



Published in final edited form as:

Ann Rheum Dis. 2009 May ; 68(5): 680–684. doi:10.1136/ard.2008.089318.

Measuring Quality of Care for Rheumatic Diseases Using an Electronic Medical Record

Jessica Agnew-Blais, BA Jonathan S. Coblyn, MD Jeffrey N. Katz, MD, MSc Ronald J. Anderson, MD Jyotsna Mehta, BPharm, MS Daniel H. Solomon, MD, MPH

Division of Pharmacoepidemiology and Pharmacoeconomics (JA-B, JM, DHS), Division of Rheumatology, Immunology, and Allergy (JSC, JNK, RJA, DHS), Brigham and Women's Hospital, Boston, Massachusetts, USA

Abstract

Objectives—The objective of this study was two-fold: one, to determine how best to measure adherence with time-dependent quality indicators (QIs) related to laboratory monitoring, and two, to assess the accuracy and efficiency of gathering QI adherence information from an electronic medical record (EMR).

Methods—We selected a random sample of 100 patients who had at least three visits with the diagnosis of rheumatoid arthritis (RA) at Brigham and Women's Hospital Arthritis Center in 2005. Using the EMR, we determined whether patients had been prescribed a DMARD (QI #1) and if patients starting therapy received appropriate baseline laboratory testing (QI #2). For patients consistently prescribed a DMARD, we calculated adherence with follow-up testing (QI #3) using three different methods, the Calendar, Interval and Rolling Interval Methods.

Results—We found 97% of patients were prescribed a DMARD (QI #1) and baseline tests were completed in 50% of patients (QI #2). For follow-up testing (QI #3), mean adherence was 60% for the Calendar Method, 35% for the Interval Method, and 48% for the Rolling Interval Method. Using the Rolling Interval Method, we calculated adherence rates that were similar across drug and laboratory testing type.

Conclusions—Results for adherence with laboratory testing QIs for DMARD use differed depending on how the QIs were measured, suggesting that care must be taken in clearly defining methods. While EMRs will provide important opportunities for measuring adherence with QIs, they also present challenges that must be examined before widespread adoption of these data collection methods.

Keywords

Rheumatoid Arthritis; Quality Indicators; Disease Modifying Anti-rheumatic Drugs

INTRODUCTION

The Institute of Medicine defines quality as the “degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.”¹ Recently, increased attention has been paid to the formation of indicators for measuring quality and ways of compensating physicians in accordance with such measures.² In late 2006, the US Congress established “G

codes”—quality indicators that physicians must meet in order to receive a 1.5% increase in Medicare payments.³

While the initial set of G codes does not include measures specific to arthritis, quality indicators (QIs) have also been developed specifically for the rheumatology community.⁴ The British Society for Rheumatology created indicators for care for persons with rheumatoid arthritis (RA), including standards for the prescription and monitoring of medications for RA.⁵ In the United States, the Arthritis Foundation Quality Indicator Project (AFQuIP) defined a set of quality indicators for rheumatology,⁶ including indicators related to analgesic use, osteoporosis prophylaxis, disease modifying anti-rheumatic drug (DMARD) monitoring, and gout. The AFQuIP indicators were developed based on extensive literature review and used a rigorous expert consensus method.⁷ The American College of Rheumatology has adopted several of the AFQuIP measures as QIs for rheumatoid arthritis, three of which we examine in this paper.

While QIs have been developed for the rheumatology community, less attention has been paid to the measurement of these indicators. Initial studies suggest that small variations in measurement methodology affect performance on QIs for RA management. For example, Kahn et al. found a discrepancy in data gathered from medical records versus patient self-report in relation to adherence to RA QIs.⁸ Such results stress the importance of clearly defining methods of measuring rheumatology QIs.

We assessed how best to measure indicators related to DMARD use. We examined adherence with these indicators using data collected from our institution's electronic medical record (EMR), as well as from physician's medical record notes. We analyzed both the validity of data collected from the EMR, as well as different methodologies for measuring adherence with temporally dependent laboratory testing QIs.

