DEVLOPMENT AND EVALUATION OF DISEASE ACTIVITY MEASURES IN RHEUMATOID ARTHRITIS USING MULTI-LEVEL MIXED MODELING AND OTHER STATISTICAL METHODOLOGIES

A Dissertation Presented

By

Mary Jane R. Bentley

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

January 28, 2010

Ph.D. Program in Clinical and Population Health Research

DEVELOPMENT AND EVALUATION OF DISEASE ACTIVITY MEASURES IN RHEUMATOID ARTHRITIS USING MULTI-LEVEL MIXED MODELING AND OTHER STATISTICAL METHODOLOGIES

A Dissertation Presented

By

Mary Jane R. Bentley

Approval as to style and content of the dissertation by:

_____
Sybil Crawford, PhD., Chair of Committee

_____
Patricia Franklin, M.D., M.B.A. Member of Committee

_____
Robert Zurier, M.D., Member of Committee

_____
Carol Oatis, PhD, Member of Committee

_____
George W. Reed, PhD, Thesis Advisor

_____
Anthony Carruthers, Ph.D., Dean of the Graduate School of Biomedical Sciences

Ph.D. Program in Clinical and Population Health Research

January 2010

DEDICATION

Dedicated to my beloved family especially:

My parents, Ruth Emmott Plante and Edmond Plante who instilled in me a love of education, a strong work ethic, and a compassion for my fellow man. I will love and admire you always.

My husband, W. Robert Bentley II, who encouraged me to reach for my dreams and persevere to achieve them. You are truly the “wind beneath my wings” and I love you.

My children, William, Christopher and Megan Campbell who inspired me to make this world a better place. You are strong, meritous, intelligent, warriors and champions and I am very proud of who you have become. I love you always and I cherish the time we have spent together.

My mother-in-law, Margaret E. Bentley, a wonderful woman with an open heart

My former father-in-law, Kenneth A. Campbell, a good, honest man who showed me kindness when I was of need.

And to Maggie, my cherished companion and friend always
and to sweet Minnie.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my thesis advisor, Dr. George Reed, for his encouragement, support and assistance throughout my graduate school studies. In addition to being an excellent mentor, he is a brilliant biostatistics professor, a wonderful role model, and a very good friend. His insight and breadth of knowledge of biostatistics provided me with the support and assistance I needed to conduct my research and become an independent investigator. I am truly grateful to him for all of his assistance and patience.

I also wish to thank my dissertation examination committee members: my chair, Dr. Sybil Crawford, Dr. Patricia Franklin, Dr, Robert Zurier and Dr. Carol Oatis. Your expert opinions and thoughtful comments and critiques were instrumental in helping me to complete my dissertation research. I am a truly grateful to you all for your support and guidance throughout this process and I know I am a better researcher due to your influences.

I would like to extend my sincere gratitude to Dr. Leslie Harrold, Dr. Jeffrey Greenberg and Dr. Joel Kremer for your assistance and helpful reviews of my manuscripts. I would also like to thank Dr. Kremer for allowing me to utilize the rich data from the CORRONA registry to conduct my research.

I wish to thank members of Dr. Reed's Biostatistics Research Group for providing me with expert programming and data analysis support throughout my graduate studies, and especially Ying Shan, Ping He and Katherine Leung. I also wish to thank Jean Villa for your administrative support.

I would like to extend sincere thanks as well to the CPHR faculty and staff, especially Dr. Carole Upshur for your efforts at establishing the program and continually improving it and Dr. Stephenie Lemon as a wonderful role model and student advisor. Thank you also to the CPHR staff: Tricia Doane, Ann Michelson and Colleen Corey for all of your help and answering my many questions.

Lastly, I would like to thank my fellow classmates of the inaugural class of the CPHR program: Dr. Hongliu Ding, Dr. Mayra Tisminstsky, and Dr. Ginny Briggs for your friendship and support. We share many memories and will always have the distinction of being the first four students and the first four graduates of the CPHR program. We did it!

ABSTRACT

Remarkable progress has been made in the development of effective treatments for patients with rheumatoid arthritis (RA). To ensure that a patient is optimally responding to treatment, consistent monitoring of disease activity is recommended. Established composite and individual disease activity measures often cannot be computed due to missing laboratory values. Simplified measures that can be calculated without a lab value have been developed and previous studies have validated these new measures, yet differences in their performance compared with established measures remain. Therefore, the goal of my doctoral research was to examine and evaluate disease activity and composite measures to facilitate monitoring of response in clinical care settings and inclusion of patients with missing laboratory values in epidemiological research.

In the first study, the validity of two composite measures, the Clinical Disease Activity Index (CDAI) and the Disease Activity Score with 28 joint count (DAS28) was examined and both were significantly associated with a rheumatologist's decision to change therapy (CDAI OR=1.58; 95% CI: 1.42, 1.76) (DAS28 OR=1.34; 95% CI 1.27,1.56). However, further evaluation using receiver operating characteristic (ROC) analysis found that they were not strong predictors of physician decisions to change therapy (AUC=0.75, 0.76, respectively). Thus, they should not be used to guide treatment decisions in the clinic.

Two measures of disease activity, erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) are often not measured and impede the computation of composite measures of disease activity. In the second study, significant factors which may predict the measurement of the ESR and CRP were identified and included physician and clinical variables but no quantitative disease activity measures. Thus the suitability of the ESR and CRP as measures of disease activity is suspect.

In the final study, I created a new composite measure, the modified disease activity score with 28 joint count (mDAS28), by replacing the laboratory value in the DAS28. The mDAS28 was then validated by comparing its performance with the DAS28. The measures were strongly correlated ($r=0.87$), and strong agreement was found between the two measures when categorizing patients to levels of disease activity ($\kappa=0.77$) and treatment response ($\kappa=0.73$). Therefore, the mDAS28 could be used in place of the DAS28 when laboratory values needed to compute the DAS28 are missing.

In summary, I found that the CDAI and DAS28 were not strong predictors of the rheumatologist's decision to change therapy. I also found that the variability in the measurement of ESR and CRP was not associated with disease activity. I was able to modify the DAS28 by replacing the laboratory measure and create a new simplified measure, the mDAS28. I also validated the mDAS28 for use in the clinic and in epidemiological research when the DAS28 is unavailable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF THE FREQUENTLY USED ABBREVIATIONS

ACR- American College of Rheumatology

APR- Acute Phase Reactant

AUC- Area under the curve

BMI- Body Mass Index

CDAI- Clinical Disease Activity Index

CI- Confidence Index

CORRONA- The Consortium of Rheumatologist Researchers of North America

CRP- C-reactive Protein

DAS- Disease Activity Score

DAS28- Disease Activity Score with based on 28 joints

DMARDs- Disease modifying anti-rheumatic drugs

EGA- Rheumatologist's global assessment of disease activity

ES- Effect Size

ESR- Erythrocyte sedimentation rate

EULAR- European League Against Rheumatism

HAQ-Health Assessment Questionnaire

ICC- Intraclass Correlation Coefficient

lnESR- Logarithm of ESR

mDAS28- Modified Disease Activity Score with 28 joint count

mEULAR- Modified European League Against Rheumatism response criteria

LIST OF THE FREQUENTLY USED ABREVIATIONS (continued)

mHAQ- Modified Health Assessment Questionnaire

MOR- Median Odds Ratio

MTX- Methotrexate

OR- Odds Ratio

PGA- Patient global assessment of disease activity

PhGA- Physician global assessment of disease activity

RA- Rheumatoid arthritis

RCT- Randomized Controlled Trial

RF- Rheumatoid factor

ROC- Receiver operating characteristic curve analysis

SD- Standard deviation

SDAI- Simplified Disease Activity Index

SJC- Swollen joint count

SRM- Standardized Response Mean

TJC- Tender joint count

VAS- Visual analog scale

VIF- Variable Inflation Factor

PREFACE

My dissertation was written in the format of three manuscripts (Chapters 2-4) to be submitted for publication. The document was reformatted to comply with the requirements of the Graduate School of Biomedical Sciences at University of Massachusetts Medical School.

Part of this thesis work has been published previously as:

- **Bentley, MJ**, Reed, GW. Simplified disease activity measures: can they be used in standard care? *J Clin Exp Rheumatol,* 26 (2)*:* 358-368, 2008. - Recommended by the Faculty of 1000 as a must read.

In addition, publications related to the PhD study but not presented in this thesis are listed as follows:

- Greenberg JD, Harrold LR, **Bentley MJ,** Kremer, J, Reed, G, Strand V. Evaluation of composite measures of treatment response without acute-phase reactants in patients with rheumatoid arthritis. *Rheumatology 2*009;48:686-690
- Harrold, Leslie R., Greenberg, Jeffrey D., Curtis, Jeffrey R., **Bentley, Mary jane**, Reed, George, Harrington, J. Timothy; Rheumatologists Prescribing Patterns for Rheumatoid Arthritis Patients with Active Disease. Arthritis Rheum 2009;60(Suppl 10) :1009

CHAPTER I

Introduction

1.1 Description and Consequences of Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a systemic, chronic condition that can cause joint damage and destruction, disability and death (1). RA occurs worldwide and affects about 1% of the adult population (2). In the United States, approximately 2.1 million people are affected (1). The disease is about 5 times more prevalent in indigenous peoples of North America compared with Caucasians in Europe and North America. Asian and African populations have a much lower prevalence of RA (1).

The disease is characterized by joint pain, stiffness and swelling due to synovial inflammation and effusion. It has a variable disease course among patients with frequent flares and fluctuations in disease activity. The disease is a heterogeneous disorder with multifaceted clinical features between patients (1). Common symptoms of RA include tenderness and swollen joints, pain, stiffness, limited range of motion in joints, loss of appetite, muscle pain and skin nodules (3). RA can begin in any joint and commonly begins in the smaller joints of the fingers, hands and wrists. Management of the disease is made difficult due to this multi-faceted nature. Frequent assessment of disease activity is necessary to ensure effectiveness of treatment, but assessment using only a single quantitative disease measure such as an acute phase reactant or joint count is inaccurate for groups of patients (4).

The cause of RA is unknown. RA occurs 2 to 3 times more often in women then in men, which suggests possible hormonal factors triggering or modulating

the onset of the disease. In addition, gender appears to influence the phenotype of the disease with more women experiencing structural joint damage and more men affected by erosions to the bones. The peak age of onset of adult RA is between 40 and 60 years old but can begin as early as the second decade. Recent research suggests that the age of onset is shifting towards later in life but this trend could be artifactual and due more to the increase in an aging population. Patients with RA may become considerably catabolic and also may experience significant anorexia and malnourishment. On the other hand, sedentary and extremely poor eating habits in other patients with RA may lead to obesity impacting already damaged weight bearing joints and other organ systems. Other factors which influence the course and outcome of the disease are formal education and marital status as well as smoking (1).

The consequences of RA are substantial (5). Affected individuals experience significant joint damage and destruction and RA is associated with several comorbidities, notably cardiovascular disease, which increases the mortality rate of affected individuals. As the disease progresses, a patients' daily activities and functional status are affected. RA has a substantial economic impact with over $26-$32 billion spent annually on long-term care and loss of employment (2). The enormous individual patient and economic consequences of RA can only be prevented by treatment with effective therapeutic agents and consistent disease activity monitoring to ensure that treatment is effective (6, 7).

1.2 Contemporary Treatment of Rheumatoid Arthritis

The therapeutic aim in rheumatoid arthritis is to achieve maximum response to treatment and the lowest disease activity level, ideally remission (8, 9). Up until the last decade, treatment options for RA were sparse and the few treatments that were available were prone to toxic side-effects and often not effective (1). Surgery was the mainstay of treatment especially to correct deformities and most patients became disabled (1). Over the last decade, the development of new and effective treatments especially disease modifying antirheumatic drugs has substantially improved the prognosis for patients with RA with many patients now experiencing lower disease activity and remission (9).

There are several different classes of drugs used to treat RA. Two kinds target pain relief and reduce inflammation but do not prevent tissue injury or progressive joint damage (10). They include: nonsteroidal anti-inflammatory drugs (NSAIDs) and analgesic drugs. NSAIDs include aspirin, ibuprofen, indomethacin, and COX-2 inhibitors such as valdecoxib and celecoxib and low dose corticosteroids or prednisone. Analgesic drugs include: acetaminophen, propoxyphene, meperidine, morphine and tramadol.

A more effective class of drugs includes the disease modifying antirheumatic drugs (DMARDs) (10). These drugs and to some extent glucocorticoids can control active synovitis and impede joint damage and destruction, but cannot reverse existing joint deformities or erosions. DMARDs are normally used in combination with NSAIDs and/or prednisone. Examples of

these drugs include: Methotrexate (MTX), injectable gold, penicillamine, azathioprine, chloroquine, hydroxychloroquine, sulfasalazine and oral gold. MTX, a component of most treatment regimens, has the highest retention rate compared with other DMARDs (6).

Recently, several new DMARDs, biologic response modifiers (biologics), were introduced, expanding the therapeutic armamentarium against RA (11). In addition to reducing inflammation, they can prevent joint damage but act differently than traditional DMARDs. These new DMARDs modify the immune system by inhibiting the production of proteins called cytokines, which contribute to inflammation. Examples of these drugs include: etanercept, infliximab, adaliumumab and anakinra. Clinical trials of these agents have reported that they may be more effective than traditional DMARDs because, in addition to halting symptoms of the disease, they also slow disease progression. Although highly effective, these agents are expensive with an annual treatment cost of $12,000 to $16,000. Clinical trials have also reported variability in the response of patients to these agents. Not all patients respond well to these agents, with many developing adverse effects, attesting to the heterogeneity of the disease (11).

Treatment recommendations from the American College of Rheumatology (ACR) published in 2008 (9) recommend aggressive treatment immediately after the patient is diagnosed to achieve tight control of disease activity and prevent bone and joint destruction. Several studies (12-14) have reported that aggressive

treatment of RA can prevent joint damage and thus prevent loss of function, loss of ability to work and expensive medical costs and surgery. Evidence that treatment of RA is more effective in the early stages includes several observational studies (12, 13, 15-17), which report substantial erosions of the bone occur in the first 2 years of the disease.

The ACR also recommends on-going measurement of disease activity to guide in the treatment decisions made by physicians and to ensure treatment effectiveness (9). A number of composite disease activity measures are available to rheumatologists but some have limitations when used in a clinical setting (18). They often cannot be calculated due to missing laboratory values, and thus, disease activity and treatment response cannot be measured in clinical care. In epidemiological research using observational registry data, patients with missing laboratory values, and thus missing composite disease activity scores, are often excluded from analyses, reducing the available cohort and most likely biasing the study results. For new DMARD therapies to demonstrate effectiveness in decreasing and preventing further joint damage, feasible measures to facilitate the consistent and accurate measurement of disease activity and response to treatment must be available for use in clinical care and in epidemiological research using registry data.

1.3 Measurement of Disease Activity and Treatment Response

1.3.1 Clinical Trials: Standard Composite Measures

Measurement of disease activity and response to treatment is complex due to the multifaceted nature of RA. No single quantitative 'gold standard' measure can assess and monitor disease activity (19, 20). Clinical trials have been the main source of information about the efficacy of new treatments for RA. In clinical trials, a variety of measures have been used (20). Up until the early 1990s, up to 10 different individual measures of disease activity representing different aspects of the disease, ranging from grip strength to joint counts to two acute phase reactants (APR) (21), the erythrocyte sedimentation rate (ESR) (22) and C-reactive protein (CRP) (23), were used to assess disease activity (24-27). Analyses of clinical trial results were hampered by these multiple individual measures for several reasons: the inability to directly compare the efficacy of new drugs to a common standard, potential reporting of only outcome measures that demonstrated an impressive result and the risk of a Type I error from multiple statistical testing (24). To standardize measurement of disease activity in clinical trials, it was recommended by the international rheumatology community that clinical trials measure a uniform core set of seven disease activity endpoints, each representing one component of the disease (28). Table 1.1 lists these endpoints:

Table 1.1 Core set of seven disease activity measures for clinical trials

Domain	Disease Activity Measures
Physician-reported	
Disease activity	Tender joint count (20, 24)
Disease activity	Swollen joint count (24)
Disease activity	Physician Global Assessment (20, 24)
Patient-reported	
Pain	Pain Assessment (20, 24)
Disease activity	Patient Global Assessment (20, 24)
Physical function	Often measured by the HAQ (20, 24)
Lab Values –measurement of one lab value	
Disease Activity	ESR(24) or CRP(24)

HAQ- Health Assessment Questionnaire; ESR-Erythrocyte Sedimentation Rate; CRP- C-reactive protein.

Use of this core set of seven individual measures in clinical trials improved the ability to compare results across trials. However, results were difficult to interpret because of multiplicity in the use of the individual measures resulting in a conflicting picture of efficacy (29). To improve interpretation of results, the core set of endpoints were combined to form a composite index which would produce one score of disease activity (30). The advantages of a composite index over individual measures include little duplicity between measures and increased ability to detect changes in disease activity over time or sensitivity to change (24, 31, 32). A study (33) reported that no single measure of disease activity could

distinguish between active or placebo treatment but a pooled index of these individual measures was effective. Therefore, the pooled index was a much more effective approach to assess clinic outcomes compared to individual measures. In addition to using composite indices in clinical trials to measure disease activity, criteria to measure change in disease activity or response to treatment were developed (34, 35).

Composite indices are continuous measures that can quantify a patients' level of disease activity into one score for ease of interpretation and comparability of clinical trial results. They are composed of the core set of individual disease activity measures (see Table 1.1) representing several aspects of the disease. The composite indices most widely used in randomized controlled trials (RCTs) are the Disease Activity Score (DAS) (19, 25, 36-38) and its modified version using 28 joint count, the DAS28 (19, 25, 39-41). The DAS and DAS28 are continuous measures with scores ranging from 0-10. Higher scores indicate worse disease activity. They combine data on swollen joint count (SJC), tender joint count, (TJC) ESR, and the patient's global health (PGA) measured by a visual analog scale. This thesis will utilize only the DAS28 and not the DAS since the DAS28 is used more often in assessing disease activity. The DAS28 is calculated using the following formula:

$$DAS28 = 0.56 * \sqrt{(28TJC)} + 0.28 * \sqrt{(28SJC)} + 0.70 * \log [ESR] + 0.014 * PGA$$

It is a simplified version of the DAS and utilizes the 28-joint tender and swollen joint counts instead of the more involved joint counts of the original DAS. The DAS28 has been shown to be as valid as the DAS (39).

A variation of the DAS28 has been developed, the DAS28-CRP and can be calculated using the laboratory value CRP instead of ESR (42). It has received little validity testing and thus is not used very often in clinics or in research. In this thesis, the DAS28 containing ESR will be used. Although there is no one "gold standard" measure in RA, many consider the DAS28 to be the gold standard and numerous studies have used the DAS28 as a comparator when validating new disease activity measures (16, 43-47).

Response criteria were established for use in RCTs to determine the effectiveness of new treatment modalities. Response can be defined as a "significant or relevant change in disease activity" (19). The two most widely used sets of response criteria in clinical trials are the American College of Rheumatology (ACR) improvement criteria (19, 25, 34) and the European League Against Rheumatism (EULAR) response criteria (25, 35).

The ACR improvement criteria were developed as a single score to measure the improvement or change from the start of treatment (baseline) to a given endpoint. The ACR criteria do not measure the actual state of disease activity. The ACR improvement criteria are comprised of the seven core set variables (see Table 1.1) while the EULAR response criteria, which are based on the DAS28 include only four (19). Usually clinical trials in RA had reported on the

average improvement (mean or median) of treated patients with the average improvement of one treatment compared to another. The efficacy of a treatment is determined by comparing group means of changes in disease activity variables. However, a significant difference between groups does not indicate the actual number of individual patients who responded to treatment. It was determined with the ACR improvement criteria that a uniform definition of improvement, the percentage of patients improving, could be compared across clinical trials. Furthermore, improvement in individual patients could be assessed. The definition of improvement determined by the ACR clinical trial patients is depicted in Table 1.2:

Table 1.2 ACR improvement criteria

- Improvement in Tender and Swollen joint counts

And improvement in 3 of the following 5 core set measures

- Pain scale
- Patient Assessment
- Physician Assessment
- Functional questionnaire
- Lab value (ESR or CRP)

ACR-American College of Rheumatology; ESR-Erythrocyte Sedimentation Rate; CRP-C-reactive protein;

The ACR response (ACR20) is defined as 20% improvement in tender and swollen joint counts and 20% improvement in 3 of the 5 core set measures in Table 2. ACR50, ACR70 or ACR90 reflect, 50%, 70% or 90% improvement in the above parameters.

The definition of response as measured by the EULAR response criteria differs conceptually from the ACR improvement. Rather than defining improvement or no improvement, the EULAR definition classifies patients into three groups, no response, moderate and good response. The EULAR response criteria reflect two components: the present level of the disease activity score using DAS28 and the change in disease activity from baseline (when the drug was initiated) to a given endpoint. A change in disease activity of 1.2 in an individual patient signifies a statistically significant change (37). The EULAR measures the relevant change in disease activity since the start of treatment and the level of disease activity at the follow-up visit. Table 1.3 depicts the measurement algorithm for the EULAR criteria.

Table 1.3 EULAR response criteria using DAS28

DAS28 at endpoint	Improvement in DAS28 from baseline		
	>1.2	>0.6 and ≤1.2	≤0.6
≤3.2	Good	Moderate	None
>3.2 and ≤5.1			
>5.1			

EULAR-European League Against Rheumatism; DAS28-Disease Activity Score with 28 joint count

Minimum response to determine active treatment from placebo in clinical trials is assessed by ACR20 and moderate EULAR values. Great improvement in the patient is indicated by ACR50, 70 and 90 and EULAR major response (48).

1.3.2 Clinical Care and Epidemiological Research: Simplified Composite Measures

The DAS28 and the response criteria measures have proven effective in clinical trials to measure disease activity level and response to treatment in groups of patients (19), but have not proven as useful in the daily practice setting (44) or in epidemiological research utilizing observational registry data (45). Since patients selected for clinical trials have a certain level of disease activity and since clinical trials follow patients for only a short period of time, it is not feasible to use the same measures of response in a clinical practice (19, 20). Computation of the DAS28 requires an acute phase reactant (APR), either the ESR or the CRP, that is oftentimes not available during the office visit preventing the score from being calculated at the clinical encounter (25, 49). Few studies have examined the frequency of measurement of ESR or CRP in clinical care or the correlates associated with measurement of ESR or CRP. Large variation in the clinical measurement of ESR and CRP has been reported (50, 51)**.** Physician practice style and training have been found to be related to measurement of ESR and CRP (50, 51). In addition, uncertainty about the value of the tests may influence physician laboratory monitoring practices (52). Studies need to be done

to determine why there is variability in the measurement of ESR and CRP and identify correlates influencing their measurement. Illumination of these factors could help improve the clinical management of RA by facilitating the consistent monitoring of disease activity to guide treatment as recommended by the ACR and also improve epidemiological research studies.

The EULAR response criteria is limited in use in the clinical care setting since it is derived using the DAS28. The ACR improvement criteria measure the percent improvement for an individual patient's response to treatment which was an important measure in clinical trials. But in clinical practice, it is more important to measure the actual amount of disease activity since the goal is to suppress disease activity then to determine the percentage of the patient's response (19). Thus the ACR response criteria will be of limited use in a clinical setting since it does not measure disease activity levels (19).

Two simplified composite disease activity indices have recently been developed to address the limitations of the DAS28 and the EULAR and ACR response criteria. These new measures, the Clinical Disease Activity Index (CDAI) (25, 53) and the Simplified Disease Activity Index (SDAI) (54) have been specifically designed to provide the physician with a feasible and accurate tool to monitor disease activity in a clinical setting. Both are numerical scores but the SDAI requires the lab value CRP to be computed and thus its use may be limited in clinical care. The CDAI does not require a lab value allowing it to be computed during the office visit to facilitate clinical decision-making while the patient is

present (49). It is comprised of tender joint count (TJC), swollen joint count (SJC), patient global assessment (PGA) and physician global assessment of disease activity (PhGA). The SDAI and the CDAI are calculated using the following formulas:

$$SDAI = 28\ TJC + 28\ SJC + PGA + PhGA + CRP$$

$$CDAI = 28\ TJC + 28\ SJC + PGA + PhGA$$

A score that can be computed without a lab value such as the CDAI was determined to be practical and was tested for validity in several studies (53) but questions still remain (18).

Because of the variable course of RA disease, clinical management of RA requires repeated assessment of disease activity and the effectiveness of treatment (9). One study (55), a multi-center, RCT, recently reported that systematic monitoring of disease activity in daily practice resulted in more patients with low disease activity due to more changes in DMARD treatment (6, 56). Measures that are feasible and valid in a clinical care setting comprised of patients with many different presentations of the disease have not been fully tested or developed (19). It is clear that research efforts are needed to evaluate the new simplified measures for validity and feasibility in a clinical care setting. The availability of valid tools to measure and monitor disease activity in the clinic and in epidemiological research could reduce disease activity in patients and lead to reduction in pain, joint damage and joint destruction. Therefore, valid

measures could reduce the economic impact and loss of employment for patients with RA and improve their quality of life.

1.4 Simplified composite disease activity measures in rheumatoid arthritis: should they be used in standard care? (published previously in Clinical and Experimental Rheumatology 2007; 33: 506-13)

(1) Abstract

Objective: To examine the validity, reliability, and predictive value of two recently developed composite disease activity measures, the Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI) in rheumatoid arthritis (RA) patients.

