30. Leeb BF, Brezinschek HP, Rintelen B. RADAI-5 and electronic monitoring tools. *Clin Exp Rheumatol* 2016;34 Suppl 101:S5–10.
31. Aletaha D, Bécède M, Smolen JS. Information technology concerning SDAI and CDAI. *Clin Exp Rheumatol* 2016;34 Suppl 101: S45–8.
32. Pincus T. Electronic eRAPID3 (Routine Assessment of Patient Index Data): opportunities and complexities. *Clin Exp Rheumatol* 2016;34 Suppl 101:S49–53.