METHODS

Quality Indicators

We assessed adherence with three ACR QIs related to DMARD use.⁴

- Quality Indicator #1: If a patient has an established diagnosis of rheumatoid arthritis, then the patient should be treated with a DMARD unless refusal or contraindication to DMARD is documented.
- Quality Indicator #2: If a patient with rheumatoid arthritis is newly prescribed DMARD, then appropriate baseline studies should be documented within an appropriate period of time from the original prescription.
- Quality Indicator #3: If a patient has established treatment with a DMARD or glucocorticoids, then monitoring for drug toxicity should be performed. The DMARDs we examined included: abatacept, adalimumab, azathioprine, cyclosporin, etanercept, gold, hydroxychloroquine, infliximab, leflunomide, methotrexate, rituximab, and sulfasalazine. The DMARDs that required laboratory testing and that were examined in QIs #2 and #3 include methotrexate, leflunomide, and sulfasalazine (azathioprine use also suggests laboratory monitoring, but no patients in our sample were prescribed azathioprine).

Study Population

The eligible study population comprised all patients seen as outpatients at Brigham and Women's Hospital (BWH) in 2005 with a primary diagnosis of RA at a particular outpatient visit, as indicated by ICD-9 code 714 for that visit. To improve the likelihood that patients

actually had RA, and to create a cohort with longitudinal care at our institution, we excluded those patients who had been seen fewer than three times with an ICD-9 code for RA. From the 1,077 patients with at least three visits, we then used a random number generator to randomly selected 100 patients to further investigate.

Data Sources

We used two different data sources to determine DMARD use and laboratory testing in the study population. The first was the EMR, which includes the medication list, a widely used feature that allows physicians to print prescriptions and send them by facsimile to pharmacies. The EMR also records laboratory testing performed at the hospital. The second source was physician medical record notes. A trained research assistant used a structured data abstraction form to gather information from the physician medical record notes, including medication start and end dates for all DMARDs and notation of any laboratory testing performed both within and outside of BWH.

Assessment of Quality Indicators

Quality Indicator #1—We assessed whether patients with RA in our cohort received a DMARD. Both the EMR medication list and medical record notes were examined to determine if any DMARDs were prescribed in 2005. Only one DMARD prescription was required.

Quality Indicator #2—The ACR quality indicators recommend baseline laboratory testing for all patients starting a DMARD (see Appendix 1). We defined baseline testing in two ways: one, within 15 days before or after starting a DMARD, or two, within 30 days before starting a DMARD. We identified patients as new starts if they began a DMARD in 2005, and then examined whether the patient had laboratory testing performed within the defined baseline time windows.

Quality Indicator #3—The ACR QIs for follow-up laboratory testing (see Appendix 2) state that testing should be administered every 8 weeks, or 56 days, for methotrexate and leflunomide, and every 12 weeks, or 84 days, for sulfasalazine. To investigate adherence with QI #3, we defined three methods of measuring compliance: the “Calendar Method”, the “Interval Method”, and the “Rolling Interval Method”.

The Calendar Method defines six 56-day periods, starting January 1st, 2005 and extending to December 3rd, 2005. We then assessed adherence with the indicator by determining whether the patient had a least one laboratory test during each of the six periods. For example, the first period extends from January 1st to February 26th. If, as in the example patient shown in Figure 1, a patient had a laboratory test on January 14th, the subject would be considered adherent for that period. One hundred percent adherence occurs when a patient has at least one laboratory test in each of the pre-defined 56-day periods, as illustrated in Figure 1.

In contrast, the Interval Method focuses not on the placement of the test within the calendar year, but on the days between tests. We defined a period as the number of days between laboratory testing, and identified the period as adherent if 56 days or less in length. For example, if testing occurred on January 4th and then again February 28th, the time between tests would be 55 days, and this interval would be considered adherent. If, however, the second set of tests was not completed until March 15th, the time between tests would be 70 days, and this period would be considered nonadherent.

The Rolling Interval Method represents a combination of the first two methods, in that it defines periods as 56-days, as in the Calendar Method, but these periods are started and stopped based on the dates of actual tests, as in the Interval Method. In the Rolling Interval Method, the first period begins on January 1st, 2005 and extends either until a laboratory test is performed, or until 56 days have passed. This is illustrated in the third panel of Figure 1. In this example, the period starts on January 1st and extends until the first laboratory test, on January 14th. This period is considered adherent since a laboratory test was performed within 56 days. The period then restarts on the test date, January 14th. This second period extends for 56 days until March 11th. Because 56 days pass with no laboratory test, this period is considered nonadherent. The next period then starts on March 11th and runs until the subsequent laboratory test on March 15th. This period is considered adherent because a laboratory test was performed within 56 days. This process continues, considering 56-day periods with a laboratory test as adherent and 56-day periods with no laboratory test as nonadherent.