Methods: A systematic review of the published literature was performed between February 2003 and November 2007. Data was extracted regarding correlations of the SDAI and CDAI with standard clinical trial measures, the predictive ability of the measures and correlations with changes in radiographic scores. The ability of the measures to categorize patients according to their disease activity status compared to standard categories was also examined.

Results: Among 17 studies initially identified, 12 provided information on the validity and reliability of the SDAI and CDAI. These measures were found to be strongly correlated with the Disease Activity Score (DAS28) with correlation coefficients ranging from 0.80 to 0.93. Areas under the curve (AUC), from receiver operating characteristic (ROC) curve analysis predicting physician responses, varied from 0.821 to 0.923. Moderate association with changes in the HAQ and radiographic scores were found with correlation coefficients ranging from 0.18 to 0.63. Several studies reported mixed results between the measures when categorizing patients according to disease severity with the SDAI and CDAI the more stringent at remission.

Conclusions: The SDAI and the CDAI were found to have concurrent validity and were highly predictive of a change in therapy, but not predictive of future functional capacity or joint damage. Differences were found when categorizing patients according to disease activity level. Further studies should be conducted especially at remission and low disease activity status before these measures are used independently in a clinical setting.

Keywords: rheumatoid arthritis, disease activity indices, disease activity scores (DAS), composite indices

(2) Introduction

Over the last 15 years, the introduction of new and effective treatment agents has revolutionized the care of rheumatoid arthritis (RA) patients. Now, progression of joint damage and subsequent disability can be retarded due to aggressive pharmacological treatment with new antirheumatic drugs (57, 58). As a result of these advances in therapy, many patients may realize minimal disease activity (MDA) and even remission of the disease. However, to optimize treatment with these agents and facilitate therapeutic decision-making, accurate and consistent monitoring of disease activity is necessary in daily clinical practice (55).

In the early 1990s, the American College of Rheumatology (ACR), the European League Against Rheumatism (EULAR), and the World Health Organization/International League Against Rheumatism (WHO/ILAR) recommended a core set of individual measures to standardize quantitative measurement in clinical trials (28, 59, 60). These individual measures include: tender joint count (TJC), swollen joint count (SJC), patient assessment of pain, patient global assessment of disease activity (PGA), physician global assessment of disease activity (PhGA), measure of an acute-phase reactant (e.g., erythrocyte sedimentation rate [ESR] or C-Reactive protein [CRP]), and an assessment of physical function. Because the presentation of RA is highly variable between individual patients, a single measure cannot successfully

determine disease activity in all patients. Thus, these individual measures of disease activity were pooled into composite indices (30). The composite disease activity indices most widely used are the Disease Activity Score (DAS) (38, 61) and its modified version using a 28 joint count (DAS28) (39). In addition to assessing disease activity, determining improvement in patients in clinical trials was necessary to determine the efficacy of various treatments. The two most widely used sets of improvement criteria in clinical trials are the ACR improvement criteria (28), and the EULAR response criteria (35).

These composite indices and improvement criteria have been validated and used extensively in clinical trials, but because the DAS and DAS28 are difficult to compute, and require extensive joint counts and laboratory results that are not immediately available, they have not been found to be as useful in daily clinical practice (62). In addition, questions remain whether these standard clinical trial measures that have been effective in measuring the disease activity and response to treatment in groups of patients can now be as effective in measurement of individual patients (63).

Two new simplified composite disease activity indices that have been recently developed specifically for use in a daily clinical setting include the Simplified Disease Activity Index (SDAI) (54) and, its modified version, the Clinical Disease Activity Index (CDAI) (53). Both have been designed to provide clinicians with simpler but valid ways to assess individual patient disease activity and therapeutic response. In addition, the CDAI does not require an assessment

of an acute-phase reactant and thus can be used to measure disease activity and response to treatment in any setting (53).

In a 2005 review (44), Aletaha *et al.* reported on the validity of the SDAI and CDAI, but the findings were primarily from the two studies (53, 54) that derived the SDAI and CDAI. The purpose of this review is to examine and summarize the evidence from the body of published literature to date regarding the validity and reliability of these two new indices as evaluative measures of disease activity in individual patients and to make recommendations regarding their use in standard care.

(3) Methods

Search Strategy and Information Sources

The literature was searched using the electronic databases PubMed, Medline, Medline Plus, Embase, Cochrane Database of Systematic Reviews and Cochrane Registry of Controlled Trials. Since this review is limited to the two new measures, the SDAI and the CDAI, the literature search was performed on literature published from January, 2003 to November, 2007. Pre-determined terms were used to conduct the search and included, but were not limited to, rheumatoid arthritis, rheumatic diseases, arthritis, disease activity index, composite index, disease activity indices, and composite disease activity indices. A second search was performed adding the names of these measures and

criteria sets to the first search: DAS, DAS28, ACR Improvement criteria and the EULAR Response criteria. A third search was conducted adding the names of the two new measures, the SDAI and the CDAI. (Table 1.4)

To be included in this review, articles had to report studies where disease activity was assessed by one of the newer activity measures (SDAI or CDAI). Articles were excluded if they were reviews, meeting abstracts or not in English.

Description of the SDAI and the CDAI

The SDAI was developed in 2003 as a simpler alternative to the DAS28. The SDAI is the numerical sum of five of the core set of endpoints: swollen and tender joint count (based on a 28-joint assessment), the PGA, the PhGA, and the level of C-reactive protein (CRP).

$$SDAI = 28TJC + 28SJC + CRP + PGA + PhGA$$

CRP is a measure of the acute phase response and has been reported to be a reliable measure (49). CRP is expressed in mg/dL with a range of values from 0.1 to 10.0. The SDAI results in an absolute number to measure disease activity as does the DAS28, but the measures differ by the weightings of the individual components, the lab value and the patient self-report measure. In the SDAI, the patient self-report measure is the PGA while in the DAS28, it is defined as the global health (GH). The GH differs from the PGA in that it contains several

different health outcomes, with some not related to RA. The SDAI is easier to calculate compared to the complicated formula of the DAS28:

$$DAS28 = 0.56 * \sqrt{(28TJC)} + 0.28 * \sqrt{(28SJC)} + 0.70 * \ln [ESR] + 0.014 * PGA$$

The SDAI originally was divided into the following disease activity categories: mild (SDAI≤ 20), moderate (SDAI >=20 and <40), and high (SDAI> 40). Major improvement was represented by a drop in the SDAI of 22 and moderate improvement is a change between 10 to 22. However, new cut points for remission, low, moderate and high disease activity were recently redefined for the SDAI and are lower than those previously used (62). The newly defined cut points are: remission <3.3, low disease activity <=11; moderate disease activity <=26 and high disease activity >26, with a possible range of 0.1 to 86.

Since the SDAI and DAS28 require the measurement of a lab value, the CRP or the erythrocyte sedimentation rate (ESR), immediate assessment of disease activity may not be possible due to a delay in the return of lab results. To address this problem, the CDAI was developed in 2005 as a modified version of the SDAI, containing the same components as the SDAI, except for the CRP. The CDAI is calculated from the following formula:

$$CDAI = 28TJC + 28SJC + PGA + PhGA$$

The CDAI has been divided into the following disease activity categories: remission <=2.8; low disease activity <=10; moderate disease activity <=22; high disease activity >22 with a possible range of 0 to 76. Criteria to define major improvement based on the CDAI have not been developed.

Assessment of validity and reliability

The validity and reliability of the indices were assessed using the measurement properties proposed by the Outcomes Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiative (64). The OMERACT filter (65), a framework for validation of outcome measures, requires that a measure have truth, be discriminatory and be feasible. These measurement properties are summarized in Table 1.5.

(4) Results

The search strategy described in Table 1.4 identified 1587 titles on composite disease activity indices in rheumatoid arthritis. After including the established composite indices, 342 titles remained. After specifying that the studies had to contain one of the new indices, the SDAI or CDAI, 17 studies remained. Of these, 5 were excluded because four studies were reviews and one study was not in English.

In the 12 studies in which the SDAI and/or CDAI were used, different measurement properties to assess the measures were investigated (several publications investigated more than one property): 1). Concurrent validity: 6 publications (16, 47, 53, 54, 62, 66) (included were 2 papers from the earlier review and 4 new publications); 2). Predictive validity: 4 publications (15, 53, 54, 67) (included were 2 papers from the earlier review and 2 new publications); 3). Construct validity: 5 publications (15, 53, 54, 62, 68) (included were 2 papers from the earlier review and 3 new publications); 4). Discrimination and reliability: 9 publications (15, 16, 47, 53, 62, 66, 67, 69-71) (included was 1 paper from the earlier review and 8 new publications); 5). Responsiveness: 2 publications (16, 62) (included was 1 paper from the earlier review and 1 new publication). 6). Face and content validity: 1 publication (62). This review will address topics 1-5 from above. None of the 12 studies in this review formally evaluated the feasibility of the SDAI and CDAI in the clinical setting and this topic is not included in this review.

Concurrent Validity:

Five studies (47, 53, 54, 62, 66) reported on correlations of the SDAI with the DAS28 (Table 1.6) and one study (16) with the DAS. Two studies (53, 54) previously showed that the SDAI was highly correlated with the DAS28, one study using data from 3 RCTs (r=0.80 to 0.93) and one using an observational and inception cohorts (r=0.86 to 0.90). Three studies (47, 62, 66) confirmed

these results. Leeb *et al.* (47) reported a strong cross-sectional correlation (r=0.89) and strong longitudinal correlations (r=0.89, 0.85 and 0.84) over four time points. A second study (66), an RCT, also reported a strong correlation (r=0.91) between the SDAI and the DAS28. A third study (62) found the same result using an observational cohort (r=0.91).

The CDAI was compared to the DAS28 in two studies (53, 66)(Table 1.6). One study (53) previously showed that the CDAI was highly correlated with the DAS28 in an observational cohort and validated the result using an inception cohort. One RCT (66) confirmed these results (r=0.89).

A sixth study (16) not shown in Table 1.6, compared the SDAI with the DAS in an observational cohort of 200 early RA patients. Good correlations between the DAS and the SDAI or CDAI were found but no data was reported.

Predictive Validity:

The SDAI and the CDAI were compared with the Health Assessment Questionnaire Disability Index (HAQ), a measure of functional disability (29) in two studies (53, 54) (Table 1.6). One study (54) previously showed moderate correlations between the SDAI and the HAQ at baseline and 6 months in 3 RCTs (range r= 0.36-0.63). The same study modified the SDAI by removing the CRP value and reported moderate correlations with the DAS28 in the 3 trials (range: r= 0.47-0.56). Aletaha *et al* (53) reported moderate correlations in observational (r=0.34, 0.48) and inception cohorts (r=0.18, 0.46) between the SDAI and the

HAQ at two time points. Moderate correlations between the CDAI and the HAQ at two time points were also found in the two cohorts (observational: r=0.34, 0.49) (inception: r=0.18, 0.50).

Two other studies (15, 62) reported on the SDAI and CDAI's ability to discriminate between groups of patients with different functional status. Aletaha *et al.* (53) in an observational cohort reported that the new SDAI cut points discriminated well between groups of patients in remission and high disease activity with different HAQ scores. Khanna *et al.* (15) in an observational cohort of 200 early RA patients reported similar results when using the new SDAI cut points and the CDAI to classify patients as either in minimal disease activity (MDA) or remission. Those classified in MDA or in remission were found to have lower median HAQ status compared to groups who were not.

Construct Validity:

Moderate correlations were found in one study (53) that assessed construct validity, by comparing the SDAI and CDAI with radiographic progression. The study using an inception cohort of 91 RA patients, found statistically significant moderate correlations (SDAI: r=0.59, CDAI: r=0.54, $p<0.0001$) between the SDAI and the CDAI and the Larsen score (30).

Several studies (15, 54, 62, 68) reported that the SDAI and CDAI discriminated well between groups of patients with radiographic progression. Two studies (54, 68) used the old SDAI cut points. Smolen *et al.* (54) using data from

1839 patients from 3 RCTs, reported that changes in the SDAI response categories (major, minor or no improvement) of disease activity correlated with similar increases in the Sharp score. In a second study, (68)1004 patients in an RCT with high SDAI scores showed greater progression of joint damage when measured by the modified SHS. Two other studies (15, 62) used the new SDAI cut points. Aletaha *et al.* (62)using an inception cohort of 56 patients to validate the new SDAI cut points, reported patients classified in remission by the SDAI had significantly smaller changes in the Larsen score ($p<0.0009$) compared to those in high disease level. In a second study, Khanna *et al*. (15) reported similar results using a cohort of 200 patients with early RA. Patients classified according to new SDAI criteria and CDAI in minimal disease activity (MDA) or remission had a smaller change in Sharp scores compared to the patients not in MDA or remission.

Discrimination and reliability:

The ability of the measures to differentiate between different disease activity levels was assessed in nine studies (15, 62, 66, 69, 70) using the SDAI and in four studies using the CDAI (15, 53, 62, 69) (Table 1.7). In four studies (15, 53, 62, 69) the agreement between the classifications of individual patients by the new SDAI cut points, CDAI and the DAS28 were compared using kappa statistics (72)(31) (Table 1.7). One study (62) using an observational cohort, reported good agreement between the activity states classified by the SDAI and

the DAS28 (k=0.70). In a second study, Aletaha *et al*. (53) reported good agreement (k=0.70) between the DAS28 and the CDAI. Two other studies (15, 69) investigated agreement between the measures when classifying patients in lower disease levels. One study (69) an observational cohort of 621 patients, reported moderate to good agreement for DAS28 and SDAI or CDAI (k=0.63 and 0.58, respectively, P<0.0001) when classifying patients in remission. The study found significantly more patients in DAS28 remission had CDAI values in low disease activity rather then in remission and similar results were observed for the SDAI and the DAS28. This study also reported residual swollen joints, a major determinant of joint destruction (73, 74) were seen in 13% of patients in DAS28 remission compared to only 5% of patients in SDAI and CDAI remission. Based on the results of this study, the SDAI and CDAI are the more stringent measures compared to the DAS28 when classifying patients in remission. In a second study (15), an observational cohort of 200 early RA patients, the agreement between classifications by the DAS28, SDAI and CDAI cut points for minimal disease activity (MDA) were good (k= 0.68, k= 0.67) and the agreement for patients in remission when classifying by the DAS28, SDAI and CDAI was moderate (k= 0.48, k= 0.52).

Two studies (66, 70) compared the abilities of the SDAI and the CDAI to predict physician responses by calculating the area under the curve (AUC) using receiver operating characteristic (ROC) curve analysis (75) (Table 1.8). One study (70) reported the DAS28 had superior discriminative ability (AUC=0.840)

when compared to the SDAI and the CDAI. The second study (66) reported that the SDAI was superior to the DAS28 (AUC=0.923). Even though the study results differed, the area under the curves for the three indices ranged from 0.923 to 0.821, indicating that each index is highly predictive of a change in therapy.

In five publications (16, 47, 62, 67, 71) various other methods to assess the categorization ability of the SDAI and the CDAI were used (Table 1.9). Three of the five publications (47, 67, 71) used the old SDAI cut points, and reported differences in the categorizations of patients by the EULAR and the SDAI. In the first study (47), the distributions of the disease activity categories were not normally distributed, and a Wilcoxon's rank test was used to compare the categories. A highly significant difference was reported ($Z=-13.078$, $p<0.0001$) between the categories. The limit for the lower disease activity level of the SDAI was reduced from 20 to 10 and the data was reanalyzed. After the data was reanalyzed, insignificant differences between the EULAR and the SDAI scores were found ($p=0.07$, Wilcoxon's rank test). The second study (71), reported a higher proportion of patients was classified as having low disease activity by the SDAI criteria then the EULAR criteria. A third study (67), compared the categorization abilities of the SDAI with the EULAR and ACR response criteria sets. The study grouped the EULAR moderate and major responders into a EULAR "overall" group and grouped the SDAI minor and major responders into a SDAI "overall" group. All of those that showed an ACR 20% response were found

in the EULAR and SDAI overall groups, but only moderate agreement between the ACR50% response levels and the EULAR good level and the SDAI major level was reported. Two studies (16, 62) used the new SDAI cut points to categorize patients. One study (62) determined the ACR responder status of an inception cohort of patients and then categorized them as either responders or nonresponders, according to the ACR criteria. The proportion of patients that were classified according to the SDAI and DAS28 were then compared in the responder and nonresponder groups. At remission, the DAS28 classified more responders then the SDAI (34.9% v. 23.8%). But at high disease activity level, the classifications of the two indices were identical. A second study (16), using an observational cohort of early RA patients, reported on the degree of agreement between the measures to classify patient response utilizing EULAR criteria. In this study the DAS was used as the referent criterion and percent agreements were calculated. The percent agreements for the DAS with the SDAI and CDAI were 80%, and 74.4%, respectively. The data were re-analyzed using the DAS28 as the referent criterion and the percent agreements with the SDAI and CDAI were 73% and 74%, respectively.

Responsiveness:

The ability of a measure to detect important changes over time in response to treatment was assessed for the SDAI and the CDAI in two studies (16, 53). Both studies measured response with the ACR improvement criteria. In

one study (53), the average change in the SDAI and the CDAI where found to be greater in patients with a higher ACR improvement. A second study (16), using an observational cohort of 200 early RA patients found that the SDAI and the CDAI were both able to detect a statistically significant difference between no ACR response at baseline and an ACR 20/50/70 response at follow-up as well as between ACR50 and ACR70. However, a significant difference was not detected by the measures between ACR20 and ACR50 responders.

(5) Discussion

With advances in treatment modalities, accurate measurement of disease activity and response in individual patients in standard care is an important issue. In the present review, 12 published reports evaluated the measurement properties of two new disease activity indices, the SDAI and the CDAI to determine validity and reliability of the measures.

Evidence to support the concurrent validity of the SDAI and CDAI was found by this review. The level of correlation between the measures and DAS28 was excellent in both observational cohorts that are representative of clinical practice and in RCTs, which are highly selected populations with higher disease severity, further supporting the generalizability of the concurrent validities of the SDAI and CDAI. Recent questions have been raised about the appropriateness of the DAS28 as a measure of individual patients. A recent study (76) reported

that the DAS28 and the SDAI may not be appropriate measures of disease activity and response when assessing individual patients since both exclude the feet in their computations. The feet are often more involved then the hands and wrists especially in early RA (77). A recent study (78) refuted these findings and reported that while the assessment of feet is an important part of the clinical evaluation of patients, reduced joint counts used in the DAS28 and SDAI are appropriate and valid for disease activity assessment. Because all but one of the 12 studies in this review used the DAS28 as a referent measure when assessing both the SDAI and CDAI, uncertainty about the appropriateness of the DAS28 as a measure of individual patients has serious implications with respect to the results of the studies included in this review. Further research is needed to resolve this issue.

Little evidence was found to support the predictive or construct validities of the measures. Both of these validities are more informative then concurrent validity because they are associated with the 'real' situation in which the measure will be used. Two studies reported moderate correlations between the SDAI or CDAI when compared with the HAQ and radiographic progression. Because the correlations are only moderate and indicate variability between the measures, it is not possible to make very accurate predictions for individual patients of future functional capacity and radiographic progression using the SDAI or CDAI. Further studies in different populations should be conducted to further examine their validity.

Four studies assessed the discriminant validity of the measures when differentiating severity level between individual patients. The ability to accurately classify patients in lower disease levels especially in remission is an important issue in standard care. Differences in agreement between the SDAI, and CDAI with the DAS28 were found when classifying individual patients in remission. In addition, one study found that 13% of patients classified in DAS28 remission had more residual joint counts, a major determinant of joint destruction, compared to 5% of patients in SDAI and CDAI remission. This supports results found in a recent study (79) where patients that met DAS28 remission still exhibited synovitis after being assessed by MRI or ultrasound. It has also been shown that patients in DAS28 still exhibit synovitis in clinical examination (62, 80). Thus, the SDAI and CDAI appear to be the most specific measures compared to the DAS28 when classifying patients in remission since minimal residual joint counts were observed in SDAI or CDAI remission. When determining treatment decisions of individual patients based on classification into disease activity levels, the divergence between the discriminant ability of the DAS28, and the SDAI or CDAI should be considered.

Several studies compared the classification of patient groups by the SDAI and CDAI with the HAQ score or radiographic progression. Two studies found that groups of patients in remission when classified by the SDAI and CDAI had low median HAQ scores. Two studies reported similar results when comparing radiographic progression to the SDAI and the CDAI patient groupings. Thus, at

the group level, the SDAI and CDAI appear to accurately discriminate groups of patients when compared to the HAQ and radiographic scores. However, whether individual patients classified by the SDAI and CDAI in remission would have low HAQ or radiographic progression has not been proven by these analyses and prediction of these future outcomes based on classification into disease severity levels cannot be made.

In two studies, areas under the curve (AUC), from receiver operating characteristic (ROC) curve analysis were calculated to determine the ability of the SDAI and CDAI to predict physician response. Both measure had high AUCs indicating that they are highly predictive of a change in therapy. Additional studies should be performed to confirm these results.

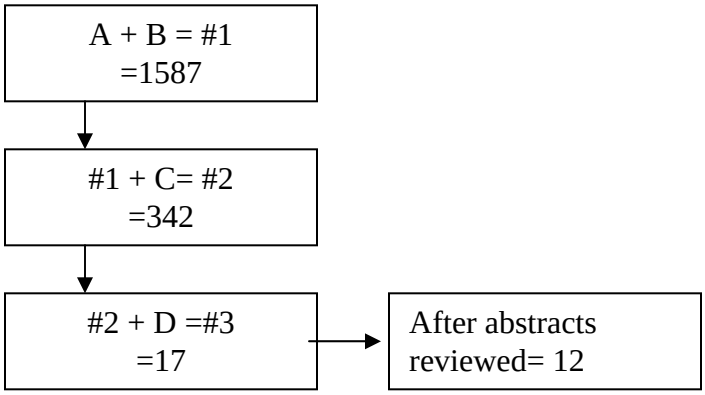
The SDAI and CDAI have met some of the criteria set forth by OMERACT but measurement of response to treatment and disease severity is complex and many issues remain to be determined. First, few studies assessed the performance of the CDAI and thus further validation of the measure, especially evidence to support the predictive and constructive validities using large observational cohorts is needed. In addition, consensus about the proposed CDAI cut points and criteria to reflect important clinical changes must be determined for use in future studies. Second, many of the studies included in this review were performed before the SDAI cut points were redefined. Only four studies were found that used the new SDAI cut points but although agreement between the ability to discriminate between patient groups of different functional

status and radiographic progression were not impacted, differences were seen between the classifications of the SDAI with the EULAR criteria depending on which cut points were used. Additional studies should be done to further validate the new SDAI criteria. Third, given that new treatment advances may help many patients attain sustained minimal disease activity or remission, it was surprising that only four studies were found that evaluated the performance of the measures in lower disease levels. Since differences were found in the discriminant ability of the measures, and none of the studies that determined correlations between the measures investigated whether the measures were correlated at lower disease levels, more studies are needed to determine which measure is most accurate and valid in low disease levels. Fourth, responsiveness of the SDAI and CDAI was evaluated in only two studies and both studies reported that the measures were able to differentiate between response categories. More studies are needed to confirm these results. Lastly, few studies evaluated the performance of the measures when assessing individual patients. Accurate assessment of disease activity in patient groups differs from the assessment of individual patients due to differences in comorbidities, treatment, and other characteristics of the individual patient. Additional studies need to evaluate the performance of the measures in individual patients.

In conclusion, this review highlights the complexity involved in determining appropriate and accurate measures of disease activity and response in standard

care. Evidence for concurrent validity of the SDAI and CDAI was found, and some evidence for responsiveness and discriminant validity when differentiating patients in different disease states and predicting change in treatment. However, divergence between the measures was reported when classifying individual patients at lower disease severity. With advances in treatments, many patients may now realize low disease activity and remission, and it will be important to have measures that are valid clinical assessment tools at these levels. This review showed little evidence to support the predictive and construct validities of the SDAI and CDAI. Studies compared patient groupings to the HAQ and radiographic progression but not individual patient classifications. More studies need to be done to evaluate the SDAI and CDAI as measures of individual patients and as accurate measures when used in remission before they can be recommended as sole quantitative measures in standard care.

Table 1.4 Search Strategy



A	B	C	D
Rheumatoid arthritis OR Arthritis OR Rheumatic diseases	Disease activity indices OR Disease Activity Index OR Composite Index OR Composite indices	ACR[a] OR ACR20[b] OR ACR50[c] OR ACR70[d] OR ACR90[e] OR EULAR[f] OR DAS[g] OR DAS28[h]	SDAI[i] OR CDAI[j]

[a]American College of Rheumatology(ACR)improvement criteria; [b]American College of Rheumatology 20% improvement criteria(ACR20); [c] American College of Rheumatology 50% (ACR50)improvement criteria; [d] American College of Rheumatology 70% (ACR70)improvement; [e]American College of Rheumatology 90% (ACR90)improvement; [f]European League Against Rheumatism(EULAR)response criteria; [g]Disease Activity Index; [h]Disease Activity Index with 28 joint count(DAS28); [i]Simplified disease Activity Index; [j]Clinical Disease Activity Index.

Table 1.5 Measurement properties of disease outcome measures

Property	Description
Truth (validity)	The ability of a measure to truly measure what it intends to measure. It includes the following domains:
Face Validity	The aggregation of the individual components in an index appears sensible.
Content Validity	The choice and importance given to the components of the index and that sample multiple domains in RA.
Criterion Validity Subtypes: Concurrent Predictive	The extent that a measure - correlates with a 'gold standard' - predicts future outcome of disease (i.e. functional disability)
Construct Validity	The ability of a measure to agree with the final outcome of the disease such as radiographic progression
Discrimination	The ability of a measure to differentiate between different disease states at one time (classification) or at different times (to measure change). It includes the following domains:
Reliability	The ability of the measure to yield the same results on repeated applications.
Responsiveness	The ability of a measure to detect important changes in the outcome after treatment has been initiated.
Feasibility	The ability of a measure to be easily applied in the setting of which it is intended.

Information from Bombardier and Tugwell (81) and Boers (82)

Table 1.6 Concurrent and predictive validities of the composite indices.

Author (Ref. no.)	Study Design	Sample Size	Sex (% female)	Correlations with DAS28[a]		Correlations with HAQ[d]		
				SDAI[b]	CDAI[c]	SDAI[b]		CDAI[c]
SMOLEN (54)	4	399	76.9	At baseline: r_{trial1}=0.91* r_{trial2}=0.91* r_{trial3}=0.80* At 6 months: r_{trial1}=0.93* r_{trial2}=0.91* r_{trail3}=0.91*		At baseline: r_{trial1}=0.46* r_{trial2}=0.44* r_{trial3}=0.36*	At 6 months: r_{trial1}=0.63* r_{trial2}=0.66* r_{trial3}=0.53*	r_{trial1}=0.56* r_{trial2}=0.56* r_{trial3}=0.47*
ALETAHA (53)	1	279	79.9	r=0.91	r=0.89	r=0.46		r=0.45
	2	105	65.2	r=0.90	r=0.89	r=0.31		r=0.30
LEEB (47)	1	767†	79.9	r=0.89				
	3	105±	65.2	r_1=0.89 r_2=0.85 r_3=0.84				
VANDER CRUYSSEN (66)	4	1242	65	r=0.91	r=0.89			
ALETAHA (62)	1	115	82.6	r= 0.91				

Study design- 1: Observational, 2: Inception, 3: Longitudinal, 4: RCT
[a]Disease Activity Index with 28 joint count(DAS28); [b]Simplified Disease Activity Index; [c]Clinical Simplified Disease Activity Index;
[d]Health Assessment Questionnaire
*Significant correlations at p<0.0001
†n=720 for correlations with HAQ
±n=104 for correlations with HAQ

Table 1.7 Agreement between the categorization ability of the composite indices*.

Author (ref. no.)	Study Design	Sample Size	Sex (% female)	SDAI[b] Cut Points	DAS28[a] v. SDAI[b]	DAS28[a] v. CDAI[c]
ALETAHA (62)	1	767	79.9	New	κ=0.70	-
ALETAHA (53)	1	767	79.9	N/A	-	κ=0.70
MIERAU (69)	1	621	78.1	New	*In remission:* κ=0.63 *	*In remission:* κ=0.58*
KHANNA (15)	1	200		New	*In remission:* κ =0.48 *In low disease activity:* κ=0.68	*In remission:* κ =0.52 *In low disease activity:* κ =0.67

* Degree of agreement indicated by kappa values from Altman (72): <0.2 = poor agreement; 0.2-0.4 = fair agreement; >0.4-0.6 = moderate agreement; >0.6-0.8 = good agreement; >0.8-1.0 = excellent agreement.
Study design- 1: Observational, 2: Inception, 3: Longitudinal, 4: RCT
[a]Disease Activity Index with 28 joint count(DAS28); [b]Simplified Disease Activity Index;
[c]Clinical Simplified Disease Activity Index;
* p=0.0001

Table 1.8 Area under the curve of the composite indices.

Author (ref. no.)	Study design	Sample Size	DAS28[a]	SDAI[b]	CDAI[c]
VANDER CRUYSSEN (66)	4	511	AUC=0.840	AUC=0.824	AUC=0.821
SOUBRIER (70)	1	204	AUC=0.872	AUC=0.923	

Study design- 1: Observational, 2: Inception, 3: Longitudinal, 4: RCT
AUC- Area under the curve
[a]Disease Activity Index with 28 joint count(DAS28); [b]Simplified Disease Activity Index;
[c]Clinical Simplified Disease Activity Index;

Table 1.9 Classification of patients by the composite indices

Author (ref. no.)	Study Design	Sample Size	Sex (% female)	SDAI[b] Cut Points	Classifications		Significance
LEEB (47)	1	399	76.9	Old	SDAI[b]: 74.4% low 21.1% moderate 4.3% high	EULAR[d]: 42.9% low 41.4% moderate 15.8% high	*p*<0.0001
LEEB (71)	1	207	75.8	Old	SDAI[b]: 95%-low 2.5% -medium 2.5%-high	EULAR[d]: 68%-low 30% -medium 2%-high	-
GÜLFE (67)	1	184	75	Old	SDAI[b]: 100%-at ACR20% level* 61%-at ACR50% level†	EULAR[d]: 100% at ACR20% level* 61% at ACR50% level†	No p-value reported
ALETAHA (62)	2	91	-	New	SDAI[b]: *In remission:* ACR20 responders- 23.8% Non-7.1% *In high disease activity:* ACR20 responders- 1.6% Non-21.4%	DAS28[a]: *In remission:* ACR20 responders-34.9% Non-10.9% *In high disease activity:* ACR20 responders-1.6% Non-21.4%	*p*<0.01
RANGANATH (16)	1	223	78.5	New	DAS[e] v. SDAI[b]: 80%** DAS28[a] v. SDAI[b]: 73%**	DAS[e] v. CDAI[c]: 74%** DAS28[a] v. CDAI[c]: 74%**	-

Study design- 1: Observational, 2: Inception, 3: Longitudinal, 4: RCT
*SDAI "overall" and EULAR "overall" response levels compared to ACR20% responders.
†SDAI "major" and EULAR "good" response levels compared to ACR20% responders.
** Per cent agreements ; [a]Disease Activity Index with 28 joint count(DAS28); [b]Simplified Disease Activity Index; [c]Clinical Simplified Disease Activity Index; [d]European League Against Rheumatology response criteria (EULAR);

1.5 Summary

For new treatment modalities such as DMARDs to be optimally effective in decreasing joint damage and destruction, clinical management of RA must focus on feasible and valid measures to facilitate the consistent monitoring of disease activity and response to treatment. In addition, feasible measures of disease activity are also needed in epidemiological research using data from large observational registries, to prevent the exclusion of patients from research studies due to missing disease activity scores.

Several studies have evaluated two new simplified composite measures, the CDAI and SDAI, and reported on their validity as measures of clinical disease activity (15, 16, 43, 45, 47, 66, 69, 70, 83, 84). Since studies assessing the performance of the composite measures using large sample sizes are lacking, a recent study by Greenburg *et al* (45) using a large observational cohort, examined the measurement agreement between the CDAI, SDAI and DAS28. New cutoff values based on this sample were determined for the CDAI and the SDAI and then compared with the DAS28. The study found the performance of the CDAI and the SDAI comparable with the DAS28 in this population, but differences between the measures remain (18). To further examine the validity of the CDAI as a clinical care measure to monitor treatment response and guide treatment decisions, in the second chapter of my thesis, I will present findings from a study where the association of the CDAI with a rheumatologist's decision to change therapy was evaluated. In addition, little previous research has been

done to examine correlates that might influence physician prescribing behavior. Thus, factors, including patient, clinic and physician characteristics, were investigated as predictors of physician's decision to change therapy using a multi-level modeling strategy.

Acute phase reactants, ESR and CRP have long been used in clinical care as measures of disease activity. However, variability in the clinical measurement of these two laboratory values has been reported (50, 51). In addition, a recent study (85) raised concerns about the appropriateness of the ESR and CRP as measures of disease activity. Thus, in the third chapter of my thesis, I will present findings from a study where correlates influencing physician clinic decision-making regarding measurement of the ESR and CRP are identified using a multi-level modeling strategy. Also the frequency of ESR and CRP testing is described and the variability in measurement of ESR and CRP is quantified.

Treatment recommendations published by the ACR in 2008 recommend measurement of disease activity to guide treatment decisions (9). However, two composite measures of disease activity currently available for use in clinical care, the DAS28 and SDAI, cannot be computed at some clinical encounters due to missing laboratory values (i.e. ESR or CRP). In epidemiological research, patients with missing values of disease activity due to missing laboratory values are often excluded from analyses. In an attempt to facilitate the inclusion of patients in epidemiological research and to improve the clinical management of

RA, in the fourth chapter of this thesis, I will present findings from a study where I developed and validated a new measure of disease activity and response, the modified DAS28 (mDAS28). A model was developed that predicted factors associated with ESR. These significant factors were then imputed into the DAS28 formula in place of the ESR value. The performance of this new measure was then evaluated by comparing its performance with the DAS28, CDAI and SDAI. As of writing this thesis, the study described in this chapter is under review by the Journal of Rheumatology. The three studies presented in my dissertation research have been reviewed and approved by the University of Massachusetts Institutional Review Board.

In conclusion, the work detailed in this thesis furthers the knowledge of measurement of disease activity in rheumatoid arthritis by determining the reliability, validity and utility of disease activity measures for use in standard clinical care settings. By validating measures of disease activity, individual patient response to effective therapeutic agents can be consistently monitored to impede further joint damage and destruction and ultimately improve patient outcomes. Results from analyses in the identification of correlates, which might influence physician practice behaviors, are also presented.

CHAPTER II

Use of CDAI and DAS28 to Inform Clinical Decision-Making In Patients with Rheumatoid Arthritis: Do They Reflect Physician Prescribing Behavior?

2.1 Abstract

Objective: To examine the association of composite disease activity indices, CDAI and DAS28, with a physician's decision to change disease-modifying antirheumatic drug (DMARD) therapy in patients with rheumatoid arthritis (RA). Also to identify patient, physician and practice characteristics associated with change in DMARD therapy.

Methods: Data were obtained from a multi-site, longitudinal, observational registry, the CORRONA registry. Initiators of a DMARD were identified and a change in therapy was assessed at each follow up visit. The association of disease activity indices and change in DMARD therapy was examined in 4,955 follow up visits using multi-level mixed multivariable models. To assess the diagnostic accuracy of the CDAI and DAS28 to discriminate a physician's decision to change DMARD therapy, areas under receiver operating characteristic (ROC) curves were employed and the predictive positive value was calculated.

Results: 4,955 follow-up visits from 1627 RA patients seen by 124 physicians at 64 clinics were eligible for analysis. Of the 4,955 follow-up visits, 14.3% (n=709) had a change in DMARD therapy with 42% had an escalation and 48% had a discontinuation. In separate models, the CDAI (OR=1.58; 95% CI: 1.42,1.76) and

DAS28 (OR=1.54; 95% CI: 1.35, 1,74) were both significantly associated with change in DMARD therapy after controlling for visit-, patient-, physician, and clinic-level characteristics. However, a change in CDAI core was not significantly related. Other factors significantly associated with change in DMARD therapy included: initiation of a nonbiologic DMARD (OR=2.67; 95% CI: 1.78, 4.00), number of prior prescribed DMARDs (OR=1.69; 95% CI: 1.47, 1.95), patient age (OR=0.99; 95% CI: 0.98-1.01), duration of disease (OR=0.66; 95% CI: 0.58,0.75), currently employed (OR=1.66; 95% CI: 1.10,2.49), physician age (OR=0.97; 95% CI: 0.95-1.00) and academic clinic (OR=1.92; 95% CI: 1.08,3.40). The CDAI and the DAS28 were not strong predictors of a change in DMARD therapy as determined from area under the receiver operating characteristic curve analysis (AUC of 0.69 and 0.72, respectively). The effect of patient characteristics on change in DMARD therapy was substantial.

Conclusion: A physician's decision to change DMARD therapy was based in part but not completely on disease activity. The CDAI and DAS28 were not strong discriminators of a change in DMARD therapy and should not be used to guide treatment decisions in a clinic setting.

Key words: Rheumatoid arthritis, Clinical Disease Activity Index, Disease Activity Score, Multi-level Mixed models, Receiver operating characteristic (ROC) curves, Discriminating ability

2.2 Introduction

Rheumatoid arthritis (RA) is a complex disease with a wide array of disease manifestations causing disease management to be difficult (1, 2). Previous treatment recommendations published in 2002 (3) by the American College of Rheumatology (ACR) focused on the use of nonbiologic disease-modifying antirheumatic drugs (DMARDs). In 2008, the ACR recently updated their recommendations, to address the rising rate of biologic DMARD use and suggested aggressive treatment with DMARDs to attenuate the risk for joint destruction and resulting disability (4).

Little research has been done to determine whether physicians are adhering to these treatment recommendations. Physician prescribing behavior may be influenced by many variables such as other physician practice behaviors, financial incentives, restrictions on treatments, factors unique to practice settings, cost and insurance, individual physician characteristics and individual patient characteristics, but specific factors have not been illuminated in the literature (5-8).

To facilitate physician clinical decision-making, the ACR has recommended several clinical factors to consider when making a therapy change: disease duration, prognostic factors and disease activity assessment (4). To assess disease activity, many individual measures and composite indices are available (9). One recent RCT, the TICORA study (10)) demonstrated that

aggressive therapy and consistent monitoring of disease activity with a composite index, the disease activity score (DAS) (11, 12), resulted in improved outcomes. Composite measures such as the clinical disease activity index (CDAI) (13) and the disease activity score with 28 joint count (DAS28) (14), if able to evaluate the effect of DMARD therapy and remaining disease activity in daily practice could be used to guide treatment and ultimately improve disease outcome. While validated in several studies (15-30), the ability of the CDAI and DAS28 to guide treatment decisions, to the best of my knowledge, has not been previously examined.

Using data from a large, multi-site, observational registry, we evaluated whether a clinical decision by a physician to change DMARD therapy could be reflected by the CDAI or DAS28. We also sought to evaluate the effect on RA management of clinic, physician and patient-level characteristics. In addition, the discriminative ability of both measures to determine treatment change was evaluated by calculating the predictive probability value. Areas under the receiver operating characteristic curves (AUC) (31-34) were calculated to assess the fit of the models and the diagnostic ability of the measures.

2.3 Methods

Data Source

Data for the analysis were from a large, multi-site, observational registry, the Consortium of Rheumatology Researchers of North America (CORRONA)

(35). The CORRONA registry contained information for more than 17,000 patients with RA, 113 rheumatology practices, and 272 rheumatologists in the United States. Information was collected on patient and physician demographics, disease activity and severity, quality of life measures, medical comorbidities, and use of DMARD drug therapy. The data are obtained from patients and their treating rheumatologist using study questionnaires. Follow-up exam visits are scheduled at three month intervals and completed during routine clinical encounters.

Study Sample

Subjects who had RA and had at least one visit where a DMARD had been newly prescribed during the study period between October 1, 2001 and November 1, 2008 were identified (n=3,172). Since some of these 3,172 patients had switched DMARD therapy and initiated another DMARD at different visits, they had more than one visit with an initiation. Thus, there were a total of 6,335 initiation visits. The baseline visit was defined as the visit where a DMARD was first initiated. DMARD initiations that occurred between two scheduled visits, and had a prior visit within 4 months of the initiation, were included in the sample. Information from the prior visit was used as the baseline visit. A 4-month interval was decided upon because most DMARD therapies do not show an effect until 3-6 months after initiation. Patient and physician information from the prior visit was utilized as the baseline visit. A total of 2,352 initiation visits were excluded

because they occurred between two visits and had a prior visit more than 4 months of the initiation, leaving 3,983 visits with an initiation (baseline). A baseline visit where an initiation had occurred had to have at least one follow up visit to be included in the sample. A total of 716 initiation visits were excluded from further analysis because they did not have at least one follow up visit, leaving 3267 initiation visits. The initiation baseline visit and its accompanying follow up visits comprised an “initiation segment “. An additional 211 initiation segments or 1356 visits were excluded due to missing CDAI scores at the baseline visit leaving 3,056 initiation segments.

In the analysis, DMARD initiators were followed over time until the outcome, change in DMARD therapy (defined in next section), occurred at a follow up visit or if no change in DMARD therapy occurred during the initiation segment, then until the last follow up visit of the initiation segment. If a change in DMARD therapy occurred at a follow up visit, these follow up visits along with any follow up visits in an initiation segment that occurred prior to the change in DMARD therapy were included in the sample. Follow up visits where a change in DMARD therapy occurred prior (or between) the visit, were excluded from the sample because patient and physician information was not collected between regular visits. Therefore, of the remaining 3056 initiation segments, 1429 initiation segments were excluded from the analysis because a change in therapy occurred between visits.

The final sample contained 4,955 follow up visits within 1627 initiation segments. A sensitivity analysis was done to test any differences between the excluded visits and the ones included. The institutional review board of the University of Massachusetts Medical School approved the study. Since the study was a secondary analysis of existing data, the need for informed consent was waived.

Outcomes of interest

The primary outcome of interest was the physician's decision to change DMARD therapy and was defined as any of the following having occurred at a follow up visit: (a). an addition of DMARD, (b). a discontinuation of baseline DMARD, and (c). a discontinuation of baseline DMARD and new DMARD added. The outcome was operationalized as a dichotomous variable, coded as '0' for no change and '1' for change (either a-c above). Individual components "a" and "b" (above) were also examined as separate outcomes in subsequent analysis and operationalized as dichotomous variables. Preliminary data analysis revealed that component "c" occurred in only 18 follow up visits. Thus, due to this low amount of occurrences, component "c" was not examined as a separate outcome variable. DMARD medications included in the analysis were 6 biologics: etanercept, adalimumab, anakinra, abatacept, infliximab, rituximab, and 7 nonbiologics: methotreexate, leflunomide, sulfasalazine, cyclosporine, azathioprine, hydroxychoroquine, and minocycline.

Independent variables

Data were organized into four levels to reflect the natural hierarchy of the data structure: visit-, patient-, physician- and clinic-level characteristics. The study utilized multi-level logistic regression mixed models to account for the correlation between the observations (visits), and the clustering of the data at the patient, physician and clinic levels. Multi-level modeling allows the simultaneous appraisal of group- and individual-level factors on individual-level outcomes (36). In the study sample, 11% of the clinics and 29% of the physicians reported no changes made in DMARD therapy. Thus we knew *a priori* that the physicians and clinics prescribing behavior varied widely.

Visit-level factors*:* The effect of the CDAI on the prescription of DMARD therapy was assessed and the effect of its individual components, tender joint count (TJC; 0-28), swollen joint count (SJC; 0-28), physician's global assessment of disease activity (PhGA; visual analog scale (VAS) 0-100 mm), and patient's global assessment of disease activity (PGA; VAS 0-100 mm) were assessed in separate models. The effect of the DAS28 was assessed in separate sub-analysis. The erythrocyte sedimentation rate (ESR; mm/h) was measured to calculate the DAS28. The CDAI and DAS28 were calculated as previously described (11, 13). The effect of other disease activity variables on rheumatologist's decision to change DMARD therapy included C-reactive protein (CRP; mg/l), modified health assessment questionnaire (mHAQ; 0-3) (37, 38),

patient assessment of pain (PAIN; VAS 0-100mm), RF positivity (RFpos), presence of erosions and joint narrowing, and morning stiffness. CDAI and DAS28 were calculated after data collection occurred so that the treating rheumatologist was not able to utilize either score to make a decision regarding a therapy change. Therapy characteristics at each follow up visit were also examined and included type of baseline initiation DMARD (i.e. biologic or nonbiologic), number of previous DMARDS prescribed, use of prednisone, and use of other therapies such as gastrointestinal medications, analgesics and antidepressants. Calendar year was included as control variable since patients may have experienced different recommendations for therapy based on a particular year.

Changes over time of the composite scores and disease activity variables were also considered. Change scores were derived by calculating the change in disease activity from the baseline initiation visit to the next follow up visit. For example, if a patient had a follow up visit at 3 months and then at 6 months, the change scores were calculated as the difference between baseline and 3 months and then the difference between baseline and 6 months.

Patient-level factors: included patient demographic variables: patient age at baseline which was operationalized as a continuous variable, patient race dichotomized as white or not white, gender defined as male or female, education level which was dichotomized as having either completed college or no college,

marital status which was dichotomized into two levels, married or not married, employment status was dichotomized as either currently employed or not employed, and body mass index (BMI) was utilized as a continuous variable. Insurance status was assessed using two dichotomous variables: private insurance and public insurance. Public insurance include both Medicare and Medicaid and both insurance variables were dichotomized as “0” for no insurance and “1” for insurance. Duration of RA was a continuous variable. Comorbidities included seven conditions (hypertension, coronary artery disease, myocardial infarction, congestive heart failure, stroke, diabetes mellitus, and cancer) and were dichotomized into two levels, either having the disease or not.

Physician-level factors: Factors at the physician level included physician gender, defined as male or female, physician age and practice years. The frequency of formulary restrictions and its influence on physician decisions was included in this level. This variable was operationalized as dichotomous with frequent formulary restrictions coded as 0 and infrequent coded as 1. Physicians also reported on the percentage of managed care in their practices. The variable was operationalized as a categorical variable with three categories: <25%, 25-50% or >50%. A variable ”prescribing tendency” to assess the past tendency of an individual physician to prescribe DMARD therapy was derived. The variable was derived by averaging the number of previous DMARDs prescribed by a physician per their total number of patients.

Clinic-level factors: Clinic level variables included in model selection were: type of clinic (academic or private) and the geographic region in which the clinic was located. Geographic location was defined according to four regions: the Northeast, Midwest, South and the West according to the Geography Division of the U.S Census Bureau (39).

Statistical analysis

Descriptive statistics are presented as means and standard deviations (SD) for continuous data or as percentages for counts for the characteristics of the follow-up visits. In this study, the unit of analysis is the follow-up visit and the proportion of follow-up visits with a DMARD change was tabulated. To account for the correlation and clustering between levels of the data, univariate multi-level logistic mixed-modeling analysis was used to examine visit, patient, physician and clinic characteristics as potential predictors of a change in DMARD therapy. Variables that had univariate associations of $P<0.10$ and/or theoretically relevant were included in subsequent multi-level multivariate mixed-modeling. Correlation among the variables was examined using the variable inflation factor (VIF) (40). Variables found to have a VIF >2 were excluded from further modeling. The assumption of linearity of continuous variables with the endpoint was tested using two way plots with loess curves. Variables not linear were transformed accordingly.

In the multilevel multivariable mixed models, concurrent values of the CDAI at each follow up visit and significant visit, patient, physician and practice-site characteristics were evaluated as potential predictors of a change in DMARD therapy. This approach accounts for the clustering of visits within patients and patients within physicians. Using a single-level logistic regression model would ignore clustering of the patients in physicians, and physicians in clinics and may exaggerate the precision of the estimates. Each aspect of the primary outcome, escalation of DMARD therapy and discontinuation, were examined in separate analyses. The effect of the DAS28 on the physician's decision to change therapy was estimated in a separate multi-level multivariate model. The CDAI and DAS28 scores were standardized for ease of parameter interpretation in the models. Standardization of the CDAI occurred by dividing the CDAI score by 6.7, the value of a statistically significant change for the CDAI previously determined (27).The DAS28 was standardized by dividing the DAS28 score by 0.6, the value of a statistically significant change for the DAS28 previously determined (41). All continuous variables were centered and physician global assessment, patient global assessment, and patient pain scores were presented by decile to aid in interpretation of the results.

The fit and performance of the models to discriminate a change in DMARD therapy was assessed using areas under the receiver operating characteristic (ROC) curve analysis (32, 33). To build the models, the sample was divided into a random training (n=2499) and test (n=2456) datasets. The

models were developed in the training set and the test set was used to validate the model by making predictions from the data. The data predictions were created using ROC analysis and then compared with the observed data using the kappa statistic (42, 43). ROC curves were plotted for each model to determine the areas under the curves (AUC), sensitivity and specificity of each model. AUC values >0.75 indicate good discriminative ability (31). Cut points for the CDAI to delineate between change in DMARD therapy and no change were determined using ROC analysis and the Youden index (44, 45). The performance of the new cut point was compared to the actual observed change values using kappa statistics. Positive predictive value of the new test cut point was determined. To quantify the proportion of residual variance attributable to variations between units of the 4 levels of hierarchy, random intercept variance, intra-cluster correlation coefficient (ICC) (46) and the median odds ratio (MOR) (46) were calculated. Residual variance is the proportion of unexplained variation in outcomes due to unmeasured characteristics at each of the 4 levels. All analyses were conducted using STATA, release 10.1 (47).

2.4 Results

A total of 4,955 follow up visits (level 1), were nested within 1,627 RA patients (level 2), and received care from 124 physicians (level 3), from 64 clinics (level 4). Baseline DMARD initiators (n=1627) had a mean CDAI score of 19.34,

had an average age of 56.8 (SD), had RA for an average of 9.92 years, 78% were female, most had insurance, and the number of past DMARDS was 1.68. Slightly more patients had initiated a biologic DMARD (n=873; 54%) at baseline, compared to a non-biologic (n=754; 46%).

As shown in Table 2.1, a total of 4,955 follow up visits were eligible for analysis. Disease activity at the visit-level was moderate with a mean CDAI score of 11.8 and a mean DAS28 score of 3.28. Moderate disease also was reflected in other disease activity measures (Table 2.1). Most patients were female and most were Caucasian. The mean age of patients was 58 years. The average percent of visits with comorbid conditions was 7% and the average number of previous DMARDS prescribed was 1.69. The mean age of the 124 rheumatologists was 56; 84% were men and had an average of 24 years in practice. Most of the 64 practice sites were private (89%) compared to academic (11%). Most clinics were located in the Northeast (52%) with only 4.5% in the West

The proportion of follow-up visits with a change in DMARD therapy was 709 (14.3%) compared to 4262 (85%) visits with no DMARD therapy change (Table 2.2). Of these 709 visits, 483 (60%) had an escalation of DMARD therapy, 208 (30%) had a discontinuation of a baseline DMARD and 18 (7%) had a discontinuation of a baseline DMARD and another DMARD prescribed.