RESULTS

The characteristics of the 100 randomly selected patients are shown in Table 1. 83% of these subjects were female and the mean age was 61.1 years. Approximately 50% had some form of commercial health insurance, followed by about one-third with Medicare. Many patients had multiple health insurances.

Table 2 illustrates the results for Quality Indicator #1, in which the most commonly prescribed DMARD was methotrexate and many patients received combination DMARD therapy. This table also compares the two data sources, the EMR and the physician medical record notes. The medication list from the EMR documented that 90% of patients were prescribed a DMARD. The medical record notes documented DMARD use in 97% (agreement Kappa, 0.96).

Table 3 illustrates Quality Indicator #2, adherence with baseline laboratory testing guidelines for subjects who started a DMARD in 2005. We defined adherence with baseline testing in two ways: one, testing 15 days before or after start, and two, testing 30 days before starting a DMARD. Fifteen patients began a DMARD in 2005. Defining baseline as 15 days before or after initiation or 30 days before initiation resulted in similar rates of adherence: about a third to a half of subjects received baseline testing. Many received testing on the same visit in which they were prescribed the DMARD.

Of these new starts, three subjects did not have any laboratory testing recorded in the EMR. To further look into the issue of a lack of any record of testing, a trained research assistant examined medical record notes of the ten patients, both new starts and consistent users, who were found to have no laboratory testing in 2005 based on the EMR laboratory module. For all of these patients with no tests recorded in the EMR laboratory module, results of tests were mentioned for all ten patients in the physician visit notes. This suggests that patients had testing at other locations, with results sent to their rheumatologists but not recorded in the EMR laboratory module. To investigate the presence of outside laboratory testing further, we again used a random number generator to randomly select 20 subjects consistently prescribed a DMARD in 2005 and examined their physician visit notes for the mention by physicians of outside laboratory tests. Of these 20 patients, three had definite mention of outside laboratory testing, one had suggestion of outside testing and 16 had no mention or suggestion of outside laboratory testing.

Figure 2 illustrates the adherence rates with QI #3, laboratory monitoring for patients consistently prescribed a DMARD, grouped by type of laboratory test; adherence for these

groups of tests (complete blood count, creatinine, and liver function) was similar. Figure 2 also shows the three different methods for measuring laboratory testing adherence. The first bar of each grouping (grouped by drug), represents the Calendar Method, which consistently gives the highest adherence. The second bar, the Interval Method, consistently yields the lowest. The third method, the Rolling Interval Method, results in adherence that falls between the first two. This trend is reflected in the overall adherence averages of the Calendar (60%), Interval (35%) and Rolling Interval Methods (48%), for all drugs, all laboratory tests. Figure 2 shows that adherence with testing is similar for the three DMARDs: using the Rolling Interval Method, we found an average adherence with monitory testing for methotrexate of 47%, leflunomide of 43%, and sulfasalazine of 54%.

DISCUSSION

The measurement of adherence with quality indicators is becoming increasingly important as groups such as insurance companies, hospitals, and physician organizations focus on establishing QIs and pay-for-performance measures. We assessed adherence with ACR QIs both to measure the performance of rheumatologists in a particular clinical practice and to examine methodological challenges in measuring adherence with these QIs.

We defined three different approaches to measure adherence with follow-up laboratory testing (QI #3). The Calendar Method does not simply sum the number of laboratory tests in one year, but also recognizes that to comply with the QI, the tests should be spaced throughout the year. The strength of the Calendar Method is the straightforward nature of measurement: the periods are the same for every subject and their number does not change. This method is easy to conceptualize and to measure. The central weakness of the Calendar Method is that it does not operate on a “biological clock,” as it does not effectively take into account the time lag between laboratory tests. For example, a patient could receive a test at the beginning the first period, on January 4th, and at the end of the second period on April 20th, and would be assessed as adherent, although the actual time between the tests, 106 days, was more than the recommended 56 days or less. This limitation means that the Calendar Method does not measure the appropriate spacing of tests. Because of this, the Calendar Method tends to inflate adherence rates.