In univariate analysis presented in Table 2.3, variables significantly associated with the decision to change DMARD therapy at a $P \leq 0.10$ level are

presented. Variables are grouped in the four hierarchical levels of the data listed earlier. Several quantitative disease activity measures were found to be significantly associated with a physician's decision to change therapy and included the CDAI, DAS28, ESR, CRP, mHAQ, and pain score. The values of each quantitative score including the CDAI and DAS28 are concurrent values from each follow up visit. The CDAI and DAS28 were standardized and thus a change of 6.7 units in a CDAI score increases the probability by almost 60% that a physician will change DMARD therapy. Similarly for the DAS28, a change in 0.6 in the DAS28 score increases the likelihood of a DMARD therapy change by almost 50%. Change in CDAI score and a change in DAS28 score were also found significantly associated with a physician's decision to change DMARD therapy. Univariate results for the second and third outcomes of interest, escalation of therapy and discontinuation were similar to the results for outcome one and are not presented.

In adjusted analysis, the parameter estimates for all fixed effects of CDAI and the visit, patient, physician and clinic on the physician's decision to change DMARD therapy are presented in Table 2.4. A change in 6.7 units of a CDAI score significantly increased the probability that the physician would change DMARD therapy by almost 60%. A baseline initiation of a nonbiologic DMARD and the number of past DMARDs prescribed at a clinical visit also significantly increased the probability that a physician would change DMARD therapy. A patient's age and duration of disease were both inversely related to a change in

DMARD therapy. Patients who were currently employed were almost 70% likely to receive a change in DMARD therapy. As a physician aged, the probability of a change in DMARD was decreased. Academic clinics compared to private had almost a 2-fold increase in probability of changing DMARD therapy. A change in CDAI score was not found to be significantly related to a physician's decision to change therapy after controlling for fixed and random effects (OR=1.01; 95% CI:0.92-1.11). In a separate adjusted analysis, the DAS28 was found to be significantly associated with the physician's decision to change therapy (OR=1.54; 95% CI: 1.35, 1.74). Multivariate analysis of the second and third outcomes, escalation and discontinuation of DMARD therapy showed similar results and are not shown.

When assessing the fit of the CDAI model (Table 2.4), the full model including fixed and random effects performed well to discriminate physician decision to change DMARD therapy (AUC=0.96). The discriminatory ability of both the CDAI and DAS28 to delineate change in DMARD therapy was similar with an AUC of 0.69 for CDAI and 0.72 for DAS28. The discriminative ability of the CDAI was improved when random physician and patient effects were accounted for in the model (AUC= 0.91) but was weaker when not including random effects (0.69). Accounting for unobserved patient and physician random effects played a role in increasing the predictiveness of the CDAI. The CDAI cut-off value for change in DMARD therapy with the highest combination of specificity and sensitivity was 14.1. When the performance of the newly derived

cut point was compared to actual physician behavior, analysis by kappa statistics showed moderate agreement. The positive predictive value of the CDAI was 41% and was calculated by taking into account the low prevalence of physician change in DMARD therapy in the sample population (14%).

The proportion of the unexplained variation in the primary outcome due to unmeasured characteristics at each of the four levels of hierarchy was determined and is presented in Table 2.5. There was strong and significant effect of clustering found at the individual patient level. In model 1 (empty model), there was variation in the physicians' decision to change DMARD therapy across physicians (τ=0.69) and across patients (τ=1.33). Yet, little variation was found across clinics ($\tau=1.8e^{-10}$). Intra-class correlation coefficients (ICC) for each level of data support these findings with 0%, 12.9% and 25% variation in the change of DMARD therapy, attributed to clinic, physician and patient level factors, respectively. However, results from the final four level model (Table 2.5, Model 5) after accounting for residual variation at the visit-, patient-, physician- and clinic-level factors, the effect was significantly attenuated at the patient level but still a significant effect. Median odds ratios (MOR), presented in table 5 for each of the models, confirmed that patient characteristics influenced the decision to change DMARD therapy. The high MOR values in all of the models at the patient level suggest that the patient heterogeneity in the decision to change DMARD therapy is substantial.

In other analyses, differences in disease activity over time and response scores were examined to discriminate the rheumatologist's decision. Those analyses indicated that change variables seem to be less important than disease activity measured concurrently at a follow up visit to discriminate the rheumatologist decision.

2.5 Discussion

The aim of the present analysis was to evaluate whether a composite index such as the CDAI or DAS28 could discriminate the treating physician's decision to change DMARD therapy in RA patients. Using data from the large, multi-site longitudinal CORRONA registry, the results of the present study indicate that the CDAI and DAS28 both were significantly associated with change in DMARD therapy. Yet, the CDAI and DAS28 were found to be poor predictors of change in DMARD therapy as indicated by a low AUC and PPV. The discriminative power of the CDAI without controlling for random effects was much lower compared to the final model which included the random effects. Thus, based on these results, the CDAI and DAS28 would not be good discriminators of a change in DMARD therapy in the clinic.

Follow up visit characteristics that were found to be predictive of a decision to change DMARD therapy included an initiation of a nonbiologic DMARD at baseline and the number of past DMARDS prescribed at the follow up visit. With an initiation of a nonbiologic DMARD, results of the study indicated

that a physician would be almost 3 times more likely to change DMARD therapy. Biologic DMARD therapy has been found to be more effective then nonbiologics and ACR treatment recommendations do suggest aggressive DMARD therapy to prevent disease progression (4), factors which provide support for our results. The number of past DMARDs prescribed increased the likelihood of a change in DMARD therapy by almost 70%. Prior research has reported that frequent switching of DMARD therapy results in improved outcomes (48). This finding may indicate that physicians are switching DMARD therapy to achieve a treatment response.

Patient characteristics significantly predictive of a DMARD therapy change included patient age, disease duration and employment status. Patient age was inversely related to a change in therapy, a finding supported by previous research which found that older patients with RA receive less switches in DMARD therapy. Also duration of disease was inversely related to change in DMARD therapy. Patients who were currently employed compared to those not employed had an almost 70% likelihood of receiving a DMARD therapy change. This finding could be due to differences in insurance coverage. Further research needs to be conducted to determine the reasons for this finding.

Only one physician characteristic was predictive of change in DMARD therapy. Physician age was inversely related to change in DMARD therapy with 3% less chance of a therapy change for every year increase in physician age. This finding may be related to physician training and practice styles.

Type of practice environment was found to influence DMARD therapy change with academic clinics strongly predictive of a change in DMARD therapy. Physicians at academic clinics may be involved in research efforts and, thus may be more likely to switch therapies. Patient populations that visit academic clinics may differ from those at private clinics and may change the prescribing tendencies of physicians.

Variation in the physicians' decision to change DMARD therapy was found to be strongly influences by patient characteristics with little variation due to clinic or physicians. Thus interventions to improve adherence of physician prescribing of DMARDs should target patient levels characteristics and not the individual physician or clinic where little variation was found.

This study has several strengths. The data source utilized for the study is a large, observational cohort of patients with RA form across the United States. Thus the results from the study are generalizable throughout the U. S. Second, the study utilized a training dataset to develop the models and a test dataset to test the models for fit and discriminative ability which adds to the robustness of the results. Third, the study utilized a multi-level mixed effects modeling strategy to account for correlation between individual visits and clustering effects at each level of the data. The study is limited by not having data on several variables which could impact a rheumatologist's decision-making process such as case-mix of the patient population or quality assurance programs that had been implemented which may result in a loss of heterogeneity in the use of therapies

for patients. Since we did not have this information, we were unable to control for these differences between patients and clinics.

In summary, disease activity as measured by the CDAI and the DAS28 was found to be significantly associated with a physician's decision to change DMARD therapy. Change in disease activity over time was not found to be associated with a change in therapy. The CDAI and DAS28 were not found to be strong predictors of a change in therapy and thus based on these results should be used in clinic care to guide treatment decisions. To our knowledge, this is the first large scale study which demonstrated the association of the CDAI and DAS28 with physicians' decision to change DMARD therapy. These results should be verified in further studies, but using the CDAI and the DAS28 in the clinical setting to guide treatment appears not to be warranted.

Table 2.1 Visit, patient, physician and clinic characteristics of follow up visits (n=4955)

Characteristics	N	Mean (SD)
Visit-level (n=4,955)		
CDAI, mean (SD)	4728	11.82 (10.89)
DAS28, mean (SD)	1982	3.28 (1.44)
TJC, mean (SD)	4935	2.86 (5.02)
SJC, mean (SD)	4936	4.24 (5.91)
ESR, mean (SD)	2078	22.08 (20.20)
CRP, mean (SD)	1688	2.91 (8.09)
PhGA, mean (SD)	4948	18.17 (16.80)
PGA, mean (SD)	4747	29.67 (24.42)
Pain, mean (SD)	4812	31.26 (24.93)
mHAQ, mean (SD)	4847	0.40 (0.45)
Number of past DMARDs	4955	1.69 (1.67)
Corticosteroids (%)	4939	31.0 (0.42)
Non-steroidal (%)	4955	59.0 (0.30)
Antidepressants (%)	4955	29.0 (0.45)
Baseline DMARD initiation (n=1627)		
Nonbiologic¥	754	1.35 (0.23)
Biologic*	873	1.67 (0.34)
Patient-level (n=1627)		
Age, mean (SD)	4944	58.80 (12.54)
Caucasian (%)	4901	87.0 (0.34)
Female (%)	4955	76.0 (0.43)
BMI (%)	4926	29.3 (7.15)
RA duration, yrs , mean (SD)	4949	11.37 (9.61)

Currently working (%)	4955	43.0 (0.31)
Private insurance (%)	3851	78.0 (0.42)
Comorbidities ± (%)	4955	7.0 (0.24)
Physician-level (n=124)		
Male gender (%)	4837	84.0 (0.37)
Age, mean (SD)	4674	56.2 (6.51)
Years in practice, mean (SD)	4777	24.5 (7.20)
% of managed care		
<25%	1184	23.9 (0.43)
25-50	1044	21.1 (0.41)
>50	2493	50.3 (0.50)
Formulary restrictions		
Frequent	2973	60.0 (0.49)
Infrequent	1830	37.0 (0.48)
Prescribing tendency	4955	1.54 (1.01)
Clinic-level (n=64)		
Academic, (%)	4764	11.0 (0.32)
Region, mean (SD)		
Northeast	2510	51.0 (0.50)
Midwest	1450	29.26 (0.46)
South	623	12.57 (0.33)
West	217	4.40 (0.20)

CDAI- clinical disease activity index; DAS28-disease activity score with 28 joint count; TJC-tender joint count with 28 count; SJC-swollen joint count with 28 count; ESR- ethrocyte sedimentation rate; CRP-C-reactive protein; PhGA-physician global assessment; PGA- patient global assessment; mHAQ-modified Health Assessment Questionnaire; BMI-body mass index; RA-Rheumatoid arthritis

±comorbidities includes hypertension, coronary artery disease, myocardial infarction, congestive heart failure, stroke, diabetes mellitus, and cancer.

¥nonbiologics includes: methotreexate, leflunomide, sulfasalazine, cyclosporine, azathioprine, hydroxychoroquine, and minocycline

*****biologics includes: etanercept, adalimumab, anakinra, abatacept, infliximab, rituximab

Table 2.2 Proportions of changes in DMARD therapy at follow up visits (n=4,955)

Changes in DMARD therapy	Number of changes
Any DMARD change	709
DMARD discontinuation	208
DMARD addition	483
DMARD Discontinuation and then addition	18
No change	4262

DMARD- disease-modifying antirheumatic drug

Table 2.3 Univariate analysis of a rheumatologist's decision to change DMARD therapy at a follow up visit (training dataset n=2499)

Variable	Odds Ratio (95% CI)	*P* Value
Level 1 : Clinical encounter		
CDAI	1.59 (1.43-1.75)	0.00
Change in CDAI	1.22 (1.13-1.31)	0.00
DAS28	1.46 (1.44-1.47)	0.00
Change in DAS28	1.14 (1.09-1.26)	0.00
ESR	0.96 (0.94-0.98)	0.00
CRP	0.94 (0.91-0.97)	0.00
mHAQ	0.67 (0.53-0.95)	0.03
Pain Score	1.11 (1.05-1.54)	0.05
Nonbiologic baseline initiation¥	2.54 (2.13-2.75)	0.00
No. of past DMARDs	1.68 (1.54-1.85)	0.00
Level 2: Patient		
Patient age	0.98 (0.97-0.99)	0.01
Female	1.05 (0.99-1.10)	0.00
White	0.84 (0.78-0.89)	0.03
BMI	1.00 (0.99-1.004)	0.00
Duration of disease	0.98 (0.95-0.990	0.00
Currently employed	0.95 (0.91-0.99)	0.05
Private insurance	1.05 (1.00-1.09)	0.06
Comorbidities±	0.87 (0.58-0.76)	0.00
Level 3: Physician		
Physician age	0.97 (0.97-0.98)	0.00
Male	0.83 (0.78-0.88)	0.00
Propensity to prescribes	1.52(1.24-1.65)	0.00
Level 4: Clinic		
Academic	1.56 (1.23-1.76)	0.00
Geographic region		
Northeast	1.00	
Midwest	1.34 (1.27-1.43)	0.02
South	0.66 (0.62-0.71)	0.02
West	1.09 (1.01-1.19)	0.01

CDAI- clinical disease activity index; DAS28-disease activity score with 28 joint count; ESR- erythrocyte sedimentation rate; CRP-C-reactive protein; mHAQ-modified Health Assessment Questionnaire; BMI-body mass index
±comorbidities includes hypertension, coronary artery disease, myocardial infarction, congestive heart failure, stroke, diabetes mellitus, and cancer.
¥nonbiologics includes: methotreexate, leflunomide, sulfasalazine, cyclosporine, azathioprine, hydroxychoroquine, and minocycline

Table 2.4 Multivariate analysis of visit, patient, physician and clinic variables associated with a DMARD change (training dataset n=2499)

Variable	Odds Ratio (95% CI)	*P* Value
CDAI	1.58 (1.42-1.76)	0.00
Nonbiologic baseline initiation¥	2.67 (1.78-4.00)	0.00
No. of past DMARDS	1.69 (1.47-1.95)	0.00
Patient age	0.99 (0.98-1.01)	0.00
Duration of disease	0.66 0.58-0.75)	0.03
Currently employed	1.66 (1.10-2.49)	0.01
Physician age	0.97 (0.95-1.00)	0.00
Academic clinic	1.92 (1.08-3.40)	0.00

DMARDs-disease modifying antirheumatic drugs; CDAI- clinical disease activity index;
¥nonbiologics includes: methotrexate, leflunomide, sulfasalazine, cyclosporine, azathioprine, hydroxychoroquine, and minocycline

Table 2.5 Sources of variation among rheumatologists to change a DMARD (training dataset: n=2499)

Measures of variation	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]	Model 5[e]
Clinic level					
Variance (SE)	$1.8e^{-10}$ (0.21)	$2.15\ e^{-9}$ (0.12)	$3.68e^{-9}$ (0.36)	$7.07e^{-12}$ (0.33)	$9.22e^{-11}$ (0.18)
ICC (%)	0.00	0.00	0.00	0.00	0.00
MOR	1.00	1.00	1.00	1.00	1.00
Physician level					
Variance (SE)	0.69 (0.11)	0.0001 (0.35)	$6.53e^{-9}$ (0.60)	$1.12e^{-11}$ (0.56)	$2.89e^{-11}$(0.27)
ICC (%)	12.9	0.00002	0.00	0.00	0.00
MOR	2.21	1.01	1.00	1.00	1.00
Patient level					
Variance (SE)	1.33 (0.11)	1.89 (0.71)	9.26 (2.2)	7.98 (2.30)	1.87 (0.86)
ICC (%)	25.0	36.5	74.0	71.0	36.0
MOR	3.00	3.71	18.22	14.8	3.69

DMARD- disease-modifying antirheumatic drug; SE- standard error; ICC- intra-class correlation; MOR- median odds ratio

[a] Model 1 is the null model containing no predictor variables

[b] Model 2 contains visit variables

[c] Model 3 contains visit and patient-level variables

[d] Model 4 contains visit, patient and physician-level variables

[e] Model 5 contains visit, patient, physician and clinic-level variables

CHAPTER III

Clinical Measurement of Erythrocyte Sedimentation Rate and C - reactive Protein in Patients with Rheumatoid Arthritis Is Not Associated with Disease Activity

3.1 Abstract

Objective: To determine the frequency of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) measurement, identify predictors of ESR and CRP measurement and quantify the variation among rheumatologists in their measurement.

Methods: Frequency of ESR and CRP measurement was tabulated. The study utilized a multi-level mixed-effects logistic regression analysis to identify significant predictors of ESR and CRP measurement at clinical encounters, patient, physician and clinic levels. The fit of the final model was examined using receiver operator characteristic (ROC) curve analysis. Unmeasured variation attributable to each cluster of data was quantified using intra-class coefficients and median odds ratios.

Results: There were 92,062 clinical encounters (level 1), nested within 17,450 patients (level 2) who were patients of 272 physicians (level 3) from a total of 100 clinics (level 4). Significant predictors included ordering other laboratory tests (OR:2.61; 95% CI: 2.3,2.9), number of previous DMARD prescriptions (OR:0.96; 95% CI: 0.94,0.98) , female gender (OR: 1.09; 95% CI: 1.00,1.18), disease duration (OR: 0.99; 95% CI:0.992,0.999) , no insurance (OR: 0.37; 95% CI: 0.27,0.49), physician propensity to order either ESR or CRP (OR: 1.46; 95%

CI:1.34,1.59) and private clinics (OR: 0.40; 95% CI: 0.20,0.81). Quantitative disease activity measures were not significantly predictive of ESR or CRP measurement. ROC curve analysis indicated a substantial fit of the final model (AUC=0.871). Variation in the measurement of an ESR and CRP was attributable to patient and clinic characteristics and not the individual physician.

Conclusion: ESR and CRP measurement was not associated with disease activity levels. The use of ESR and CRP tests in the clinical management of RA should be reevaluated.

Key words: ESR, CRP, Acute phase reactant, DAS28, disease activity measures

3.2 Introduction

Evaluative laboratory tests, such as acute phase reactants (APRs) – erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), have long been used in the clinical management of RA to assess disease activity and guide treatment (1-4). They are part of the American College of Rheumatology (ACR) core data set (5) for measuring disease activity and are used in the calculation of three composite disease activity measures: the disease activity score (6, 7)(DAS) and its modified version with 28 joint count (DAS28) (8) and the simplified disease activity index (SDAI) (9). They are also part of criteria to assess response to treatment (10-12).

With the advent of new therapeutic biologic agents, assessment of disease activity to guide treatment decisions was recently recommended by the ACR (13). Variability in the use of laboratory tests by rheumatologist's in the United States (U.S.), such as the ESR and CRP, to assess disease activity in clinic care has been reported in prior studies (14, 15). And, results of the two tests are often not available to the clinician at the time of the clinical encounter, thus, preventing the assessment of disease activity and the calculation of the DAS28 and SDAI and their associated response criteria (16).

Only two studies (14, 15) were found that have examined the level of variation between rheumatologists in the measurement of ESR and CRP or the factors associated with their measurement. Henke *et al* (14) found, in a sample

of 66 U.S. rheumatologists', wide variation in the use of ESR to monitor patients with RA. Forty-three per cent never ordered an ESR test while 3% were found to order an ESR test at every clinical encounter. In contrast, Donald *et al* (15) in a survey of 575 US rheumatologists found that 43% reported ESR and/or CRP testing at every visit or almost every visit. Henke *et al* (14) found that differences in the individual physician practice styles were more important in the measurement of ESR then patient characteristics or practice incentives. In contrast, Donald et al (15) found that factors most important to testing practices included the clinical experience with the patient, evidence from the literature and the rheumatologist's training. Clinician financial considerations were found to be of minimal importance when ordering laboratory tests. This variation in the measurement of ESR or CRP may also be the result of the rheumatologists' uncertainty surrounding the understanding of the value of using the tests (14). The relative value of the two tests is uncertain due to sparse and inconclusive literature. Some studies (17, 18) have found ESR to be the better test while other studies (19, 20) have found CRP superior. Wolfe *et al* (3) reported that both, ESR and CRP perform similarly in the clinic although CRP is the better test to measure acute phase. But a recent study (21) reported that neither ESR nor CRP is the better measure in the clinical setting and found no correlation between ESR or CRP and disease activity measures.

To improve the clinical management of RA and facilitate consistent monitoring of disease activity in the clinic, an understanding of the measurement

practices of ESR and CRP in a clinical setting would be useful. Therefore, the goals of the study were (1) to measure the frequency of the measurement of ESR and CRP using data from a large, representative sample of patients in clinical care, the Consortium of Rheumatology Researchers of North America (CORRONA) registry (22); (2) to identify predictors of the measurement of ESR and CRP at a clinical encounter; and (3) to quantify the variability among rheumatologists when measuring the ESR and CRP. The study utilized a multi-level modeling strategy to identify individual-level and population-level factors of the data instead of aggregating them only to the individual level. While most related work has focused on the individual patient level, this study focused on factors associated with clinical encounters, patients, physicians and their practice sites that may affect measurement practices of ESR or CRP in the clinic.

3.3 Methods

Design and study population

In this longitudinal, population-based study to assess the measurement of ESR and CRP among rheumatologists, all clinical encounters of patients with RA, during the study period between October 1, 2001 and November 1, 2008, were included in the study sample (n=92,062). The sample had a four-level hierarchical data structure: clinical encounters (level 1) nested within patients (level 2), treated by an individual physician (level 3) who practiced from a clinic

(level 4). The study utilized a multi-level mixed-effects logistic regression analysis to account for the correlation between the observations (encounters), and clustering of the data at the patient, physician and clinic levels. The units of investigation in the study were the clinical encounter and each cluster. The study was approved by the local institutional review board.

Data source

Data for the study was obtained from the Consortium of Rheumatology Researchers of North America registry (CORRONA) (22). CORRONA is an independent registry of patients with RA, osteoarthritis (OA), psoriatic arthritis (PsA) and osteoporosis from multiple sites across the United States. Since 2001 the registry has enrolled 113 clinics, 272 rheumatologists and over 17,000 patients with RA. Data are collected prospectively from both patients and their treating rheumatologists using survey questionnaires. Information collected from the patient includes: demographics, functional status, assessment of pain, family history, medical co morbidities and the presence of morning stiffness and fatigue. Data collected by the physician includes: disease activity measures as measured using the ACR core criteria, data on laboratory measures including ESR, CRP, and rheumatoid factor (RF), radiographic evidence of erosion and joint space narrowing, medication usage including disease modifying anti-rheumatic drugs (DMARDs), anti-inflammatory drugs (NSAIDs), biologics, and corticosteroids, treatment patterns, and adverse events . A separate questionnaire administered

to the physician upon enrollment into the registry collected information on physician demographics and practice characteristics. Follow-up examinations occur at three month intervals and are completed during routine clinical encounters.

Outcome measures

To examine the measurement of ESR and CRP, three outcomes of interest were evaluated. The first outcome, “APR measurement”, was a composite of either ESR or CRP measurement. It was operationalized as a dichotomous variable coded ‘1’ if measurement occurred and ‘0’ if there was no measurement. The measurement of the individual components of this outcome variable, the ESR and CRP were also examined in separate analyses. Both outcomes were operationalized as dichotomous and coded ‘1’ if ESR or CRP were measured and coded as ‘0’ if no measurement occurred. If ESR or CRP or both were measured at a clinical encounter, the values were collected at the time of the clinical encounter and the physician recorded the values of these lab tests on a physician questionnaire.

Predictors

Potential predictors of the outcomes were abstracted from the 4 levels of the data and included (a) clinical encounter, (b) patient, (c) physician and (d) clinic. Variables corresponding to each category were defined and measured at

their respective level. Descriptions of these variables, their source and classifications are presented in Table 3.1.

Clinic-encounter factors: Since APR's have been used to monitor inflammatory activity, disease activity may be predictive of their measurement. Therefore, the following physician-reported disease activity measures were assessed as predictors: presence of erosive disease and joint space narrowing, rheumatoid factor, tender joint with 28 count (TJC), swollen joint with 28 count (SJC), patient global assessment (PGA), physician global assessment (PhGA), and a composite measure, the Clinical Disease Activity Index (CDAI) (23). The DAS, DAS28 and SDAI were not included in the models because they contain an APR (ESR or CRP), but descriptive statistics for the measures were tabulated. Three patient-reported measures of disease activity were assessed for their influence on the measurement of an APR. These measures included the modified Health Assessment Questionnaire (mHAQ) (24, 25) a measure of functional status, level of pain (visual analog scale 0-100) and morning stiffness. Since an APR is often measured in conjunction with the ordering of other laboratory tests, a variable, "other laboratory tests" was derived to identify encounters where other lab tests had been ordered. The dichotomous indicator was coded as '0' if no other tests were measured and '1' if other tests were measured at a visit. Other laboratory tests in this variable included: cyclic citrullinated peptide, platelets, hematocrit, aspartate aminotransferease, alanine

aminotransferease, white blood cells, creatinine, albumin, PPD. Previous history of measurement of an APR may influence a measurement at a clinical encounter. Thus, a variable, “history of APR measurement”, was derived to indicate history of previous APR testing. It was derived by summing the number of previous encounters per patient where an ESR or CRP had been measured. Increased inflammation may result in therapeutic changes. Thus, the influence of current or past use of therapies to treat RA as predictors of the measurement of an ESR or CRP was examined. These included: the number of previous disease modifying anti-rheumatic drugs (DMARDs), current use of prednisone, and analgesics. The influence of treatment with gastrointestinal medications and antidepressants on the measurement of an APR was also examined. Calendar year of examination was included as a potential factor because treatment recommendations for RA may have changed depending on the year.

Patient-level factors: Patient characteristics examined in the analysis included: age, gender, duration of RA, marital status, education level, body mass index (BMI), smoking status, and insurance status. To examine the influence of medical comorbidities on the measurement of an APR, a variable, “comorbidity” was derived and included coronary artery disease (CVD), myocardial infarction (MI), congestive heart failure (CHF), stroke, diabetes, and cancer. It was operationalized as dichotomous with ‘0’ indicating no comorbidity and ‘1’ indicating the presence of a comorbidity from the above list. Anemia and

infections have been found to raise the level of ESR (26) and thus the presence of these conditions was assessed as predictors.

Physician-level factors: Factors at the physician-level that were included in the analysis were physician gender, age, years in practice, the proportion of managed care in the physician's practice, and the influence of formulary restrictions. If the physician has a high proportion of managed care in their practice, then frequency of laboratory testing may be restricted. The physician's past propensity to order an APR and its association with ordering an APR was examined. A variable, "physician propensity to order APR tests", was derived by calculating the mean number of tests previously ordered by an individual physician across their patient visits.

Clinic-level factors: Two clinic characteristic variables were assessed in the analysis and included the type of clinic (academic or private) and its geographical location.

Statistical analysis

Descriptive statistics of the predictors of measurement of an APR from each data level were calculated according to their APR measurement status and their mean values and standard deviations are presented. The frequency of a rheumatologist's measurement of the three outcomes was tabulated. To improve

the robustness of the study results, the study sample was randomly and evenly divided into a training (n=46,091) and test (n=45,971) data set. The training dataset was used to develop the models and the test dataset was used to test the model fit. The model-building process consisted of 3 stages. First, univariate models were constructed using multi-level univariate mixed-effects logistic regression to examine the relationship between all potential predictors and the outcome. Second, variables that were significantly associated with an outcome at the $p<0.10$ level were retained for inclusion in subsequent multi-level multivariable mixed-effects logistic regression models and all other variables were excluded. Third, all remaining significant predictors were evaluated in a final multi-level multivariate mixed-effects logistic regression model and retained if $p<0.05$. This final model incorporated the most predictive variables from each of the four data levels. Separate modeling was repeated for the measurement of an ESR and the measurement of a CPR. All analyses were performed using the STATA software 10.1 (27).

Multi-level Models

Five multi-level mixed-effects logistic regression models were constructed separately for each of the three outcomes. The models are mixed-effects because some of the coefficients are modeled as fixed and others are random. In these analyses, the fixed effects were clinical encounter, patient, physician and clinic characteristics and each cluster defined as patient, physician and clinic,

were the random effects. The first model was an empty model containing no predictors. The second model (level 1), contained the encounter-level variables significant in univariate analysis. The third model (level 2), was extended to also include significant patient-level variables. The fourth model (level 3) was extended to include significant physician-level variables. The fifth model (level 4) was the combined model and extended to include significant clinic-level variables. Since the multi-level analysis involved 4 levels, a 4-stage system of equations was conceptualized. The probability of the outcomes for clinical encounters *i* of the patient *j* seeing the physician *k* from the clinic *l* was examined.

Level 1 model:

The level 1 model is specified to compare clinical encounters with the outcome which were from the same patients (Equation 1). For example, an APR measurement for an encounter *i* in a patient *j* is modeled as:

$$\text{Logit}\,(APR_{ij}) = \beta_{0j} + \beta_{1j}X_{ij} \quad (1)$$

where the dependent variable, APR measurement (APR), distinguished between those clinical encounters that had an APR measured (1) and those that did not (0) and X_{ij} is the clinical encounter variable. The probability of an APR

measurement for each encounter is modeled as a function of factors at the encounter-level which were significant in bivariate analysis.

Level 2 model:

The level 2 model takes into account the differences between the patients and explains these differences in terms of patient characteristics at this level. In equation 2, the intercept from level 1 (β_{0j}), equation 2, was modeled as a function of significant patient-level predictors of the outcome. This modeling means that within-patient random intercepts (μ_{0j}) of each patient vary systematically with the patient characteristics that were significant in bivariate analysis.

$$\beta_{0j} = \beta_{00} + \beta_{01} X_j + \beta_{02} X_j + \mu_{0j} \qquad (2)$$

Level 3 model:

The level 3 model takes into account the differences between the physicians and explains these differences in terms of physician characteristics at this level. In equation 3, the intercept from level 2 (β_{00}), equation 2, was modeled as a function of significant physician-level characteristics from the bivariate analysis. This modeling means that within-physician intercepts (μ_{0k}) of each physician vary systematically with physician characteristics.

$$\beta_{00} = \beta_{0k} + \beta_{01} X_{jk} + \beta_{02} X_{jk} + \beta_{03} X_{jk} + \mu_{0k} \qquad (3)$$

Level 4 model:

The level 4 model takes into account the differences between the clinics and explains these differences in terms of clinic characteristics at this level. In equation 4, the intercept from level 3 (β_{0k}), equation 3, was modeled as a function of significant predictors at the clinic level. This modeling means that within-clinic random intercepts of each clinic (μ_{0l}) vary systematically with these significant predictors.

$$B_{0k} = \beta_{0j} + \beta_{01} X_l + \beta_{02} X_l + \mu_{0l} \qquad (4)$$

Combined Model:

The combined model is described in equation 5 by substituting equations 4 into 3 then 3 into 2 and then 2 into 1.

$$\text{Logit}(\text{APR}_{jkl}) = \beta_{00} + \beta_{1j} X_{ij} + \beta_{01} X_j + \beta_{02} X_j + \beta_{03} X_{jk} + \beta_{04} X_{jk} + \beta_{05} X_{jk} + \beta_{06} X_l + \beta_{07} X_l + \mu_{0j} + \mu_{0k} + \mu_{0l} \qquad (5)$$

Fixed effects (measures of association)

The associations of APR measurement and the statistically significant factors at the clinical encounter-, patient-, physician- and clinic-levels are presented as odds ratios (OR) with their 95% confidence interval (CI).

Random effects (measures of variation)

Measures of random effects included the random intercept variance, intra-cluster correlation (ICC), and median odds ratio (MOR). The ICC is a measure of dependence or correlation among the dichotomous responses for the same subject or between subject heterogeneity. It was calculated according to the formula used by Rabe-Hebeth *et al* (28). An alternative measure of heterogeneity, the median odds ratio (MOR) (29, 30) was also used to quantify the unexplained cluster heterogeneity.

Model fit

Multi-collinearity or the correlation among independent variables was examined using the variable inflation factor (VIF) (31). Variables found to have a VIF >2 were excluded from further modeling. To test the assumption of linearity of the independent continuous variables with an outcome, the Box-Tidwell transformation (32) was performed. In this test, interaction terms which are the cross-product of each independent variable time its natural logarithm are calculated and then added to the model. If these transformations are significant then there is nonlinearity in the logit. Linearity was also tested graphically by plotting the log odds for each of the probabilities of the outcome compared to the independent variable. Model fit to the observed data was determined using receiver operating characteristic (ROC) curve analysis (33) and plotting the area under the curve (AUC) (34).

3.4 Results

Sample characteristics

As shown in Table 3.2, there were 92,062 clinical encounters (level 1), nested within 17,450 patients (level 2) who were patients of 272 physicians (level 3) from a total of 100 clinics (level 4). Among the 92,062 encounters, an APR measurement occurred in 47,164 (51%) visits. Of theses visits, 39,400 (43%) were measurements of ESR, and 28,327 (31%) were measurements of CRP. Both the mean ESR (23.86 mm/h) and the mean CRP (2.63mg/dL) values were normal (abnormal ESR>25mm/h and abnormal CRP≥0.5mg/dL) (35).

Clinic encounters where an APR was measured had a mean (SD) CDAI and DAS28 scores of 11.66 (11.5) and 3.4 (1.5), respectively, indicating low to moderate disease activity. The proportion of other disease activity measures was similar across both types of visits. Two variables, “other laboratory tests” and “history of APR measurement” differed between the encounters. In encounters where an APR had been measured, the average number of other laboratory tests ordered was 50% higher compared to encounters with no APR measurement. In encounters with an APR measurement, the number of previous APRs was almost 4 times higher then in encounters with no APR measurement.

Patients in the study sample were mostly female (75%) aged 60.3 years on average and had an average disease duration of 11.7 years. Their average (SD) BMI was 29.15 (9.3) and 15% were smokers. Most had private insurance or

Medicare and approximately 3% had no insurance. The proportion of patient characteristics were similar in encounters where an APR was measured compared to those where no measurement occurred.

Most physicians in the study sample where male (85%), were on average 55.89 years old had had practiced on average 24.46 years. 70% had at least 25% or more managed care in their practices and more had frequent formulary restrictions (55% to 44%). The proportion of physician characteristics was similar across encounters with the exception of physician practice years. In encounters with an APR measurement, physicians had been in practice slightly less years compared to visits with no APR measurement. Clinics were primarily private and located mostly in the Northeast and Midwest.

Frequency of APR measurement

Large differences were found among the rheumatologists in their measurement of the three outcome variables of interest (Table 3.3). On average a rheumatologist measured one APR test (i.e.-ESR or CRP or both) at a clinical encounter. Measurement of ESR tests was done in 51% of the visits of the average rheumatologist, but one quarter of the rheumatologists ordered this test in 38% or less of their encounters and another one quarter in 83% or more. Testing for CRP was performed in 31% of the clinical encounters.

Univariate analysis

Descriptive results from univariate analysis for variables significantly associated with a rheumatologists' decision to measure an ARP at a $p \leq 0.10$ level are presented in Table 3.4. Variables are grouped in the four hierarchical levels of the data listed earlier. More then half (n=21) of the 40 factors assessed for associations with the measurement of an APR were statistically significant when compared in univariate analysis. Disease activity measures such as the CDAI were not found to be significant predictors of the measurement of an APR and were not retained for further analysis. Univariate results for the second and third outcomes of interest, measurement of an ESR and measurement of a CRP were similar to the results for outcome one and are not presented.

Fixed effects (measures of association)

The multivariate analysis identified seven variables that were most strongly associated with APR measurement at $p \leq 0.05$ level (Table 3.5). Odds ratios indicate the likelihood that an event will occur relative to the likelihood that it will not. Having other laboratory tests ordered at the visit increased the odds of having an APR measured by more than 2- fold. Encounters where there was a past history of DMARD use were 4% less likely to have a measurement of an APR. Female patients had a 10% increase in the odds of having an APR measurement and patients without any insurance coverage had a 60% less likely chance of an APR measurement. Disease duration was inversely related to

having an APR measurement. Increased propensities by a rheumatologist to order an APR increased the likelihood of an APR measurement. A clinic that was private had significantly lower odds of measurement of an APR test compared to an academic clinic. The variable “history of APR testing” was dropped from the model because of multi-collinearity. Using the test dataset (n=45,971), the fit of the final model was assessed by ROC curve analysis. The area under the curve was 0.8137 and indicated that the model had substantial ability to discriminate between encounters with an APR measurement compared to those without (Figure 3.1).

Multivariate results for the second and third outcome, measurement of an ESR and measurement of CRP, respectively, were very similar to the results listed above and thus are not reported. The only exception was the type of clinic was not a significant predictor in the measurement of ESR.

Random effects (measures of variation)

The amount of residual or unexplained variation attributable to variability between patients, physicians and clinics for outcome 1, measurement of an APR, is presented in Table 3.6. Unmeasured differences between clinics and between patients represent the largest amount of variability not explained by the model. In model 1 (empty model), there was variation in the log odds of the measurement of an APR across the clinics (τ =2.03), across physicians (τ =0.89) and across patients (τ =0.67). According to the intra-clinic and intra-physician and intra-

patient correlation coefficients, 29.5%, 12.9% and 9.7% of the variance in the measurement of an APR could be attributed to clinic, physician and patient factors, respectively. After controlling for visit-, patient-, physician- and clinic-level factors in the full model (Table 3.6, Model 5), the variance in the measurement of an APR was decreased for intra-physician to 5.1% but increased for intra-clinic and intra-patient to 29.6%, and 14.8%, respectively.

Results form the MOR also confirmed evidence of patient and clinic characteristics shaping physician attitudes towards measuring an APR. The high MOR (3.09) in the final model (Table 3.6, Model 5) between encounters in a clinic suggests that the clinic heterogeneity in the measurement of an APR is substantial. The MOR between encounters within physicians with a higher compared to a lower propensity to measure an APR was 1.60. This low odds ratio suggests that the clustering effect is low at the physician level. Thus there was little evidence of variations between physicians when measuring an APR.

3.5 Discussion

Using data from a large observational cohort of patients' representative of a clinical care setting, results of the study indicate less than half of the visits had a measurement of ESR and less than a third had a CRP measurement, a finding in contrast to results of a recent study where most rheumatologists reported measuring ESR and CRP at most of their clinical encounters (15). Several

significant predictors of the measurement of an APR were identified in multivariate analysis including the ordering of other laboratory tests at the clinical encounter, the number of previous DMARDs prescribed, several patient characteristics, the propensity of the physician to order APR tests and the type of clinic. The study also found a wide variation among rheumatologists when ordering the measurement of an ESR or CRP. Most of the clustering effects were due to clinic and patients characteristics rather than individual physicians. Thus efforts to further explore the variability in the measurement of an APR should be targeted towards the clinics and also at individual patients and not at individual physicians.

The discrepancy between the frequency of measurement of an ESR and CRP in our study compared to a prior study which reported over 85% of physicians measuring ESR and CRP at every clinical encounter may be due to the sampling inconsistencies. In the prior study, data was collected using a mail survey with only 29% of rheumatologists responding. Thus, the finding that most rheumatologists are measuring ESR or CRP may be overinflated especially if the rheumatologist who did not return the questionnaires did not routinely measure ESR or CRP. Since our study utilized data from a large observational registry of patients with RA from across the United States that are representative of clinical care setting, our finding of limited ESR and CRP testing are most likely accurate.

The ordering of other laboratory tests at a clinical encounter significantly increased the likelihood that an APR would be ordered. This finding suggests

that rheumatologists may be primarily ordering the measurement of an APR because they are part of a standard battery of routinely ordered tests and not because of a specific reason. Further research needs to be conducted to verify this finding.

The number of previous DMARDs prescribed was inversely related to the measurement of an APR, and thus an increased number of past DMARD prescriptions would mean a less likelihood that an APR would be ordered. Frequent switching of DMARDs is recommended by the ACR treatment recommendations if maximum treatment response is not achieved (13). This finding seems to indicate physicians may not be using ESR or CRP measurements to guide therapy decisions, a finding supported form a recent study (15).

Patient characteristics predictive of an APR measurement were female gender, having no insurance coverage and having greater disease duration. Female patients were found to have an almost 10% increased probability of having an APR measurement. This could be due to several factors such as patient demand for more testing or because values of ESR are known to be higher in women than men. This result would need to be further evaluated to determine why female gender increases the probability of having an APR done. Predictably, having no insurance coverage was directly related to not receiving an APR measurement with 60% less likely an APR measurement would be ordered. Thus this finding indicates insurance status does play a role in the

measurement of an APR. And patients with a longer duration of disease had lower odds of having a measurement of an APR. This is reflective of previous research that reports older patients receive less therapy changes (36).

A physicians' past propensity of ordering laboratory tests such as the ESR or CRP influenced the measurement of an APR with an almost 50% increase in the probability of a measurement if the physician had a higher past propensity. Thus physicians who have a proclivity to order APR tests in the past will be more likely to order them in the future. This finding might be the result of an individual physicians' past training or practice style. One clinic characteristic was found to be significantly predictive of an APR measurement. This study found that an APR would be more likely to be measured at an academic clinic rather than a private clinic. This result is consistent with a similar result found in another study (15). This finding could be due to academic sites having superior testing capability with laboratory facilities compared to private clinics that would most likely not.

ESR and CRP have long been thought to be used to monitor disease activity and response in patients with RA but the results of this study found the measurement of an APR was not associated with disease activity measures. This supports the results reported by Donald et al that physicians are measuring ESR but not using it to guide therapy. Also this confirms results found recently by Keenan *et al* (21) that ESR or CRP were not correlated with any disease activity measures. Thus, based on the results of this study, rheumatologists are not measuring an APR to monitor disease activity or treatment decisions.

The study has several strengths. First, it is a large, population-based nationally-representative study with data from numerous private and academic clinics from four geographic regions across the United States. Thus the results can be generalized to most areas of the country. Second, the study design was longitudinal and the effects of correlation and clustering were accommodated by the utilization of a multi-level modeling strategy. A multilevel approach analyzes the social, cultural and economic context in which individual patient experience health outcomes. Beyond their physicians, individual patients will be influence by the clinics that their physician practices from and how they perceive measurement of an APR. Therefore, using a multilevel approach, we were able to study the associations of different levels that correspond to health care and provide more robust evidence about individual, physician and clinic factors that influence the measurement of an APR. Understanding the relative contribution of these factors from these levels is important towards understanding how physicians determine whether to measure an APR. Third, the study utilized training and test datasets to test the fitted model and thus increased the robustness of the results. Lastly, the number of included clinics, physicians and patients in the study and geographic diversities strengthen the findings from the study. A potential limitation to this study is that there was no information about testing equipment in the clinics which might explain why the APR measurement was low in private clinics.

Multilevel analysis was used to control for clustered data but other statistical procedures also account for clustering such as GEE modeling or alternating logistic regression (ALR). A multi-level modeling approach was utilized because the study was interested in the average change at the individual level and to allow the effects of factors at different levels to be used. Conversely, ALR and GEE modeling estimate a population-averaged approach which was not related to the study's objective.

In conclusion, the results of this study indicate that measurement of an APR in the clinic is influenced by physician practice decisions. If other laboratory testing are ordered and if the physician has a propensity to order an APR, then the probability of an APR measurement was increased. Several patient characteristics were associated with APR measurement including gender, and insurance status. The study did find that most of the unexplained variability in the measurement of an APR was due to differences at the clinic and patient levels and not the individual physician. Lower variation among physicians may indicate that most physicians are unclear as to when to measure an APR, suggested by Wolfe *et al* (3). Since many other tests are available to physicians, measurement of ESR and CRP may only be done in conjunction with other lab tests, a finding supported by our results. Based on these findings, future public health interventions to standardize APR testing in an effort to contain costs could be focused on clinic and patient levels instead of the individual physician. Surprisingly, the study found that disease activity levels were not associated with

the decision to measure an APR. This finding may suggest that measurement of an APR is not ordered to monitor disease level or to guide treatment decisions. Rethinking the role of ESR and CRP in the clinical management of RA may be warranted.

Table 3.1 Variables used in the models and associated data sources and definitions

Factor	Source	Classification
Level 1: Clinical encounter level		
Rheumatoid factor	Physician	yes vs. no (reference)
Erosive disease	Physician	yes vs. no (reference)
CDAI	administrative	continuous variable
Tender joint count	Physician	0-28
Swollen joint count	Physician	0-28
PGA	Patient	0-100
PhGA	Physician	0-100
Pain Score	patient; visual analog scale	0-100
Functional Status	patient; measured by mHAQ	0-3
Morning Stiffness	patient	continuous variable
Calendar year of examination	Physician	continuous variable
Morning stiffness	patient	continuous variable
Other laboratory tests±	Physician	yes vs. no (reference)
No. of previous APR tests	Physician	continuous variable
No. of previous DMARDs	Physician	continuous variable
DMARDs	Physician	yes vs. no (reference)
Prednisone	Physician	yes vs. no (reference)
Gastrointestinal¶	Physician	yes vs. no (reference)
Analgesics*	Physician	yes vs. no (reference)
Antidepressants§	Physician	yes vs. no (reference)
Level 2: Patient level		
Patient gender	patient	male (reference) vs. female
Race: white	patient	yes vs. no (reference)
Patient age , year	patient	continuous variable
Duration of RA	physician	continuous variable
Married	patient	yes vs. no (reference)
Education level	patient	completed college vs. not completed (reference)
BMI	patient	continuous variable
Smoker	patient	yes vs. no (reference)
Insurance status	patient	
Private	patient	yes vs. no (reference)
None	patient	yes vs. no (reference)
Medicare	patient	yes vs. no (reference)
Medicaid	patient	yes vs. no (reference)
Comorbidity€	physician	yes vs. no (reference)
Infection	Physician	yes vs. no (reference)
Anemia	patient	
Level 3: Physician level		
Physician gender	Physician	male vs. female (reference)
Physician age, years	Physician	continuous variable
Physician years in practice	Physician	continuous variable
Managed care	Physician	<25% (reference), 25-50%, >50%; percent in practice
Formulary restrictions	Physician	frequently (reference) vs. infrequently; influence on

		prescribing
Level 4: Clinic level		
Clinic type	Physician	academic (reference) vs. private
Geographic location£	Physician	northeast (reference), midwest, south, west

±other lab tests include: cyclic citrullinated peptide, platelets, hematocrit, aspartate aminotransferease, alanine aminotransferease, white blood cells, creatinine, albumin, PPD
BMI-body mass index; APR-acute phase reactants (APR refers to ESR or CRP in this study)
mHAQ- modified Health Assessment Questionnaire; CDAI- Clinical Disease Activity Index; DMARDs- disease modifying anti-rheumatic drugs; PGA-physician global assessment; PhGA-physician global assessment.
*Analgesics include:
¶ Gastrointestinal medications include:
§Antidepressants include:
€comobidities include: hypertension, cancer, diabetes mellitus, congestive heart failure, coronary artery disease, myocardial infraction, stroke
£ northeast –Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New York, New Jersey, Pennsylvania; Midwest-Indiana, Illinois, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; south- Dellaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, Texas; west-Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming, Alaska, California, Hawaii, Oregon, Washington.

Table 3.2 Descriptive characteristics of clinical encounters by APR* measurement status (n=92,062)

Characteristics	N	Total Sample (n=92,062)	APR Measurement (n=47,164)	No APR Measurement (n=44,898)
Level 1: Clinical encounter (n=92,062)				
ESR, mean (SD)	39400	23.86(21.0)	23.86 (21.0)	----
CRP, Mean (SD)	28327	2.63(7.6)	2.63 (7.6)	----
Rheumatoid factor, % (SD)	19703	0.71(0.46)	0.71 (0.45)	0.69 (0.46)
Erosive disease, % (SD)	18299	0.50 (0.49)	0.49 (0.50)	0.49 (0.49)
Joint space narrowing, % (SD)	18489	0.64 (0.48)	0.65 (0.47)	0.63 (0.48)
CDAI, mean (SD)	85276	11.9 (11.71)	11.66 (11.54)	12.15 (11.81)
DAS28, mean (SD)	36829	3.38 (1.49)	3.38 (1.49)	----
SDAI, mean (SD)	26917	14.1 (14.71)	14.09 (14.71)	----
Tender joint count, mean (SD)	91234	3.11 (5.2)	3.09 (5.1)	3.12(5.3)
Swollen joint count, mean, (SD)	91248	3.97 (5.6)	4.15(5.9)	3.81 (5.3)
PGA, mean (SD)	86139	28.48 (24.9)	28.17 (24.9)	28.80 (24.9)
PhGA, mean (SD)	91767	19.51 (18.5)	19.17 (18.13)	19.87 (18.78)
Pain Score, mean (SD)	87430	30.81(25.4)	30.26 (25.26)	31.39 (25.53)
mHAQ, mean (SD)	88736	0.37 (0.46)	0.36 (0.46)	0.37 (0.46)
Morning stiffness, % (SD)	88350	0.94 (2.12)	0.93 (2.16)	0.96 (2.07)
Other laboratory tests±, % (SD)	92062	0.09 (0.28)	0.12 (0.33)	0.06 (0.23)
No. of previous APR tests, mean (SD)	92062	2.6 (2.93)	3.79 (3.12)	1.28 (2.84)
DMARD, % (SD)	92062	0.14 (0.34)	0.14 (0.35)	0.13 (0.34)
No. of previous DMARDs, mean (SD)	92062	1.31 (1.56)	1.27 (1.55)	1.34 (1.57)
Prednisone, % (SD)	91415	0.32 (0.47)	0.31 (0.46)	0.33 (0.49)
Gastrointestinal, % (SD)	92062	0.35 (0.48)	0.35 (0.48)	0.36 (0.48)
Analgesics, % (SD)	92062	0.60 (0.49)	0.60 (0.49)	0.60 (0.48)
Antidepressants, % (SD)	92062	0.26 (0.49)	0.25 (0.43)	0.27 (0.46)
Level 2: Patient (n=17,450)				
Age , mean (SD), y	91694	60.3 (13.2)	60.41 (13.3)	60.1 (13.8)
Sex , % female (SD)	92019	0.75 (0.43)	0.76 (0.43)	0.75 (0.43)
Ethnic origin, % Caucasian (SD)	91084	0.86 (0.39)	0.85 (0.34)	0.87 (0.34)
Disease duration, mean(SD), y	91604	11.7 (9.90)	11.7 (10.11)	11.7 (9.8)
BMI, mean (SD)	90065	29.15 (9.3)	29.07 (9.1)	29.23 (9.8)
Smoker, % (SD)	89055	0.14 (0.35)	0.14 (0.35)	0.15 (0.35)
Insurance status, % (SD)				
Private	80377	0.73 (0.44)	0.73 (0.44)	0.73 (0.44)
No insurance	80377	0.01 (0.11)	0.01 (0.10)	0.02(0.13)
Medicare	80377	0.42 (0.49)	0.43 (0.49)	0.41 (0.49)
Medicaid	80377	0.06 (0.23)	0.06 (0.24)	0.06 (0.23)
Comorbidities, % (SD)	92062	0.076 (0.27)	0.077 (0.27)	0.076 (0.27)
Anemia, % (SD)	88767	0.25 (0.43)	0.25 (0.43)	0.25 (0.43)
Infections, % (SD)	91909	0.17 (0.37)	0.18 (0.38)	0.15 (0.35)
Level 3: Physician (n=272)				
Age, mean (SD)	83898	55.9 (6.60)	55.6 (6.97)	56.2 (6.17)
Sex, % male (SD)	86580	0.85 (0.36)	0.85 (0.36)	0.85 (0.35)
Years in practice, mean (SD)	84222	25.47 (52.4)	24.13 (18.18)	26.86 (72.3)
Percent of managed care in practice, % (SD)				

<25%	23646	0.26 (0.44)	0.29 (0.45)	0.22 (0.41)
25-50%	24045	0.26 (0.44)	0.24 (0.43)	0.28 (0.45)
>50%	36925	0.40 (0.45)	0.38 (0.49)	0.42 (0.49)
Formulary restrictions, % (SD)				
Frequently	47556	0.52 (0.49)	0.51 (0.49)	0.52 (0.49)
Infrequently	39820	0.41 (0.49)	0.41 (0.49)	0.41 (0.49)
Level 4: Clinic (n=100)				
Type of clinic, % (SD)				
Academic	10975	0.12 (0.32)	0.13 (0.31)	0.10 (0.30)
Private	73146	0.79 (0.40)	0.79 (0.42)	0.82 (0.38)
Geographical region, % (SD)				
Northeast	43357	0.47 (0.49)	0.47 (0.49)	0.47 (0.48)
Midwest	22592	0.25 (0.43)	0.28 (0.45)	0.21 (0.41)
South	13667	0.15 (0.35)	0.10 (0.30)	0.20 (0.39)
West	5683	0.06 (0.24)	0.06 (0.24)	0.06 (0.23)

*APR-acute phase reactants (in this study refers to erythrocyte sedimentation rate and C-reactive protein)
±other lab tests include: cyclic citrullinated peptide, platelets, hematocrit, aspartate aminotransferease, alanine aminotransferease, white blood cells, creatinine, albumin, PPD.
ESR-erythrocyte sedimentation rate; CRP- C-reactive protein; CDAI- Clinical Disease Activity Index; DAS28- Disease Activity Score 28; SDAI=Simplified Disease Activity Index; BMI-body mass index; mHAQ- modified health assessment questionnaire; PGA-patient global assessment; PhGA-physician global assessment.