To address this limitation, we developed the Interval Method. The Interval Method focuses on the actual testing dates, defining the length of the period by actual test dates. But it also has a significant weakness, illustrated in Figure 2. By measuring only the number of days between tests, a patient who received five tests, all spaced greater than 56 days apart (i.e., all not adherent), could have the same poor adherence as a patient who received no testing for the entire year. It could be argued that the five tests, because they were not within the recommended time window, never succeeded in meeting the QI. However, it seems illogical to equate a pattern of five slightly delayed tests with that of no tests at all. Because of this, the Interval Method tends to underestimate adherence.

A potential solution to the problems with the Calendar and Interval Methods is the Rolling Interval Method. We found this method the most convincing because it addresses the weaknesses of the first two and produces adherence rates that fall within the more extreme rates produced by the Calendar and Interval Methods. This Rolling Interval Method produces adherence rates that take into account lack of testing, as well as adherence due to inappropriate spacing of tests, resulting in a composite measure of adherence. Unfortunately, the Rolling Interval Method is also the most complicated method to explain and to measure.

We examined both EMR data and physician visit notes to assess QIs related to DMARD use. We found that aspects of the EMR influence how adherence is measured. For example,

certain necessary components of our analysis, such as the DMARD start date, were not automatically recorded in our hospital's EMR, but had to be abstracted from the physician visit notes. As noted in the results, we found a certain portion of patients (3 out of 20 randomly sampled subjects) received laboratory testing outside of the hospital, which underestimated adherence to QIs.

Our study is limited by the fact that it was conducted at only one medical center, thus these results may be unique to our particular institution. Different clinical settings may have different EMRs, or no EMR at all. Certain aspects of our clinical practice, such as the high number of patients receiving testing at other sites, may be less (or more) common in other settings. As mentioned above, if a physician's patients receive laboratory testing at outside institutions, her adherence rates would erroneously appear low compared to other physicians. This limits one's ability to compare individual physicians or practices using our EMR based method. Additionally, we only measured whether patients actually received laboratory testing. It is entirely possible that physicians were diligent in reminding patients of the need for testing and patients did not follow up, but are unable determine where specifically the breakdown in completing laboratory testing occurred based on our data sources.

Another limitation of this study was that we did not have pharmacy filling data for subjects, and therefore did not have the date at which the patients actually picked up their DMARD prescription from the pharmacy. The start dates for DMARDs were defined as the date in the medical record notes when the physician noted first prescribing the drug and it is possible that the patient did not actually fill the prescription, or began the DMARD at a later date. Additionally, we did not examine how different forms of insurance coverage, such as government-sponsored Medicare versus comprehensive commercial insurance or self-pay, could have affect adherence with the QI. In the United State's healthcare system, it is possible that the various insurance types could have helped or hinder adherence, and this is an issue that should be examined in future research.

After determining how best to measure adherence with laboratory testing QIs, steps should be taken to determine how adherence with these QIs could be improved. Some research has found adherence improvements using an automated reminder system.⁹ An EMR that comprehensively collects laboratory data could be used to both track adherence and generate reminders for testing by flagging patients who are scheduled for or have missed tests.

A focus on quality indicators must also include an emphasis on determining the most accurate and efficient methods of assessing QI adherence. As our results show, both the data sources as well as the specifications of the QIs can have important effects on the results of adherence with indicators. It is critical to clearly specify how QIs are defined and measured if conclusions and comparisons are to be made from their assessment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

There was no specific support for this work. Dr. Katz receives salary support from the NIH (K24 AR 02123 and P60 AR 47782). Dr. Solomon receives salary support from the NIH (R21 AG027066 and P60 AR47782) and from Pfizer.

References

1. Crossing the Quality Care Chasm: The IOM Health Care Quality Initiative. Jul 202006 [Access January 8, 2008.]. <http://www.iom.edu/CMS/8089.aspx>
2. Epstein AM, Lee TH, Hamel MB. Paying physicians for high-quality care. *N Engl J Med* 2004;350:406–10. [PubMed: 14736934]
3. Physician voluntary reporting program. Oct 282006 [Accessed January 8, 2008.]. <http://www.cms.hhs.gov/apps/media/press/release.asp?Counter=1701>
4. Quality Measures Starter Set. Feb2006 [Accessed January 8, 2008.]. <http://www.rheumatology.org/practice/qmc/starterset0206.asp>
5. Kennedy C, McCabe C, Struthers G, Sinclair H, Chakravaty K, Bax D, Shipley M, Abernethy R, Palferman T, Hull R. BSR guidelines for standards of care for persons with rheumatoid arthritis. *Rheumatology* 2005;44:553–556. [PubMed: 15728419]
6. MacLean CH, Saag KG, Solomon DH, Morton SC, Sampsel S, Klippel JH. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's Quality Indicator Set. *Arthritis Rheum* 2004;51:193–202. [PubMed: 15077259]
7. Sackman, Harold. Delphi Assessment: Expert Opinion, Forecasting, and Group Process. [Accessed January 8, 2008.]. <http://www.rand.org/pubs/reports/R1283/>
8. Kahn KL, Maclean CH, Wong AL, Rubenstein LZ, Liu H, Fitzpatrick DM, Harker JO, Chen WP, Traina SB, Mittman BS, Hahn BH, Paulus HE. Assessment of American College of Rheumatology quality criteria for rheumatoid arthritis in a pre-quality criteria patient cohort. *Arthritis Rheum* 2007;57:707–715. [PubMed: 17530663]
9. Feinstein A, et al. Electronic Medical Record Reminder Improves Osteoporosis Management After a Fracture: A Randomized, Controlled Trial. *J Am Geriatr Soc* 2006;54(3):450–45. [PubMed: 16551312]