Table 3.3 Observed Variation: Propensity to use tests among rheumatologists in clinical encounters (n=92,062)

			Distribution				
Outcome	Mean	SD	Minimum	First quartile	Median	Third quartile	Maximum
No. APR tests done	1.18	0.72	0.00	0.46	1.08	1.88	2.00
No. ESR done	0.51	0.41	0.00	0.38	0.58	0.83	1.00
No. CRP done	0.31	0.27	0.00	0.12	0.35	0.48	1.00

APR-acute phase reactant; ESR-erythrocyte sedimentation rate; CRP- C-reactive protein

Table 3.4 Univariate analysis of rheumatologist's decision to measure an APR at a clinical encounter (training dataset: n= 46,091)

Variable	Odds Ratio (95% CI)	*P* Value
Level 1 : Clinical encounter		
Other laboratory tests±	2.59 (2.36-2.83)	0.00
History of APR tests	1.46 (1.44-1.47)	0.00
No of previous DMARDs	0.96 (0.94-0.98)	0.00
Prednisone	0.94 (0.91-0.97)	0.00
Level 2: Patient		
Patient age	1.00 (1.00-1.004)	0.01
Female	1.05 (0.99-1.10)	0.00
White	0.84 (0.78-0.89)	0.03
Duration of RA	1.00 (0.99-1.004)	0.00
Married	0.95 (0.91-0.99)	0.05
College	1.05 (1.00-1.09)	0.06
No insurance	0.87 (0.58-0.76)	0.00
Medicaid	1.12 (1.04-1.19)	0.02
Infections	1.03 (1.00-1.07)	0.03
Level 3: Physician		
Physician age	0.97 (0.97-0.98)	0.00
Male	0.83 (0.78-0.88)	0.00
Practice years	0.99 (0.991-.0997)	0.00
% managed care		
<25%	1.00 (Reference)	
25-50%	0.71 (0.67-0.75)	0.01
>50%	0.72 (0.68-0.76)	0.02
Infrequent formulary restriction	1.05 (1.00-1.10)	0.03
Propensity to order lab tests	1.52(1.24-1.65)	0.00
Level 4: Clinic		
Private	0.73 (0.68-0.77)	0.00
Geographic region		
Northeast	1.00 (Reference)	
Midwest	1.34 (1.27-1.43)	0.02
South	0.66 (0.62-0.71)	0.02
West	1.09 (1.01-1.19)	0.01

±other lab tests include: cyclic citrullinated peptide, platelets, hematocrit, aspartate aminotransferease, alanine aminotransferease, white blood cells, creatinine, albumin, PPD.
DMARDs –disease modifying anti-rheumatic drugs; APR- acute phase reactant

Table 3.5 Multivariate model predicting rheumatologists' decision to measure an APR at a clinical encounter (training dataset: n= 46,091)

Variable	Odds Ratio (95% CI)	*P* Value
Other laboratory tests±	2.61 (2.33-2.93)	0.00
No. of Previous DMARDs	0.96 (0.94-0.98)	0.003
Female	1.09 (1.00-1.18)	0.03
Duration of disease	0.99 (0.992-0.999)	0.036
No insurance coverage	0.37 (0.27-0.49)	0.00
Propensity to measure APR tests	1.46 (1.34-1.59)	0.00
Private clinic	0.40 (0.20-0.81)	0.01

±other lab tests include: cyclic citrullinated peptide, platelets, hematocrit, aspartate aminotransferease, alanine aminotransferease, white blood cells, creatinine, albumin, PPD.
DMARDs –disease modifying anti-rheumatic drugs; APR- acute phase reactant;

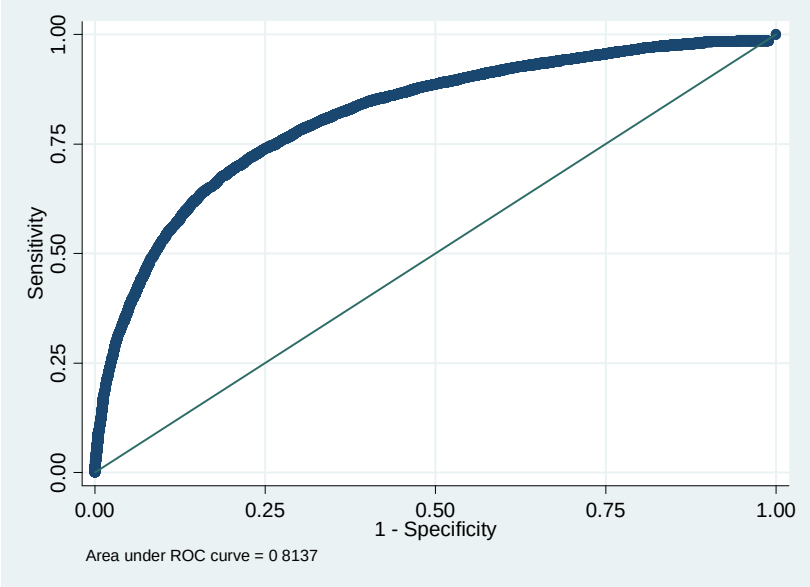


Figure 3.1 Receiver Operating Curve Analysis of full model to discriminate measurement of an APR (test dataset: n=45,971)

CHAPTER IV

A Modified Rheumatoid Arthritis Disease Activity Score With 28 Joint Count (mDAS28) For Epidemiological Research

The following work of the same name is currently under review by the Journal of Rheumatology.

4.1 Abstract

Objective. To examine the measurement properties and validity of a modified version of the DAS28 (mDAS28), as an alternative measure of disease activity when laboratory results are unavailable.

Methods. In a cross-sectional development cohort (5,729 patients), statistically significant predictors of the logarithm of erythrocyte sedimentation rate (lnESR) were identified using linear regression analysis. After computation of the mDAS28, a cross-sectional validation cohort (5,578 patients) was used to evaluate criterion and construct validity of the mDAS28. The ability of the mDAS28 to discriminate between various disease states was also assessed. A second validation cohort (longitudinal-336 pairs of patient visits) was used to assess sensitivity to change.

Results. Significant predictors of lnESR included tender joint count, swollen joint count, modified health assessment questionnaire (mHAQ), physician global assessment (PhGA) and patient pain score (VAS). The mDAS28 had high criterion validity with strong Spearman-rank correlations with the DAS28, SDAI and CDAI (r= 0.87, r=0.91, r=0.96). Predictive validity was demonstrated by good correlation with the mHAQ (r=0.58). The mDAS28 showed substantial agreement with the DAS28, SDAI and CDAI when discriminating between disease states

(κ=0.70-0.77) and moderate to substantial agreement between response levels (κ=0.52-0.73). Both mDAS28 and DAS28 measures classified patients similarly in remission compared to the SDAI and CDAI. The mDAS28 was superior when detecting change (SRM =0.58) followed by the DAS28, CDAI and the SDAI.

Conclusion. The mDAS28 is a valid and sensitive tool to assess disease activity and can be used as an alternative measure to the DAS28 when laboratory values are missing.

Key words: DAS28, mDAS28, EULAR, disease activity measures, CDAI, SDAI

4.2 Introduction

The Disease Activity Score with 28 joint count (DAS28) (1) is one of the most widely used and validated composite measures of disease activity in rheumatology. It is a modified version of the original Disease Activity Score (DAS), which was developed by van der Heijde *et al* (2, 3) in 1990. The DAS28 is calculated using the following parameters: tender joint count (TJC), swollen joint count (SJC), visual analog scale of patient general health (PGA) and a laboratory value, the erythrocyte sedimentation rate (ESR). The DAS28 comprises the European League Against Rheumatism (EULAR) response criteria (4) used to measure response to treatment in clinical trials. Cut points to delineate disease activity levels and change in disease levels have been derived (4, 5).

The DAS28 has been a reliable measure of treatment efficacy in clinical trials along with the American College of Rheumatology criteria (6) and its use has been recommended by the EULAR (7). In the clinical management of RA, the DAS28 is required by several regulatory bodies when determining patient eligibility for biologic treatments. In several research studies, the DAS28 has been a benchmark for validation of new composite indices (8-12).

However, the utility of the DAS28 has been limited, in epidemiological research using disease registry data, due to missing ESR laboratory values (13). In clinical trials, ESR values are available as mandated by study protocol thus allowing the computation of the DAS28. But in some practice settings, the ESR

laboratory test is not routinely ordered or not available during the patient visit, preventing the DAS28 from being calculated and omitting those patients from the analysis (14).

New simplified measures such as the Simplified Disease Activity Index (SDAI) (10) and the Clinical Disease Activity Index (CDAI) (8), offer some advantages over the DAS28 because of their simplified formulas. And the CDAI in particular does not require a lab value to be calculated (8). A recent literature review (15), reported that the SDAI and CDAI were found to be strongly correlated with the DAS28 in several studies. However, differences between the measurement performances of these simplified measures and the DAS28 were found. When assessing individual patients, classification of patients into remission differed between the measures.

To facilitate the calculation of the DAS28 in epidemiological research using disease registry data, one approach would be to modify the DAS28 by replacing the natural logarithm of ESR (lnESR) with disease activity measures that are measurable at every clinical encounter. Thus, the DAS28 could be calculated when ESR values were unavailable. Therefore, the aim of this study was to develop a modified version of the DAS28 that does not contain the ESR: the modified DAS28 (mDAS28) and to assess its comparability with the DAS28 and its validity according to the Outcomes Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) recommendations (16, 17).

4.3 Patients and Methods

Subjects were selected from a large North American registry, the Consortium of Rheumatology Researchers of North America (CORRONA) (18). The methods of this registry have been described in detail elsewhere (19). Patients eligible for the study had all of the parameters necessary to calculate the DAS28. Patients who did not have all of these components were excluded from further analysis. The demographic and disease activity measures of excluded and included patients were compared in a separate analysis to determine whether any bias was entered due to sample selection.

The study utilized 3 samples, cross-sectional “development” and “validation” datasets and a longitudinal “validation” dataset (Table 4.1).

The first two datasets were from a cohort of 11,307 RA patients. A cross-section of this cohort was obtained by utilizing information from the patient’s most recent visit. This cross-sectional cohort was randomly and evenly divided into “development” and “validation” datasets. The cross-sectional development dataset (n=5729) was used to build a prediction model to identify statistically significant predictors of the lnESR and the cross-sectional validation dataset (n=5578) was used to subsequently validate the mDAS28.

The third dataset, a longitudinal “validation” dataset was from a cohort of 703 patients with RA who had two paired visits. The first visit contained an initiation of a disease modifying anti-rheumatic drug (DMARD) and the second

visit occurred at least three months after the first. This longitudinal cohort was then divided randomly and evenly into “development “(n= 336 pairs) and “validation” datasets (n=367 pairs). The longitudinal validation dataset was used to evaluate the mDAS28 as a measure of response.

Disease activity parameters needed to compute the DAS28 were collected by a rheumatologist. Measures also collected included the modified Health Assessment Questionnaire (mHAQ) score (20), a measure of functional status, the patient visual analog score (VAS) pain score, physician global assessment of disease activity (EGA) and duration of morning stiffness. The DAS28 values were calculated according to its formula (2):

$$DAS28 = 0.56 * \sqrt{(28TJC)} + 0.28 * \sqrt{(28SJC)} + (0.70 * lnESR) + 0.014 * PGA$$

SDAI and CDAI values were also calculated according to their formulas (8, 10) and used as comparators along with the DAS28 to validate the mDAS28. To validate the mDAS28 as a measure of response, response criteria, namely the EULAR response criteria, were used. Response criteria for the SDAI and the CDAI, based on previously published absolute cut points (21) and change cut points (22) were also calculated and used as additional comparators.

Statistical analysis

To develop the mDAS28**,** statistically significant predictors of the lnESR were identified in the cross-sectional development cohort, using univariate linear regression analysis. Candidate predictors of lnESR included TJC, SJC, VAS, PhGA, PGA, mHAQ, and the duration of morning stiffness. Candidate variables that were significant at an alpha level of 0.10 were included in a multivariate model. Forward and backward stepwise regression was used to identify the most significant independent variables using a *P* value of 0.10 as the removal criterion. Multi-collinearity between the significant independent variables, specified in the multivariable model, was determined using the variance inflation factor (VIF) (23). Variables found to be collinear (VIF>2) were dropped from the model.

The validity of the mDAS28 as a measure of disease activity and response was then evaluated in the cross-sectional and longitudinal validation datasets. Criterion validity or the extent that a measure correlates with a “gold standard” was examined by correlating mDAS28 scores with DAS28, SDAI and CDAI. Predictive validity, the ability of a measure to predict the future outcome of the disease was examined by correlating the mDAS28 scores with the mHAQ. Both validities were assessed by Spearman-rank **c**orrelation coefficients (24). The amount of agreement between the mDAS28 and the other disease activity indices to discriminate between different disease states of individual patients (remission, low, moderate, high) and between good, moderate or none levels of response based on the EULAR response criteria was examined using weighted

kappa statistics (25). The EULAR response criteria were calculated according to its algorithm in Figure 4.1. Since new cut points were not derived for the mDAS28, modified EULAR (mEULAR) response criteria were calculated according to the EULAR response criteria, using mDAS28 scores and DAS28 cut points.

The SDAI and CDAI response criteria were derived in the same manner as the EULAR response criteria, with the exception that the absolute cut points previously defined by Aletaha *et al* (21) and the change cut points defined by Ranganath *et al* (22) for both measures were used instead of DAS28 cut points.

The sensitivity to change or responsiveness of the mDAS28, the ability of a measure to detect important changes over time after a treatment has been initiated, was evaluated by calculating the effect size (ES) (26) and standardized response mean (SRM) (27). The effect size was calculated by taking the mean differences of the disease activity scores between the baseline and second study visits (mean change scores) and then divided by the standard deviation of the baseline scores. The SRM was calculated by taking the mean change scores and then dividing the result by the standard deviation of the change scores. The values of the ES were small with a range of 0.2-0.5, moderate if 0.5-0.8 or large if >0.8 (26). The SRMs were interpreted similarly (28). Statistical analysis was carried out using Stata version 10.0 (Stata Corporation, College Station, TX) (29).

4.4 Results

A total of 11,307 patients were eligible for the cross-sectional cohort and 703 pairs of patients with an initiation of a DMARD and at least 3 months until the first follow up visit, were eligible for the longitudinal cohort. Demographics and clinical characteristics for the cross-sectional development and validation dataset were generally similar (Table 4.1). Both cross-sectional datasets exhibited mild to moderate disease levels but, in the longitudinal validation dataset, disease activity measures had higher values. This would be expected since the patients in the longitudinal dataset were initiators of DMARDs. A sensitivity analysis was performed between patients utilized in the study and the subset of patients who did not have sufficient parameters to calculate the DAS28 and their demographic and clinical characteristics were found to be comparable (data not shown).

The modified DAS28 (mDAS28)

In the unadjusted univariate analysis, all seven candidate predictors (TJC, SJC, PGA, PhGA, VAS, mHAQ, and morning stiffness) were found to significantly predict lnESR. A forward and backward stepwise regression analysis resulted in the same multivariable model with the following significant predictors: TJC, SJC, mHAQ, PhGA, and VAS (Table 4.2).

The PGA and duration of morning stiffness were significant in the unadjusted model but became insignificant when entered into the multivariable

model. Upon examination of the functional associations between the lnESR and several of the candidate predictors, transformations of TJC and SJC were performed to better fit the assumption of linearity. The multivariable model was refit after transforming the TJC and SJC to their logarithmic forms. A separate model was fit using TJC and SJC in the forms of log of (TJC+1) and (SJC+1) due to values of 0. No differences in the amount of variance were explained by these models were found when compared to the original model. In another series of models, TJC and SJC were both transformed to their square roots and refit in the multivariable model. Again no difference was found with the amount of variance explained by this model compared to the original model. It was decided to use the model containing the square roots of TJC and SJC since the transformed forms could be combined with the squared forms of TJC and SJC that were already present in the DAS28 formula. Every possible interaction between the variables was also explored but no significant interactions were found. The final model consisted of the five significant predictors of lnESR: TJC, SJC, mHAQ, VAS and PhGA. The regression equation for the lnESR was as follows:

$$lnESR = 2.42 - (0.0378 * \sqrt{28TJC}) + (0.0401 * \sqrt{28SJC}) + (0.35 * mHAQ) + (0.0014 * VAS) + (0.0077 * PhGA)$$

The possibility of multi-collinearity between the significant predictors was explored using the variance inflation factor (VIF) and no collinearity was indicated. All VIF values were less then 2.0 (range: 1.40-1.94).

Imputation of the fitted regression equation of the lnESR into the DAS28 formula in place of the observed lnESR resulted in the following modified version of the DAS28:

$$mDAS28 = 0.56 * \sqrt{(28TJC)} + 0.28 * \sqrt{(28SJC)} + 0.70\ \mathbf{[2.42 - (0.0378 * \sqrt{28TJC}) + (0.0401 * \sqrt{28SJC}) + (0.35 * mHAQ) + (0.0014 * VAS) + (0.0077 * PhGA)]} + 0.014 * PGA$$

The formula was simplified to its final form by combining the squared TJC and SJC terms:

$$mDAS28 = 0.534 * \sqrt{(28TJC)} + 0.31 * \sqrt{(28SJC)} + 0.245 * mHAQ + 0.001 * VAS + 0.005 * PhGA + 0.014 * PGA + 1.694$$

Validation of the mDAS28

Measure of disease activity

Distributional properties. The distributions of mDAS28 scores and the DAS28 scores differed, with the DAS28 being normally distributed and the

mDAS28 exhibiting right-skewness (Figure 4.2). The SDAI and CDAI values were also right-skewed as has been reported in several other studies (30, 31).

The means (SD) of the DAS28 and mDAS28 were 3.42 (1.54) and 3.41 (1.38), respectively in the cross-sectional development dataset and were similar in the two validation datasets (Table 4.1). Upon further examination, it was found that both the mDAS28 and the DAS28 were almost identical in detecting remission and low disease activity. Using the DAS28: 894 (16%) of patients had scores of <= 3.2 and >2.6 (low disease) and 1,871 (34%) had scores of < =2.6 (remission). When the mDAS28 was used: 915 (17%) of patients had scores of <=3.2 and >2.6 and 1,939 (35%) had scores of <=2.6 (Table 4.3).

When the CDAI was used to classify patients, a greater proportion of patients were classified into low disease and fewer into remission compared to the DAS28 and mDAS28. The proportion of patients classified into remission and low disease activity by the SDAI was similar to that of the CDAI (Table 4.3).

Criterion validity. On a group level, the mDAS28 was strongly correlated with the DAS28, SDAI and CDAI (r=0.87, 0.91 and 0.96, respectively). All correlations were significant at P<0.0000.

Predictive Validity. The mDAS28 was significantly correlated with the mHAQ, (r= 0.58, P< 0.0000). The DAS28, CDAI and SDAI were also found to be significantly correlated with the mHAQ but not as strongly (r= 0.51, r= 0.51, r= 0.51, P< 0.0000).

Ability to discriminate. To determine the ability of the mDAS28 to classify individual patients by disease level, weighted kappa coefficients were used and indicated strong agreement between the mDAS28 with the DAS28 (κ= 0.70). Kappa values >0.60 indicate a substantial relationship (32). There was strong agreement between the mDAS28 and the CDAI (κ=0.77) and between the mDAS28 and the SDAI (κ =0.71). Similar results were found between the DAS28 and the CDAI and the DAS28 and the SDAI (κ =0.62, κ =0.63, respectively).

Measure of response to treatment

Ability to discriminate. Substantial agreement between the EULAR and the mEULAR when classifying individual patients was found (κ= 0.74). However, only a moderate agreement was found when the mEULAR was compared to the CDAI response criteria (κ= 0.52.) and the SDAI response criteria (κ= 0.52). Moderate agreements were also found when the EULAR response criteria was compared with the CDAI (κ= 0.46) and the SDAI response criteria (κ= 0.47).

Sensitivity to change. The mean change in scores of the mDAS28 and the DAS28 from the baseline initiation visit to the follow-up visit were similar (Table 4). The mDAS28 was the most sensitive measure to detect change over time compared to the DAS28, CDAI and SDAI. The mDAS28 had moderate ES and SRM values (0.50, 0.58) while the DAS28 and CDAI both had moderate SRM values (0.57, 0.52) but small ES values (0.47, 0.45). The SDAI was the weakest

measure to detect change with ES and SRM values of 0.37 and 0.45, respectively.

4.5 Discussion

This study demonstrates that a modified version of the DAS28 calculated without the ESR, the mDAS28, performs as well as the DAS28 as both a measure of disease activity and response and could be used as an alternative measure to the DAS28 in epidemiological research when missing ESR values are unavailable.

Measures such as the DAS28 have been used successfully in clinical trials where the goal was to measure the efficacy of therapies by comparing groups of patients. In this study, the mDAS28 was strongly correlated with the DAS28, and also strongly correlated with the SDAI and CDAI on a group level. In addition, when compared to the other disease activity indices, had the strongest association with the mHAQ (r=0.58). This would be expected given that the mHAQ is a component of the mDAS28. Makinen *et al* (12) noted when developing the Mean Overall Index for Rheumatoid Arthritis (MOI-RA) that one of the limitations of the DAS28 was that it did not contain the Health Assessment Questionnaire (HAQ), (33) considered the best predictor of outcomes in RA (34-36). In several studies comparing the HAQ with the mHAQ, both measures were strongly correlated (37) and sensitive to change of treatment (38-40). Since a

measure should have face validity, the addition of the mHAQ as part of the mDAS28 strengthens the overall credibility of the measure.

The mean baseline values of the mDAS28 were almost identical to the mean baseline values of the DAS28 in all 3 cohorts (cross-sectional development: 3.41 (1.38) vs. 3.42 (1.57); cross-sectional validation: 3.40(1.37) vs. 3.41(1.54); longitudinal validation: 4.21(1.41) vs. 4.19 (1.59). When classifying proportions of individual patients into the disease states of remission, low, moderate and high disease activity, the mDAS28 performed again almost identically to the DAS28. The DAS28 classified 16% and 35% of patients into remission and low disease activity, respectively while the mDAS28 classified 16% and 34% into remission and low disease activity.

Since measurement tools need to assess individual patients, we examined the agreement of the measures when classifying individual patients according to disease levels, the mDAS28 strongly agreed with the DAS28 ($\kappa = 0.70$), despite the absence of the ESR as a component. The mDAS28 was also compared to the DAS28 when classifying patients according to their level of response using the EULAR response criteria. Strong agreement was found between the mEULAR and the EULAR ($\kappa = 0.74$). However, only moderate agreement was found when the mEULAR was compared with the CDAI and SDAI response criteria. The EULAR was also moderately in agreement with the CDAI and SDAI response criteria.

The mean change (SD) in scores of the mDAS28 and the DAS28 from baseline initiation to the follow-up visit were similar (Δ=0.698 (1.20) vs. Δ=0.732 (1.28)). In addition, the mDAS28 demonstrated similar sensitivity to detect disease activity changes after initiation of a DMARD when compared to the DAS28. This finding has important clinical implications for consistent monitoring of treatment response in the clinical setting and for observational research.