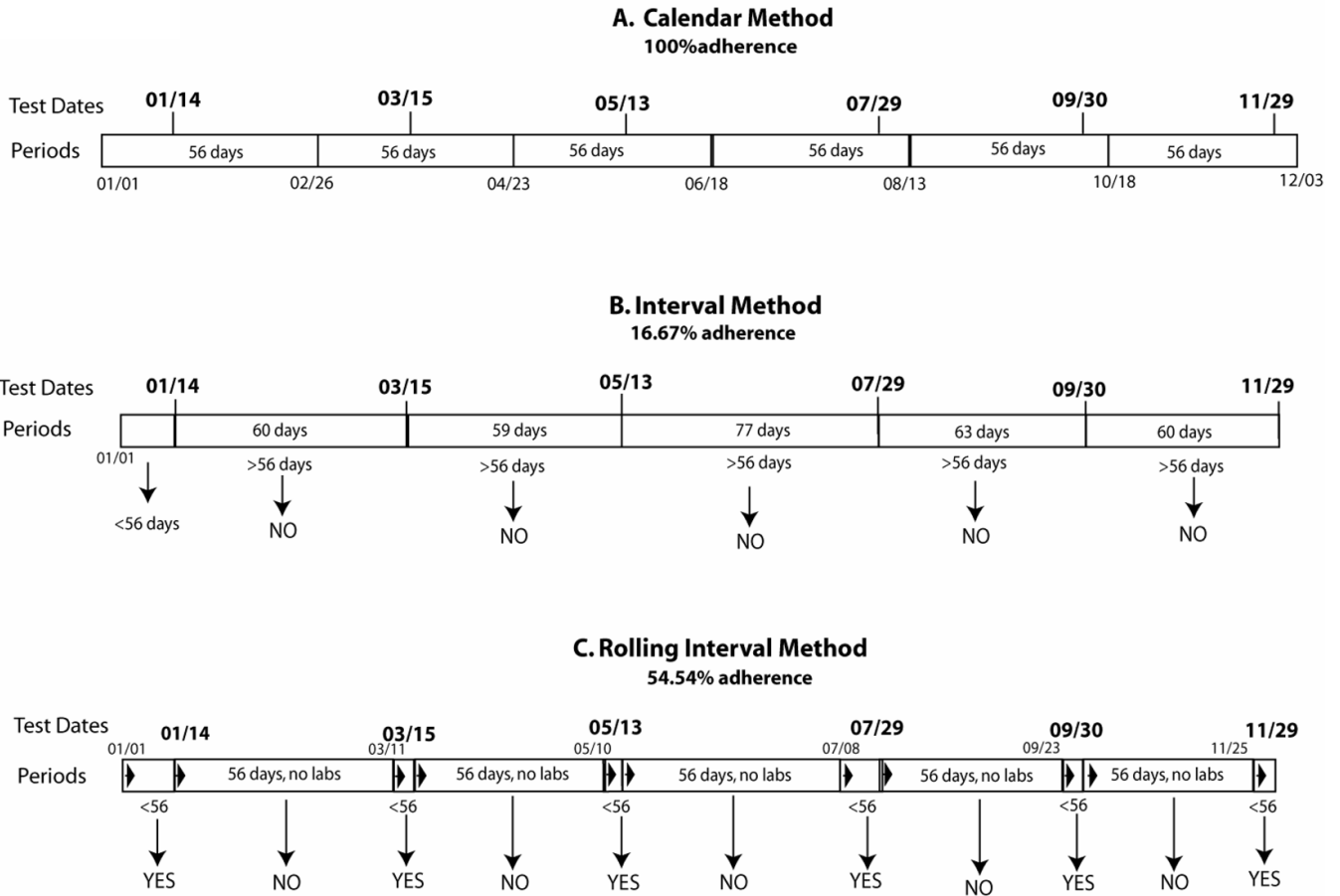


Figure 1. This figure shows the laboratory testing date for one example patient and illustrates the three methods for assessing compliance with follow-up laboratory testing. Panel A, the Calendar Method, illustrates the six periods of 56-days, with each of the six laboratory testing dates falling into each of these periods. This results in 100% adherence for this patient using the Calendar Method to assess compliance. Panel B, the Interval Method, shows that the periods are defined by the laboratory testing dates. The bars of the periods show the number of days between laboratory testing dates. Of the six periods that are defined by the laboratory testing dates, Panel B, only the period of January 1st to January 14th is less than 56 days long. This results in only 16.67% adherence for this patient using the Interval Method, a far lower adherence percentage that previously assessed using the Calendar Method. Panel C, the Rolling Interval Method, shows the periods starting and re-starting based on the laboratory testing dates. The period begins on January 1st, as in the other two methods. When a laboratory test occurs on January 14th, the first period is completed and considered adherent because it is shorter than 56 days. The second period starts on January 14th, continues for 56 days, and is considered non-adherent, as there are no laboratory tests within this 56-day period. This results in six periods of adherence and five periods of nonadherence, or an adherence rate of 54.54%