A limitation of the study is the use of only one observational data set to develop and validate the measure. Additional validations of the mDAS28 should be performed in other populations such as a clinic trial dataset. Another potential criticism of the study could be that the patients used had low to moderate disease activity. Again additional investigations of the mDAS28 using populations with greater ranges of disease levels including high disease activity should be undertaken. The reproducibility and reliability of the mDAS28, although not examined in this study, have been satisfied, based on the proven reliability of its individual components.

Our intent at modifying the DAS28 by substituting other measures for the ESR was not to diminish the importance of the ESR as a measure of RA disease activity. In effect, we suggest physicians continue to order lab measures regularly in the clinic as the ESR is an important measure of disease activity and long term outcomes. Modification of the DAS28 was done in an effort to allow the computation of a comparable measure to the DAS28 to be computed in standard care and for research settings where laboratory values such as the ESR are

either not available, or the results are available after the clinical encounter and not incorporated into the decision making process by providers.

In this observational study, we have developed a modified version of the DAS28, calculated without the ESR value and then demonstrated that it is comparable to the DAS28 when measuring RA disease activity and response. The mDAS28 was also found to be a valid outcome measure as it fulfilled most of the criterion recommended by the OMERACT initiative. The mDAS28 can be calculated when ESR values are unavailable, preventing patients from being omitted in epidemiological research using disease registries. Further testing of the mDAS28 in other patient populations is recommended.

Table 4.1 Demographic and clinical characteristics in cross-sectional and longitudinal cohorts*

	Cross-sectional		Longitudinal†
	Development	Validation	Validation
Patients (n)	5729	5578	367
Age, years	60.1 (13.7)	60.3 (13.8)	58.7 (12.5)
Gender, % female	75.8	75.3	77.9
Race, % Caucasian	82.7	82.5	85.3
Rheumatoid factor (% positive)	67.8	70	82
Disease duration, years	11.3 (10.1)	11.1 (9.9)	11.8 (10.0)
Disease activity characteristics			
Tender joints 0-28	3.37 (5.54)	3.38 (5.4)	5.65 (6.10)
Swollen joints 0-28	3.89 (5.54)	3.86 (5.53)	5.98 (5.50)
ESR (mm; normal <20)	24.6 (22.3)	24.6 (22.1)	25.9 (22.6)
CRP (mg/dL; normal <1.0)	2.96 (8.5)	2.96 (8.7)	2.8 (7.8)
Pain (VAS) assessment 0-100	32.2 (26.6)	31.6 (25.9)	39.4 (26.7)
mHAQ, 0-3	0.40 (0.49)	0.39 (0.48)	0.50 (0.52)
Patient global assessment, 0-100	29.9 (25.9)	29.8 (25.9)	37.9 (26.84)
Physician global assessment, 0-100	19.5 (19.0)	19.4 (19.2)	29.4 (20.3)
Duration of Stiffness (hours)	1.03 (2.4)	0.94 (2.11)	1.34 (3.11)
Disease activity composite measures			
mDAS28	3.41 (1.38)	3.40 (1.37)	4.21 (1.41)
DAS28	3.42 (1.54)	3.41 (1.54)	4.19 (1.59)
SDAI	15.0 (16.1)	14.9 (16.0)	20.5 (17.0)
CDAI	12.2 (12.0)	12.1(12.0)	18.4 (13.0)

*Values are the mean (SD) unless otherwise indicated; † Initiators of disease-modifying anti-rheumatic drugs (DMARDs); ESR= erythrocyte sedimentation rate; CRP=C-reactive protein; VAS= visual analogue scale; mHAQ= modified Health Assessment Questionnaire score; mDAS28= modified Disease Activity Score with 28 joint count; DAS28= Disease Activity Score with 28 joint count; SDAI=Simplified Disease Activity Index; CDAI= Clinical Disease Activity Index.

Table 4.2 Results of Forward and Backwards Stepwise Linear Regressions

Predictors	Coefficient (β)	Standard Error	P>|t|	95% CI
Disability Index (mHAQ)	0.345	0.033	0.000	(0.280 to 0.411)
Physician Global Assessment	0.077	0.001	0.000	(0.006 to 0.009)
Swollen Joint Score	0.041	0.011	0.000	(0.019 to 0.062)
Patient VAS for pain	0.001	0.001	0.033	(0.000 to 0.003)
Tender Joint Score	-0.037	0.013	0.004	(-0.064 to -0.012)
Constant	2.423	0.022	0.000	(2.378 to 2.467)

mHAQ= modified Health Assessment Questionnaire; VAS= Visual analog scale

Table 4.3 Proportion of patients classified in disease levels using composite indices in the cross-sectional validation cohort (n=5578)*

Measure	Remission N (%)	Low N (%)	Moderate N (%)	High N (%)
mDAS28	1939 (35)	915 (17)	1,936 (35)	732 (13)
DAS28	1871 (34)	894 (16)	1,970 (35)	843 (15)
CDAI	1309 (23)	1874 (34)	1,387 (25)	1,008 (18)
SDAI†	618 (20)	1050 (34)	852 (28)	537 (18)

*Values are numbers (%); mDAS28= modified Disease Activity Score with 28 joint count; DAS28=Disease Activity Score with 28 joint count; CDAI= Clinical Disease Activity Index; SDAI= Simplified Disease Activity Index;
†For the SDAI, n=3057 had SDAI scores.

Table 4.4 Sensitivity to change assessed by effect size (ES) and standardized response mean (SRM)*

Measure	Baseline Mean (SD)	Follow-up Mean (SD)	Change Mean (SD)	ES	SRM
mDAS28	4.21 (1.41)	3.50 (1.31)	0.698 (1.20)	0.50	0.58
DAS28	4.22 (1.55)	3.48 (1.48)	0.732 (1.28)	0.47	0.57
CDAI	18.37 (13.05)	12.49 (11.19)	5.88 (11.28)	0.45	0.52
SDAI	20.54 (17.04)	14.23 (11.75)	6.33 (14.17)	0.37	0.45

*Values are the mean (SD) unless otherwise indicated; mDAS28= modified Disease Activity Score with 28 joint count; DAS28=Disease Activity Score with 28 joint count; CDAI= Clinical Disease Activity Index; SDAI= Simplified Disease Activity Index.

Value of composite measure at endpoint	Improvement in composite measure from baseline		
	>1.2†	>0.6 and ≤1.2	≤0.6
≤3.2*	Good	Moderate	None
>3.2 and ≤5.1			
>5.1			

* DAS28 absolute cut points; †DAS28 change cut points;

Figure 4.1 Algorithm to calculate the EULAR Response Criteria using published absolute and change cut points.

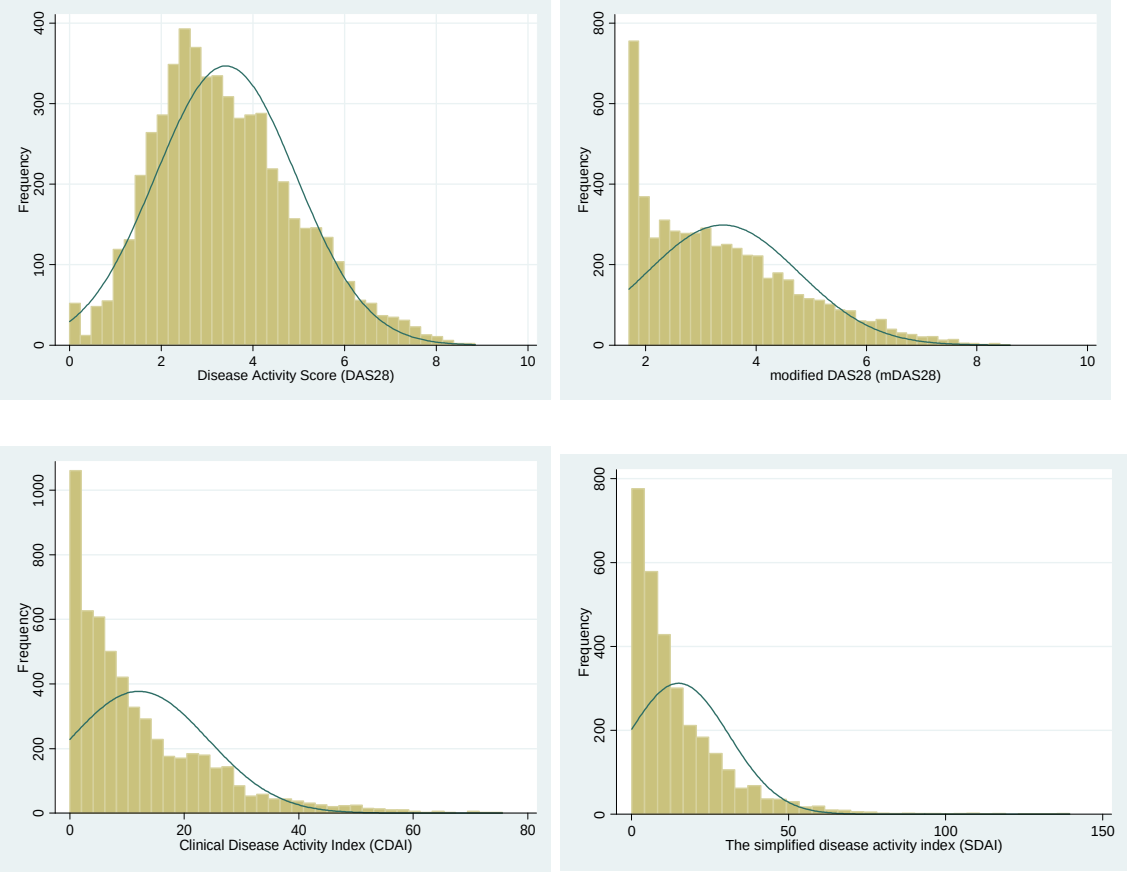


Figure 4.2 Distribution properties of composite disease activity indices in the cross-sectional validation cohort (n=5578).

CHAPTER V

Final Conclusions and General Discussions

The research presented in this dissertation is focused on improving clinical management of rheumatoid arthritis and epidemiological research efforts by evaluating and developing clinical measures of disease activity and treatment response to prevent further disease progression. Factors predictive of physician practice behaviors specifically in the measurement of acute phase reactants and prescription of disease-modifying antirheumatic drug (DMARD) therapy, were also elucidated in this work. The first study in this thesis showed that the two measures of disease activity, the Clinical Disease Activity Index (CDAI) (1) and the Disease Activity Score with 28 joint count (DAS28) (2, 3) are valid measures for use in clinical care to assess disease activity but not to guide DMARD treatment decisions. Further, several patient and clinic characteristics were found to be predictive of physician DMARD prescribing behavior and the variability in the prescription of DMARDs by rheumatologists was quantified. The second study elucidated factors predictive of the measurement of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). These factors included ordering of other laboratory tests, the number of previous DMARDs prescribed, patient age, no insurance, physician propensity to order ESR or CRP and private clinic. While a previous study (4) reported that physician practice style influenced ESR and CRP measurement, our results found that individual physician characteristics had little to do with their measurement but patient and clinic characteristics were very influential. In addition, disease activity was found not to be associated with a physician's decision to measure ESR or CRP. In the third

study, a new measure of disease activity, the modified DAS28 was developed and validated. The results of the study showed that it was a comparable measure to the DAS28 and could be used in epidemiological research studies when the DAS28 was not calculable due to missing ESR values.

Validation of CDAI and DAS28 as Measures to Guide Clinical Treatment Decisions

Previous studies have examined the performance of the CDAI and the DAS28 for use in the clinical setting to consistently monitor disease activity and treatment response to ensure optimal treatment effectiveness (5-13). Yet, whether the CDAI and DAS28 can be used to guide treatment decisions in the clinic has not been illuminated, to the best of my knowledge in the literature. The association of CDAI and the rheumatologist's decision to change DMARD therapy was examined using a multi-level mixed logistic regression modeling strategy to accommodate the hierarchical data structure of the data. The DAS28 as a predictor of treatment decision was also examined in a second analysis. Correlates that might influence the rheumatologist to change DMARD therapy were identified and variability in the rheumatologists' decisions to change DMARD therapy was quantified. The study showed that the CDAI and DAS28 were both significantly associated with the physicians' decision to change DMARD therapy. Correlates that were significantly predictive of change in treatment included nonbiologic DMARD initiation at baseline, number of past

DMARDS prescribed, patient age, duration of disease, employment status, physician age and academic clinic. The effect of patient characteristics on the rheumatologists' decision to change DMARD therapy was found to be substantial with little variation related to clinic or individual physicians. Based on these results, future interventions to improve DMARD prescription behavior should be targeted at the heterogeneity of patients and not physicians or clinics. The abilities of the CDAI and DAS28 to discriminate a change in DMARD therapy were fair with an AUC of 0.69 and 0.72, respectively. Thus, the study could not validate the CDAI and DAS28 as measures to guide treatment in the clinic.

ESR and CRP Measurement Not Associated with Disease Activity

Acute phase reactants (APRs), ESR and CRP, long used in the clinic to measure disease activity in rheumatoid arthritis were recently investigated and concerns over their appropriateness as measures of disease activity were raised (14). In addition, variability in their measurement in the clinic has been reported, impeding the calculation of the DAS28 and SDAI. To examine why variability in the measurement of these APRs exists, correlates that may be influencing measurement of ESR and CRP in the clinic, were examined. Other laboratory tests ordered at the same clinical encounter, number of past DMARDs prescribed, type of clinic and insurance status were the strongest predictors of the measurement of ESR and CRP. Remarkably, since ESR and CRP testing has been long utilized to monitor patient's disease activity, no quantitative

measures of disease activity were significantly predictive of the measurement of ESR or CRP. This finding supports the results of the recent study noted above and warrants the further investigation of the ESR and CRP as measures of disease activity. Variability in the measurement of ESR and CRP was found to be due to unmeasured clinic and patient characteristics and not the individual physician, in contrast to the results from a prior study (4) that found measurement variation was due to the individual physician. In addition, large variation in the measurement of ESR and CRP was found in our study again in contrast with another prior study that found little variation (15).

A New Simplified Measure of Disease Activity and Treatment Response

Since the variability in the measurement of ESR and CRP may impede the calculation of the DAS28 and the SDAI and limit their use in the clinic, we developed and validated a new simplified composite measure. The measure was developed by identifying calculable disease activity measures that might significantly predict the value of ESR. These predictive factors were then imputed into the DAS28 formula in place of the ESR value. The new measure, named the modified DAS28 (mDAS28), was then validated as a measure of disease activity and treatment response. Our findings demonstrated that the performance of this new measure, mDAS28, was comparable with the DAS28 and we concluded that the mDAS28 could be used in the clinical setting to consistently monitor disease activity and response to treatment when the DAS28 could not be calculated.

Summary

In summary, the three studies detailed in this dissertation have focused on improving the clinical management of patients with rheumatoid arthritis and epidemiological research efforts in RA. The findings that two composite measures, CDAI and DAS28, cannot guide treatment decisions in the clinic, that two established disease activity measures, ESR and CRP are not associated with disease activity and that the DAS28 can be modified by replacing the ESR with calculable clinical variables and its modified version validated, have significantly advanced our knowledge of disease activity and treatment response measurement in the clinic. Elucidation of these aspects of disease activity measurement may enable the improvement in the management of rheumatoid arthritis in the clinic setting, facilitate the inclusion of more patients in epidemiological research studies and ultimately improve the disease outcomes for patients.

REFERENCES

Chapter 1

1. St. Clair EW, Pisetsky DS, Haynes BF. *Rheumatoid arthritis*. Philadelphia: Lippincott, Williams & Wilkins; 2004.

2. American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis: 2002 update. *Arthritis Rheum*. 2002; 46(2): 328-46.

3. Grassi W, De Angelis R, Lamanna G, Cervini C. The clinical features of rheumatoid arthritis. *Eur J Radiol*. 1998;27 (Suppl 1):S18-24.

4. Aletaha D, Stamm T, Smolen JS. Measuring disease activity for rheumatoid arthritis. *Z Rheumatol*. 2006.

5. Albers JM, Kuper HH, van Riel PL, Prevoo ML, van 't Hof MA, van Gestel AM, et al. Socio-economic consequences of rheumatoid arthritis in the first years of the disease. *Rheumatology (Oxford).* 1999; 38(5): 423-30.

6. Smolen JS, Sokka T, Pincus T, Breedveld FC. A proposed treatment algorithm for rheumatoid arthritis: Aggressive therapy, methotrexate, and quantitative measures. *Clin Exp Rheumatol* 2003; 21(Suppl. 31): S209-10.

7. Sokka TM, Makinen H. Improving outcomes in rheumatoid arthritis: What determines decisions to change ineffective therapy? *J Rheumatol*. 2006; 33(7):1213-5.

8. Smolen JS, Aletaha D. What should be our treatment goal in rheumatoid arthritis today? *Clin Exp Rheumatol*. 2006;24(Suppl 43):S7-S13.

9. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR *et al.* American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Rheum*. 2008; 59(6):762-84.

10. Smolen JS, Aletaha D, Machold KP. Therapeutic strategies in early rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2005;19(1):163-77.

11. Smolen JS, Aletaha D, Koeller M, Weisman MH, Emery P. New therapies for treatment of rheumatoid arthritis. *Lancet*. 2007; 370(9602):1861-74.

12. Albers JM, Paimela L, Kurki P, Eberhardt KB, Emery P, van 't Hof MA, et al. Treatment strategy, disease activity, and outcome in four cohorts of patients with early rheumatoid arthritis. *Ann Rheum Dis*. 2001; 60(5): 453-8.

13. Allaart CF, Goekoop-RuitermanYP, de Vries-Bouwstra JK, Breedveld FC, Dijkmans BA. FARR study group. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: The BeSt study. *Clin Exp Rheumatol*. 2006;24(6 Suppl 43): 57-82.

14. Breedveld FC, Kalden JR. Appropriate and effective management of rheumatoid arthritis. *Ann Rheum Dis*. 2004; 63(6): 627-33.

15. Khanna D, Oh M, Furst DE, Ranganath V, Gold RH, Sharp JT *et al.* Evaluation of the preliminary definitions of minimal disease activity and remission in an early seropositive rheumatoid arthritis cohort*. Arthritis Rheum* 2007; 57 (3): 440-7.

16. Ranganath VK, Yoon J, Elashoff DA, Park GS, Khanna D, Furst DE *et al*. Comparison of composite measures of disease activity in an early seropositive rheumatoid arthritis cohort. *Ann Rheum Dis* 2007; 66: 1633-40.

17. Schoels M, Kapral T, Stamm T, Smolen JS, Aletaha D. Step-up combination versus switching of non-biological disease-modifying antirheumatic drugs in rheumatoid arthritis: Results from a retrospective observational study. *Ann Rheum Dis.* 2007; 66(8):1059-65.

18. Bentley MJ, Reed GW. Simplified composite disease activity measures in rheumatoid arthritis: Should they be used in standard care? *Clin Exp Rheumatol* 2008; 26(2):358-66.

19. van Riel PL, Schumacher HR Jr. How does one assess early rheumatoid arthritis in daily clinical practice? *Best Practice Res Clin Rheumatol* 2001;15(1): 67-76.

20. Pincus T, Sokka T. Quantitative measures and indices to assess rheumatoid arthritis in clinical trials and clinical care. *Rheum Dis Clin North Am*. 2004; 30(4):725-51.

21. Blackburn WD,Jr. Validity of acute phase proteins as markers of disease activity. *J Rheumatol Suppl* 1994; 42: 9-13.

22. Brigden ML. Clinical utility of the erythrocyte sedimentation rate. *Am Fam Physician.* 1999; 60(5):1443-50.

23. Kushner I. C-reactive protein in rheumatology. *Arthritis Rheum*. 1991; 34(8):1065-8.

24. Cheah SY, Clark C, Goldberg L, Li Wan Po A, Phillips R. Outcome measures, pooled index and quality of life instruments in rheumatoid arthritis. *J Clin Pharm Ther* 1996; 21(5): 297-316.

25. Aletaha D, Smolen JS. The definition and measurement of disease modification in inflammatory rheumatic diseases. *Rheum Dis Clin North Am*. 2006; 32(1):9-44.

26. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *J Rheumatol*. 1997; 24(8):1477-85.

27. Walsh L, Davies P, McConkey B. Relationship between erythrocyte sedimentation rate and serum C-reactive protein in rheumatoid arthritis. *Ann Rheum Dis*. 1979; 38(4): 362-3.

28. Felson D T, Anderson J J, Boers M, Bombardier C, Chernoff M, Fried B *et al*. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The committee on outcome

measures in rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993; 36(6): 729-40.

29. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis*. Arthritis Rheum* 1980; 23(2): 137-45.

30. Boers M, Tugwell P. The validity of pooled outcome measures (indices) in rheumatoid arthritis clinical trials. *J Rheumatol* 1993; 20(3): 568-74.

31. Roberts RS. Pooled outcome measures in arthritis: The pros and cons. *J Rheumatol*. 1993; 20(3): 566-7.

32. Goldsmith CH, Smythe HA, Helewa A. Interpretation and power of a pooled index. *J Rheumatol*. 1993; 20(3): 575-8.

33. Fuchs HA. The use of the disease activity score in the analysis of clinical trials in rheumatoid arthritis. *J Rheumatol*. 1993; 20(11):1863-6.

34. Felson, D T Anderson, J J Boers, M Bombardier, C Furst, D Goldsmith, C Katz, L M Lightfoot, R Paulus, H Strand,V. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*. 1995; 38(6): 727-35.

35. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk M H, 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(1): 34-40.

36. Scott DL. A simple index to assess disease activity in rheumatoid arthritis. *J Rheumatol*. 1993; 20(3): 582-4.

37. van der Heijde DM, nan't Hof MA, van Riel PL, van . Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol*. 1993; 20(3): 579-81.

38. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. *Ann Rheum Dis* 1990;49 (11): 916-20.

39. 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(1): 44-8.

40. Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie GL *et al*. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum*. 1995; 38(1): 38-43.

41. Soubrier M, Dougados M. How to assess early rheumatoid arthritis in daily clinical practice. *Best Pract Res Clin Rheumatol*. 2005;19(1):73-89.

42. DAS-score NL. home of the DAS: Disease activity score in rheumatoid arthritis [homepage on the Internet]. Available from: http://www.das-score.nl/www.das-score.nl/indfex.html.

43. Aletaha D, Stamm T, Smolen JS. Validation of the simplified disease activity index (SDAI) in an observational cohort of patients with rheumatoid arthritis. *Ann Rheum Dis*. 2004; 63:111.

44. Aletaha D, Smolen, JS. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23(Suppl 39): S100-S108.

45. Greenberg JD, Harrold LR, Bentley MJ, Kremer J, Reed G, Strand V. Evaluation of composite measures of treatment response without acute-phase reactants in patients with rheumatoid arthritis. *Rheumatology (Oxford).* 2009; 48(6): 686-90.

46. Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of disease activity score (DAS28)- erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values. *Ann Rheum Dis*. 2007 Mar;66(3):407-9.

47. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T *et al*. Disease activity measurement of rheumatoid arthritis: Comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum* 2005; 53(1): 56-60.

48. van Gestel AM, Anderson JJ, van Riel PL, Boers M, Haagsma CJ, Rich B *et al.* ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American College of Rheumatology European league of associations for rheumatology. *J Rheumatol*. 1999; 26(3):705-11.

49. Pincus T. Advantages and limitations of quantitative measures to assess rheumatoid arthritis: Joint counts, radiographs, laboratory tests, and patient. *Bull NYU Hosp Jt Dis.* 2006; 64(1-2): 32-9.

50. Donald F, Ward MM. Evaluative laboratory testing practices of United States rheumatologists. *Arthritis Rheum*. 1998; 41(4): 725-9.

51. Henke CJ, Epstein WV. Practice variation in rheumatologists' encounters with their patients who have rheumatoid arthritis. *Med Care*. 1991 Aug; 29(8):799-812.

52. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. *J Rheumatol*. 1994;21(7):1227-37.

53. 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(4): R796-R806.

54. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Ebert G, van Riel PL, Tugwell P. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford).* 2003; 42(2): 244-57.

55. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R *et al*. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): A single-blind randomised controlled trial. *Lancet* 2004; 364:263-9.

56. Fransen J, Moens HB, Speyer I, van Riel PL. Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: A multicentre, cluster randomised controlled trial. *Ann Rheum Dis.* 2005; 64(9):1294-8.

57. Furst DE, Breedveld FC, Kalden JR, Smolen JS, Burmester GR, Emery P *et al*. Updated consensus statement on biological agents for the treatment of rheumatic diseases. *Ann Rheum Dis*. 2006 Nov; 65 (Suppl 3): S2-15.

58. Smolen JS, Sokka T, Pincus T, Breedveld FC. A proposed treatment algorithm for rheumatoid arthritis: Aggressive therapy, methotrexate, and quantitative measures. *Clin Exp Rheumatol*. 2003 ;21 (Suppl 31): S209-10.

59. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, Smolen JS, Khaltaev N, Muirden KD. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Supple* 1994; 41: 86-9.

60. Smolen JS. The work of the EULAR standing committee on international clinical studies including therapeutic trials (ESCISIT). *Br J Rheumatol*. 1992;31 (4): 219-20.

61. van der Heijde DM, van't Hof MA, van Riel PL, Van de Putte L. Disease activity score. *Ann Rheum Dis* 1992;51(1):140.

62. Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: Defining criteria for disease activity states. *Arthritis Rheum*. 2005 Sep; 52(9): 2625-36.