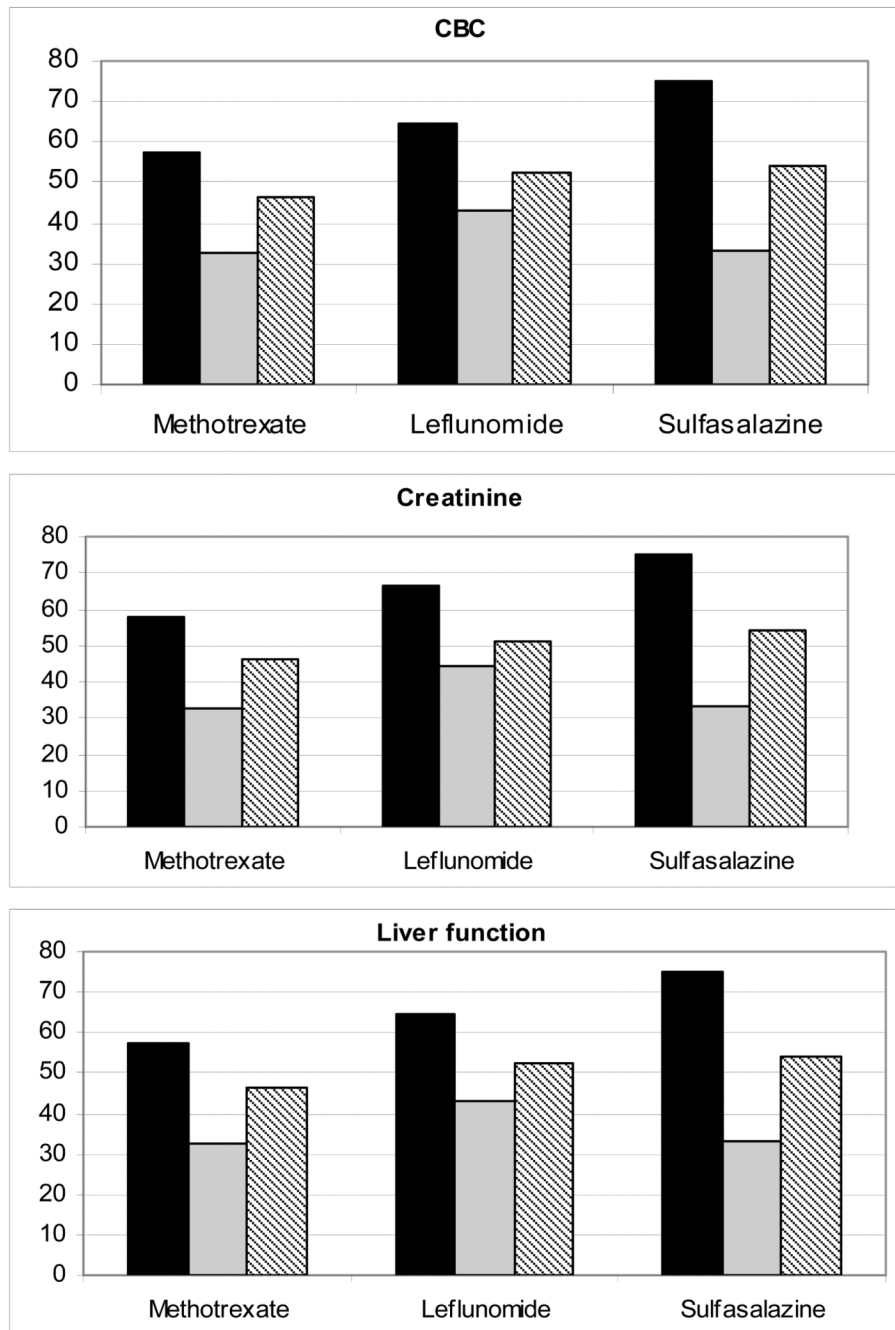


Figure 2. Abbreviations: CBC=complete blood count, MTX= methotrexate, LEF= leflunomide, SSZ= sulfasalazine. This figure illustrates the results of the three methods of measuring follow-up laboratory testing for three sets of laboratory tests: complete blood testing (platelets, white blood cell count, and hemoglobin), liver function test (aspartate aminotransferase/alanine aminotransferase, and albumin) and creatinine level

Table 1

Characteristics of 100 randomly selected rheumatoid arthritis patients seen at Brigham and Women's Hospital in 2005

Characteristics	N (%)
Females	83
Age (years)	61.1
Insurance	
Commercial	84 (55.6)
Medicare [†]	50 (33.1)
Medicaid [‡]	8 (5.3)
Dual-eligible [*]	4 (2.5)
Free Care ^{**}	4 (2.5)
Self Pay	1 (0.7)

* Dual-eligible patients qualify for both Medicare and Medicaid.

[†] Medicare and Medicaid are government-sponsored health insurance programs. Medicare is offered to the elderly while Medicaid is provided based on income level.

** Free care is free or reduced cost health care provided by hospitals and health centers to uninsured patients.

Table 2

Disease modifying anti-rheumatic drug use among study population of 100

Medication	Electronic Medical Record**	Medical Record§
Any DMARD*	90 (90)	97 (97)
methotrexate	52 (52)	61 (61)
etanercept	24 (24)	26 (26)
hydroxychloroquine	13 (13)	15 (15)
leflunomide	13 (13)	17 (17)
adalimumab	13 (13)	17 (17)
infliximab	6 (6)	11 (11)
sulfasalazine	3 (3)	5 (5)
anakinra	1 (1)	0 (0)
azathioprine	1 (1)	1 (1)

* DMARD: disease modifying anti-rheumatic agent

** Based on the medication list in the Brigham and Women's Hospital electronic medical record.

§ Based on review of physician notes in the electronic medical record.

Table 3

Baseline laboratory testing of patients starting disease modifying anti-rheumatic drugs (Quality Indicator #2)

	N	Met QI (+/- 15 days)		Met QI -30 days)	
		N (%)	N (%)	N (%)	N (%)
All new starts	15	6 (40)		5 (33.3)	
methotrexate	13	5 (38.5)		5 (38.5)	
leflunomide	2	1 (50)		0 (0)	
sulfasalazine	0	0 (0)		0 (0)	
azathioprine	0	0 (0)		0 (0)	