63. Wolfe F, Michaud K, Pincus T, Furst D, Keystone E. The disease activity score is not suitable as the sole criterion for initiation and evaluation of anti-tumor necrosis factor therapy in the clinic: Discordance between assessment measures and limitations in questionnaire use for regulatory purposes. *Arthritis Rheum*. 2005 Dec; 52(12): 3873-9.

64. Boers M, Brooks P, Simon LS, Strand V, Tugwell P. OMERACT: An international initiative to improve outcome measurement in rheumatology. *Clin Exp Rheumatol* 2005;23 (Suppl 39): S10-S13.

65. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998; 25(2): 198-9.

66. Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F *et al*. DAS28 best reflects the physician's clinical judgment of response

to infliximab therapy in rheumatoid arthritis patients: Validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther* 2005; 7(5): R1063-R1071.

67. Gülfe A, Geborek P, Saxne T. Response criteria for rheumatoid arthritis in clinical practice: How useful are they? *Ann Rheum Dis* 2005; 64(8): 1186-89.

68. Smolen JS, Van Der Heijde DM, St Clair EW, Emery P, Bathon JM, Keystone E *et al.* Predictors of joint damage in patients with early rheumatoid arthritis treated with high-dose methotrexate with or without concomitant infliximab: Results from the ASPIRE trial. *Arthritis Rheum* 2006; 54(3): 702-10.

69. Mierau M, Schoels M, Gonda G, Fuchs J, Aletaha D, Smolen JS. Assessing remission in clinical practice. *Rheumatology (Oxford)* 2007 Jun; 46(6): 975-9.

70. Soubrier M, Zerkak D, Gossec L, A?yral X, Roux C, Dougados M. Which variables best predict change in rheumatoid arthritis therapy in daily clinical practice? *J Rheumatol* 2006; 33(7): 1243--6.

71. Leeb BF, Andel I, Leder S, Leeb BA, Rintelen B. The patient's perspective and rheumatoid arthritis disease activity indexes. *Rheumatology (Oxford*) 2005; 44(3): 360-5.

72. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.

73. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: Development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. OMERACT committee. *J Rheumatol* 1993; 20(3): 561-5.

74. van Leeuwen MA, van der Heijde DM, van Rijswijk MH, Houtman PM, van Riel PL, van de Putte LB *et al.* Interrelationship of outcome measures and process variables in early rheumatoid arthritis. A comparison of radiologic damage, physical disability, joint counts, and acute phase reactants. *J Rheumatol* 1994; 21(3): 425-9.

75. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39(4): 561-77.

76. Landewe R, van der Heijde D, van der Linden S, Boers M. Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: A comparison with the original DAS remission. *Ann Rheum Dis* 2006; 65(5): 637-41.

77. van der Heijde DM, Jacobs JW. The original "DAS" and the "DAS28" are not interchangeable: Comment on the articles by Prevoo *et al. Arthritis Rheum* 1998; 41(5): 942-5.

78. Kapral T, Dernoschnig F, Machold KP, Stamm T, Schoels M, Smolen JS *et al.* Remission by composite scores in rheumatoid arthritis: Are ankles and feet important? *Arthritis Res Ther* 2007; 9(4): R72.

79. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E *et al*. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: Evidence from an imaging study may explain structural progression. *Arthritis Rheum* 2006; 54(12): 3761-73.

80. Makinen H, Kautiainen H, Hannonen P, Sokka T. Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis? *Ann Rheum Dis* 2005; 64(10): 1410-3.

81. Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: Statistical and judgmental approaches. *J Rheumatol* 1982; 9(5): 753-7.

82. Boers M. International consensus on which measures to use in rheumatoid arthritis clinical trials. *Neth J Med* 1993; 43(1-2): 55-8.

83. Cuomo G, Molinaro G, La Montagna G, Migliaresi S, Valentini G. A comparison between the simplified disease activity index (SDAI) and the disease activity score (DAS28) as measure of response to treatment in patients undergoing different therapeutic regimens. *Reumatismo*. 2006; 58(1): 22-5.

84. Smolen JS. A comparison of the SDAI and DAS28 as measures of response in adalimumab (HUMIRA (TM)) clinical trials in rheumatoid arthritis (RA). *Arthritis Rheum* 2003; 48(9): S107.

85. Keenan RT, Swearingen CJ, Yazici Y. Erythrocyte sedimentation rate and C-reactive protein levels are poorly correlated with clinical measures of disease activity in rheumatoid arthritis, systemic lupus erythematosus and osteoarthritis patients. *Clin Exp Rheumatol*. 2008 Sep-Oct;26(5):814-9.

Chapter 2

1. Pincus T, Sokka T. Complexities in the quantitative assessment of patients with rheumatic diseases in clinical trials and clinical care. *Clin Exp Rheumatol*. 2005; 23(Suppl 39): S1-S9.

2. Pincus T, Yazici Y, Sokka T. Complexities in assessment of rheumatoid arthritis: Absence of a single gold standard measure. *Rheum Dis Clin North Am*. 2009;35(4): 687-97.

3. American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis: 2002 update. *Arthritis Rheum*. 2002 Feb; 46(2): 328-46.

4. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Rheum*. 2008; 59(6): 762-84.

5. Ward MM. Clinical and laboratory measures. In: E. William St. Clair, David S. Pisetsky, Barton F. Haynes, editor. *Rheumatoid arthritis.* Philadelphia: Lippincott Williams & Wilkins; 2004. p. 51-63.

6. Landewe RB, van der Heijde DM. Principles of assessment from a clinical perspective. *Best Pract Res Clin Rheumatol*. 2003;17(3): 365-79.

7. Criswell LA, Such CL, Neuhaus JM, Yelin EH. Variation among rheumatologists in clinical outcomes and frequency of office visits for rheumatoid arthritis. *J Rheumatol*. 1997; 24(7): 1266-71.

8. Criswell LA, Henke CJ. What explains the variation among rheumatologists in their use of prednisone and second line agents for the treatment of rheumatoid arthritis? *J Rheumatol*. 1995; 22(5): 829-35.

9. Aletaha D, Stamm T, Smolen JS. Measuring disease activity for rheumatoid arthritis. *Z Rheumatol*. 2006.

10. Grigor C, Capell H, Stirling A, McMahon A, Lock P, Vallance R *et al.* Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): A single-blind randomised controlled trial. *Lancet* 2004; 364:263-9.

11. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49(11): 916-20.

12. van der Heijde DM, van't Hof MA, van Riel PL, Van de Putte L. Disease activity score. *Ann Rheum Dis* 1992; 51(1): 140.

13. 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(4): R796-R806.

14. 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(1): 44-8.

15. Cuomo G, Molinaro G, La Montagna G, Migliaresi S, Valentini G. A comparison between the simplified disease activity index (SDAI) and the disease activity score (DAS28) as measure of response to treatment in patients undergoing different therapeutic regimens. *Reumatismo.* 2006;58(1): 22-5.

16. Fransen J, Creemers MC, Van Riel PL. Remission in rheumatoid arthritis: Agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford).* 2004; 43(10): 1252-5.

17. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T *et al.* Disease activity measurement of rheumatoid arthritis: Comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum* 2005; 53(1): 56-60.

18. Makinen H, Kautiainen H, Hannonen P, Mottonen T, Korpela M, Leirisalo-Repo M, et al. Disease activity score 28 as an instrument to measure disease activity in patients with early rheumatoid arthritis. *J Rheumatol*. 2007; 34(10): 1987-91.

19. Makinen H, Kautiainen H, Hannonen P, Sokka T. Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis? *Ann Rheum Dis*. 2005; 64(10): 1410-3.

20. Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F *et al.* DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: Validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther* 2005; 7(5): R1063-R1071.

21. Bentley MJ, Reed GW. Simplified composite disease activity measures in rheumatoid arthritis: Should they be used in standard care? *Clin Exp Rheumatol* 2008; 26(2): 358-66.

22. Greenberg JD, Harrold LR, Bentley MJ, Kremer J, Reed G, Strand V. Evaluation of composite measures of treatment response without acute-phase reactants in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2009; 48(6): 686-90.

23. Leeb BF, Haindl PM, Maktari A, Nothnagl T, Rintelen B. Patient-centered rheumatoid arthritis disease activity assessment by a modified RADAI. *J Rheumatol.* 2008; 35(7):1294-9.

24. Rintelen B, Haindl PM, Maktari A, Nothnagl T, Hartl E, Leeb BF. SDAI/CDAI levels in rheumatoid arthritis patients are highly dependent on patient's pain perception and gender. *Scand J Rheumatol*. 2008; 29:1-4.

25. Aletaha D, Smolen JS. The simplified disease activity index (SDAI) and clinical disease activity index (CDAI) to monitor patients in standard clinical care. *Best Pract Res Clin Rheumatol*. 2007; 21(4): 663-75.

26. Khanna D, Oh M, Furst DE, Ranganath V, Gold RH, Sharp JT *et al*. Evaluation of the preliminary definitions of minimal disease activity and remission in an early seropositive rheumatoid arthritis cohort. *Arthritis Rheum* 2007; 57(3): 440-7.

27. Ranganath VK, Yoon J, Elashoff DA, Park GS, Khanna D, Furst DE *et al*. Comparison of composite measures of disease activity in an early seropositive rheumatoid arthritis cohort. *Ann Rheum Dis* 2007; 66: 1633-40.

28. Makinen H, Hannonen P, Sokka T. Definitions of remission for rheumatoid arthritis and review of selected clinical cohorts and randomised clinical trials for the rate of remission. *Clin Exp Rheumatol*. 2006; 24(Suppl 43): S22-S28.

29. Ranganath VK, Khanna D, Paulus HE. ACR remission criteria and response criteria. *Clin Exp Rheumatol*. 2006; 24(Suppl 43): S014-S021.

30. Aletaha D, Smolen JS. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23 (Suppl 39): S100-S108.

31. Pepe MS. An interpretation for the ROC curve and inference using GLM procedures. *Biom.* 2000 Jun;56(2):352-9.

32. van Erkel A R, Pattynama P M. Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. *Eur J Radiol* 1998;27(2):88-94.

33. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39(4): 561-77.

34. Zweig MH. ROC plots display test accuracy, but are still limited by the study design. *Clin Chem* 1993; 39(6): 1345-6.

35. Kremer J. The CORRONA database. *Ann Rheum Dis*. 2005; 64 (Suppl 4): 37-41.

36. Diez Roux AV, Auchincloss AH. Understanding the social determinants of behaviours: Can new methods help? *Int J Drug Policy*. 2009; 20(3): 227-9.

37. Pincus T, Swearingen C, Wolfe F. Toward a multidimensional health assessment questionnaire (MDHAQ): Assessment of advanced activities of daily living and psychological status in the patient-friendly health assessment questionnaire format. *Arthritis Rheum*. 1999 Oct; 42(10): 2220-30.

38. Uhlig T, Haavardsholm EA, Kvien TK. Comparison of the health assessment questionnaire (HAQ) and the modified HAQ (MHAQ) in patients with rheumatoid arthritis. *Rheumatology (Oxford)*. 2006; 45(4): 454-8.

39. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

40. Miles, J & Shevlin, M. *Applying Regression & Correlation*. London: Sage Publications; 2001.

41. Wells GA, Boers M, Shea B, Brooks PM, Simon LS, Strand CV *et al*. Minimal disease activity for rheumatoid arthritis: A preliminary definition. *J Rheumatol*. 2005 Oct; 32(10): 2016-24.

42. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther* 2005; 85(3): 257-68.

43. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall; 1991.

44. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950; 3(1): 32-5.

45. Perkins NJ, Schisterman EF. The Youden index and the optimal cut-point corrected for measurement error. *Biom J*. 2005; 47(4): 428-41.

46. Rabe-Hesketh S. SA. *Multilevel and Longitudinal Modeling Using STATA*. Second ed. College Station, Texas: Stata Press; 2008.

47. Stata Corporation. Stata statistical software release 10.1. 2009.

48. Schoels M, Kapral T, Stamm T, Smolen JS, Aletaha D. Step-up combination versus switching of non-biological disease-modifying antirheumatic drugs in rheumatoid arthritis: Results from a retrospective observational study. *Ann Rheum Dis.* 2007 Aug; 66(8): 1059-65.

Chapter 3

1. Ward MM. Evaluative laboratory testing. assessing tests that assess disease activity. *Arthritis Rheum*. 1995; 38(11): 1555-63.

2. Ward MM. Clinical and laboratory measures. In: E. William St. Clair, David S. Pisetsky, Barton F. Haynes, editor. *Rheumatoid arthritis*. Philadelphia: Lippincott Williams & Wilkins; 2004. p. 51-63.

3. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *J Rheumatol*. 1997; 24(8):1477-85.

4. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. *J Rheumatol*. 1994; 21(7): 1227-37.

5. Pincus T. The American College of Rheumatology (ACR) core data set and derivative "patient only" indices to assess rheumatoid arthritis. *Clin Exp Rheumatol.* 2005; 23(Suppl 39): S109-13.

6. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA *et al*. Judging disease activity in clinical practice in rheumatoid

arthritis: First step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49(11): 916-20.

7. van der Heijde DM, van't Hof MA, van Riel PL, Van de Putte L. Disease activity score. *Ann Rheum Dis* 1992; 51(1): 140.

8. 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(1): 44-8.

9. Smolen JS, Breedveld FC, Schiff MA, Kalden JR, Emery P, Ebert S, van Riel PL, Tugwell P. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003; 42(2): 244-57.

10. Ranganath VK, Khanna D, Paulus HE. ACR remission criteria and response criteria. *Clin Exp Rheumatol*. 2006;24(Suppl 43): S014-21.

11. Fransen J, van Riel PL. The disease activity score and the EULAR response criteria. *Clin Exp Rheum* 2005; 23 (5 Suppl 39) S93-9.

12. 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(1): 34-40.

13. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR *et al*. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Rheum*. 2008; 59(6): 762-84.

14. Henke CJ, Epstein WV. Practice variation in rheumatologists' encounters with their patients who have rheumatoid arthritis. *Med Care*. 1991; 29(8): 799-812.

15. Donald F, Ward MM. Evaluative laboratory testing practices of United States rheumatologists. *Arthritis Rheum*. 1998; 41(4): 725-9.

16. Aletaha D, Smolen JS. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheum*. 2005; 23 (Suppl 39): S100-S108.

17. Hassell AB, Davis MJ, Fowler PD, Clarke S, Fisher J, Shadforth MF *et al.* The relationship between serial measures of disease activity and outcome in rheumatoid arthritis. *Q J Med*. 1993 Sep; 86(9): 601-7.

18. Mottonen TT. Prediction of erosiveness and rate of development of new erosions in early rheumatoid arthritis. *Ann Rheum Dis*. 1988 Aug; 47(8): 648-53.

19. Grindulis KA, Calverley M, Constable TJ, Forster PJ, Ahmed ME, McConkey B. A comparison between clinical and laboratory tests in rheumatoid arthritis. *Scand J Rheumatol* 1983;12 (3):285-8.

20. Blackburn WD, Jr. Validity of acute phase proteins as markers of disease activity. *J Rheumatol* 1994; 42: 9-13.

21. Keenan RT, Swearingen CJ, Yazici Y. Erythrocyte sedimentation rate and C-reactive protein levels are poorly correlated with clinical measures of disease activity in rheumatoid arthritis, systemic lupus erythematosus and osteoarthritis patients. *Clin Exp Rheumatol*. 2008; 26(5): 814-9.

22. Kremer J. The CORRONA database. *Ann Rheum Dis*. 2005; 64 (Suppl 4): 37-41.

23. 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(4): R796-R806.

24. Uhlig T, Haavardsholm EA, Kvien TK. Comparison of the health assessment questionnaire (HAQ) and the modified HAQ (MHAQ) in patients with rheumatoid arthritis*. Rheumatology (Oxford)*. 2006; 45(4): 454-8.

25. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): Analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol*. 2001; 28(5): 982-9.

26. Brigden ML. Clinical utility of the erythrocyte sedimentation rate. *Am Fam Physician.* 1999; 60(5):1443-50.

27. Stata Corporation. Stata statistical software release 10.1. 2009.

28. Rabe-Hesketh S. SA. *Multilevel and Longitudinal Modeling Using STATA*. Second ed. College Station, Texas: Stata Press; 2008.

29. Larsen K, Petersen JH, Budtz-Jorgensen E, Endahl L. Interpreting parameters in the logistic regression model with random effects. *Biom* 2000; 56(3): 909-14.

30. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: Integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol*. 2005; 161(1): 81-8.

31. Miles, J & Shevlin, M. *Applying Regression & Correlation*. London: Sage Publications; 2001.

32. Box G.E Transformation of the independent variables. *Technometrics.* 1962; 4: 531-50.

33. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39(4): 561-77.

34. Pepe MS. An interpretation for the ROC curve and inference using GLM procedures. *Biom* 2000; 56(2): 352-9.

35. Sokka T, Pincus T. Erythrocyte sedimentation rate, C-reactive protein, or rheumatoid factor are normal at presentation in 35%-45% of patients with

rheumatoid arthritis seen between 1980 and 2004: Analyses from Finland and the United States. *J Rheumatol*. 2009 Jul; 36(7):1387-90.

36. Fraenkel L, Rabidou N, Dhar R. Are rheumatologists' treatment decisions influenced by patients' age? *Rheumatology (Oxford)* 2006; 45(12): 1555-7.

Chapter 4

1. 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.

2. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49: 916-920.

3. van der Heijde DM, van 't Hof MA, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol* 1993; 20: 579-581.

4. 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.

5. Prevoo ML, van Gestel AM, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Remission in a prospective study of patients with rheumatoid

arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score. *Br J Rheum* 1996; 35:1101-5.

6. Felson D T, Anderson J J, Boers M, Bombardier C, Furst D, Goldsmith C *et al*. American College of Rheumatology: preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 727-35.

7. Fransen J, van Riel PL. The disease activity score and the EULAR response criteria. *Clin Exp Rheumatol*. 2005; 23 (Suppl 39): S93-9.

8. Aletaha D, Nell VPK, 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-R806.

9. Smolen JS. A comparison of the SDAI and DAS28 as measures of response in Adalimumab (HUMIRA (TM)) clinical trials in rheumatoid arthritis (RA). *Arthritis Rheum* 2003; 48: S107.

10. Smolen JS, Breedvald FC, Schiff MH, Kalden JR, Emery P, Eberl G *et al*. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* (Oxford) 2003; 42: 244-57.

11. Aletaha D, Stamm T, Smolen JS. Validation of the simplified disease activity index (SDAI) in an observational cohort of patients with rheumatoid arthritis. *Ann Rheum Dis* 2004; 63: 111.

12. Makinen H, Kautiainen H, Hannonen P, Sokka T. A new disease activity index for rheumatoid arthritis: Mean overall index for rheumatoid arthritis (MOI-RA). *J Rheumatol*. 2008; 35: 1522-1527.

13. Pincus T, Sokka T. Complexities in the quantitative assessment of patients with rheumatic diseases in clinical trials and clinical care. *Clin Exp Rheumatol* 2005; 23: S1-S9.

14. Greenberg JD, Harrold LR, Bentley MJ, Kremer J, Reed G, Strand V. Evaluation of composite measures of treatment response without acute-phase reactants in patients with rheumatoid arthritis. *Rheumatology* (Oxford) 2009; 48(6): 686-90.

15. Bentley MJ, Reed GW. Simplified composite disease activity measures in rheumatoid arthritis: Should they be used in standard care? *Clin Exp Rheumatol* 2008; 26: 358-366.

16. Boers M, Brooks P, Simon LS, Strand V, Tugwell P. OMERACT: An international initiative to improve outcome measurement in rheumatology. *Clin Exp Rheumatol* 2005; 23 Suppl 39: S10-3.

17. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998; 25: 198-199.

18. Kremer J. The CORRONA database. *Ann Rheum Dis*. 2005; 64 (Suppl 4): 37-41.

19. Greenberg JD, Bingham CO 3rd, Abramson SB, Reed G, Sebaldt RJ, Kremer J. Effect of cardiovascular comorbidities and concomitant aspirin use on selection of cyclooxygenase inhibitor among rheumatologists. *Arthritis Rheum* 2005; 53:12-7.

20. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983; 26:1346-53

21. Aletaha D, Smolen JS. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheum* 2005; 23 (Suppl 39): S100-S108.

22. Ranganath VK, Yoon J, Khanna D, Park GS, Furst DE, Elashoff DA, et al. Comparison of composite measures of disease activity in an early seropositive rheumatoid arthritis cohort. *Ann Rheum Dis* 2007; 66: 1633-40.

23. Miles J, Shevlin M. *Applying Regression & Correlation.* London: Sage Publications; 2001.

24. Altman DG. *Practical Statistics for Medical Research.* London: Chapman and Hall; 1991.

25. Fleiss JL. Statistical methods for rates and proportions, 2nd ed. New York: John Wiley; 1981: 38-46.

26. Cohen J. *Statistical Power Analysis for the Behavioural Sciences,* 2nd ed. Lawrence Erlbaum; 1988.

27. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; 28: 632-42.

28. Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clinical Epi* 1997; 50: 79-93.

29. Stata Corporation. Stata statistical software: Release 10.0. College Station, TX: Stata Corporation; 2007.

30. Smolen JS, Aletaha D. Activity assessments in rheumatoid arthritis. *Curr Opin Rheumatol*. 2008 May; 20(3): 306-13.

31. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T, et al. Disease activity measurement of rheumatoid arthritis: Comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum* 2005; 53: 56-60.

32. Landis, JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-174.

33. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. *Br J Rheumatol* 1986; 25: 206-09.

34. Sultan N, Pope JE, Clements PJ. The health assessment questionnaire (HAQ) is strongly predictive of good outcome in early diffuse scleroderma: Results from an analysis of two randomized controlled trials in early diffuse scleroderma. *Rheumatology (Oxford)* 2004; 43: 472-8.

35. Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. *Arthritis Rheum* 2003; 48:1530-42.

36. Sokka T, Häkkinen A, Krishnan E, Hannonen P. Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population. *Ann Rheum Dis* 2004; 63: 494-7.

37. Uhlig T, Haavardsholm EA, Kvien TK. Comparison of the Health Assessment Questionnaire (HAQ) and the modified HAQ (MHAQ) in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2006; 45: 454-8.

38. Ziebland S, Fitzpatrick R, Jenkinson C, Mowat A, Mowat A. Comparison of two approaches to measuring change in health status in rheumatoid arthritis: The Health Assessment Questionnaire (HAQ) and modified HAQ. *Ann Rheumatic Dis* 1992; 51:1202-05.

39. Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon G *et al.* Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide Rheumatoid Arthritis Investigators Group. *Arch Intern Med* 1999; 159: 2542-50.

40. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-10.

Chapter 5

1. 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(4): R796-R806.

2. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. *Ann Rheum Dis.* 1990; 49(11): 916-20.

3. van der Heijde DM, van't Hof MA, van Riel PL, Van de Putte L. Disease activity score. *Ann Rheum Dis*. 1992; 51(1): 140.

4. Henke CJ, Epstein WV. Practice variation in rheumatologists' encounters with their patients who have rheumatoid arthritis. *Med Care* 1991; 29(8): 799-812.

5. Aletaha D, Smolen JS. The simplified disease activity index (SDAI) and clinical disease activity index (CDAI) to monitor patients in standard clinical care. *Best Pract Res Clin Rheumatol.* 2007; 21(4): 663-75.

6. Greenberg JD, Harrold LR, Bentley MJ, Kremer J, Reed G, Strand V. Evaluation of composite measures of treatment response without acute-phase reactants in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2009; 48(6): 686-90.

7. Khanna D, Oh M, Furst DE, Ranganath V, Gold RH, Sharp JT *et al*. Evaluation of the preliminary definitions of minimal disease activity and remission in an early seropositive rheumatoid arthritis cohort. *Arthritis Rheum* 2007; 57(3): 440-7.

8. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T, et al. Disease activity measurement of rheumatoid arthritis: Comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum* 2005; 53(1): 56-60.

9. Mierau M, Schoels M, Gonda G, Fuchs J, Aletaha D, Smolen JS. Assessing remission in clinical practice. *Rheumatology (Oxford)* 2007; 46(6): 975-9.

10. Soubrier M, Dougados M. How to assess early rheumatoid arthritis in daily clinical practice. *Best Pract Res Clin Rheumatol* 2005; 19(1): 73-89.

11. Soubrier M, Zerkak D, Gossec L, Ayral X, Roux C, Dougados M. Which variables best predict change in rheumatoid arthritis therapy in daily clinical practice? *J Rheumatol* 2006; 33(7): 1243-6.

12. Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F *et al*. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: Validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther*. 2005; 7(5): R1063-R1071.

13. Ranganath VK, Yoon J, Elashoff DA, Park GS, Khanna D, Furst DE *et al*. Comparison of composite measures of disease activity in an early seropositive rheumatoid arthritis cohort. *Ann Rheum Dis*. 2007; 66: 1633-40.

14. Keenan RT, Swearingen CJ, Yazici Y. Erythrocyte sedimentation rate and C-reactive protein levels are poorly correlated with clinical measures of disease activity in rheumatoid arthritis, systemic lupus erythematosus and osteoarthritis patients. *Clin Exp Rheumatol* 2008; 26(5): 814-9.

15. Donald F, Ward MM. Evaluative laboratory testing practices of United States rheumatologists. *Arthritis Rheum* 1998; 41(4): 725-9.

- **Bentley MJ,** Reed G., Development and evaluation of a modified disease activity score (mDAS28) in rheumatoid arthritis for use in standard care. *Ann Rheum Dis* 2009;68(Suppl 3):530.