Rate of Change in Functional Limitations for Patients with Rheumatoid Arthritis: Effects of Sex, Age, and Duration of Illness

T. JOSEPH SHEEHAN, SARAH DuBRAVA, JUDITH FIFIELD, SUSAN REISINE, and LAURIE DeCHELLO

ABSTRACT. ***Objective.*** To estimate the rate of change in functional limitations for patients with rheumatoid arthritis (RA) as a function of age, duration of illness, and sex.

Methods. Patients with RA (n = 700) aged 21–65 years in 1988 were interviewed yearly for 6 years in The National Rheumatoid Arthritis Study. Functional limitations scores based on a Rasch measurement model of 20 Health Assessment Questionnaire items were analyzed in mixed-effects models to estimate the rate of change in functional ability as a function of age, duration of illness, sex, and interactions.

Results. Models for both patient age and duration of illness significantly predicted limitations in functional ability for men and women. The model for age included a significant cubic effect; the model for duration of illness included a significant linear effect only. Sex was significant in both models and no interactions were significant in either model. The AIC index of fit, an indicator of the information value of the model, favored the model for duration of illness over the model for age. While both models showed higher levels of functional limitations in women than men, the rate of change for women was similar to men.

Conclusion. Limitation in functional ability in RA progressed in a linear manner with duration of illness and progressed at the same rate for both men and women, but functional limitations were greater for women. (J Rheumatol 2004;31:1286–92)

Key Indexing Terms:
RHEUMATOID ARTHRITIS
FUNCTIONAL LIMITATIONS
RASCH
FUNCTIONAL ABILITY
MIXED-EFFECTS MODELS

Rheumatoid arthritis (RA) is chronic and is associated with increasing difficulty in functional ability as indicated by the performance of activities of daily living (ADL)[1-3]. With new and promising biological therapies such as infliximab and etanercept[4,5], it is likely that difficulties in functional ability will stabilize in existing patients and that future patients will experience less difficulty in ADL and ultimately less disability over time. Therefore, it is important to establish reliable baseline estimates of the way in which functional ability changes prior to widespread use of new therapies. This will enable the tracking of improvements with new therapies, contributing to the validation of their efficacy and effectiveness. This study estimates the rate of change in functional ability among patients with RA, under the care of a rheumatologist, but not receiving any of the new biological therapies during a 6-year period from 1992 to 1997.

Studies have reported that limitation in ADL among RA patients is slow to develop[1], does not change in a linear fashion over time[1,3], is reported at higher levels among women[6,7], and changes at different rates for men and women[1]. However, several factors limit our confidence in these findings. First, most samples are drawn from well studied clinic populations and may not be representative of RA patients generally. Second, measurement models used have not been appropriate for ordinal data[8]. This study addresses limitations in samples, using patients from a national panel, and uses the Rasch measurement model[9] to transform the nonlinear ordinal Likert scores into linear interval measures. It also uses a mixed-effects statistical model for analyzing changes over time, a more appropriate statistical model for longitudinal data. These refinements should improve estimates of how limitations in functional ability change over time in patients with RA.

MATERIALS AND METHODS

Sample. Subjects were drawn from The National Rheumatoid Arthritis Study, a prospective cohort study with followup interviews each year for 10 years between 1988 and 1997[10]. Participants were patients who had been diagnosed with classical or definite RA[11], and were under the care of one of 56 physicians of 116 randomly selected board certified rheumatologists

From the Department of Community Medicine and Health Care and Department of Family Medicine, University of Connecticut School of Medicine; and Department of Behavioral Sciences and Community Health, University of Connecticut School of Dental Medicine, Farmington, Connecticut, USA.

Supported by a grant from the Arthritis Foundation and the Claude Pepper Older Americans Independence Center.

T.J. Sheehan, PhD; S. DuBrava, MS; L.M. DeChello, BA, Department of Community Medicine and Health Care; J. Fifield, PhD, Department of Family Medicine; S. Reisine, PhD, Department of Behavioral Sciences and Community Health.

Address reprint requests to Dr. T.J. Sheehan, Department of Community Medicine (MC-6325), University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030.

Submitted March 31, 2003; revision accepted January 22, 2004.

who were members of the American College of Rheumatology in 1987. Data from the first 4 years of the study were not analyzed because interviews included only 13 of the 20 Health Assessment Questionnaire (HAQ) items, while the remaining 6 years used the 20-item HAQ[12]. Participants were interviewed on the telephone for these 6 years, answering the 20-item HAQ. Seven hundred of the original 988 study participants, 551 women and 149 men, completed at least one interview between the fifth and tenth years of the study.

Measures. There are numerous scales to measure functional ability, including the Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL)[13], and the HAQ[12], each of which combines responses into a single measure of functional ability. The HAQ was first introduced by Fries in 1980, and was developed specifically to study rheumatology outcomes. Test-retest reliability of the HAQ has been studied, with resulting correlations ranging from 0.87 to 0.96. Studies have also shown that the HAQ is sensitive to change in patients' functional ability[14,15]. Administration of the HAQ by telephone interview has been validated in several studies[14]. Subjects were asked how often they have difficulty performing a series of tasks of daily living, each with the following possible responses: "with no difficulty," "with some difficulty," "with much difficulty," and "unable to do." These answers were coded 0, 1, 2, and 3, respectively. Because these responses are ordinal, there is no metric for judging the distance between responses, although equal intervals are often assumed. Further, the ordinal nature of the responses does not lend itself to algebraic manipulation, a prerequisite for many statistical analyses, even for summary measures such as means[8].

Measurement model. The ordinal nature of the HAQ scales was addressed by fitting a Rasch model[9] to estimate functional status, item difficulty, and distance between response categories on the same interval scale[16]. Equation 1 models patient responses to the 20 HAQ items to estimate a patient's functional limitation status, the difficulty of each HAQ item, and the distance between response categories. There are 4, or *k*, categories of responses to each HAQ item, and the model estimates the probability that the patient will respond to category k rather than level *k-1* to item *i* where:

$$\log\left[\frac{P_{nik}}{P_{nik-1}}\right] = B_n - D_i - F_k$$

where P_{nik} = probability of Patient *n* choosing functional status level *k* on item *I*; P_{nik-1} = probability of Patient *n* choosing functional status level *k–1* on Item *I*; B_n = functional ability level of Patient *n*; D_i = difficulty level of Item *I*; F_k = difficulty of choosing functional status Level *k* relative to Level *k–1*.

The equation estimates the log-odds that a patient with functional ability B_n will choose functional status Level *k* rather than functional ability Level *k–1* on an Item with Difficulty D_i. Equation 1 is the partial credit version of the Rasch model developed by Andrich[17]. The Rasch model estimates the functional limitations of each person, B_n, and the difficulty of each item, D_i, and the distances between each response category, F_k, and relates these estimates to the same logit scale. A summary of the analyses for the first study year is shown in Figure 1. Figure 1 shows the logit scale in column 1, the distribution of patients in column 2, the distribution of items in column 3, and the distances between Likert scale categories in column 4, where the width of the interval within each category is indicated by the broken lines. Patients with positive logit scores have more functional limitations. Items with positive logit scores are the most difficult to perform. The threshold for logit positive scores is about halfway between response category 1 and response category 2.

The Rasch model assumes that the item responses are influenced by a single construct, or underlying variable — in this case, functional ability. If this assumption is true, there should be no meaningful correlation among item residuals after a Rasch model is fit, an assumption that can be tested by performing a principal components analysis (PCA) on the residuals after

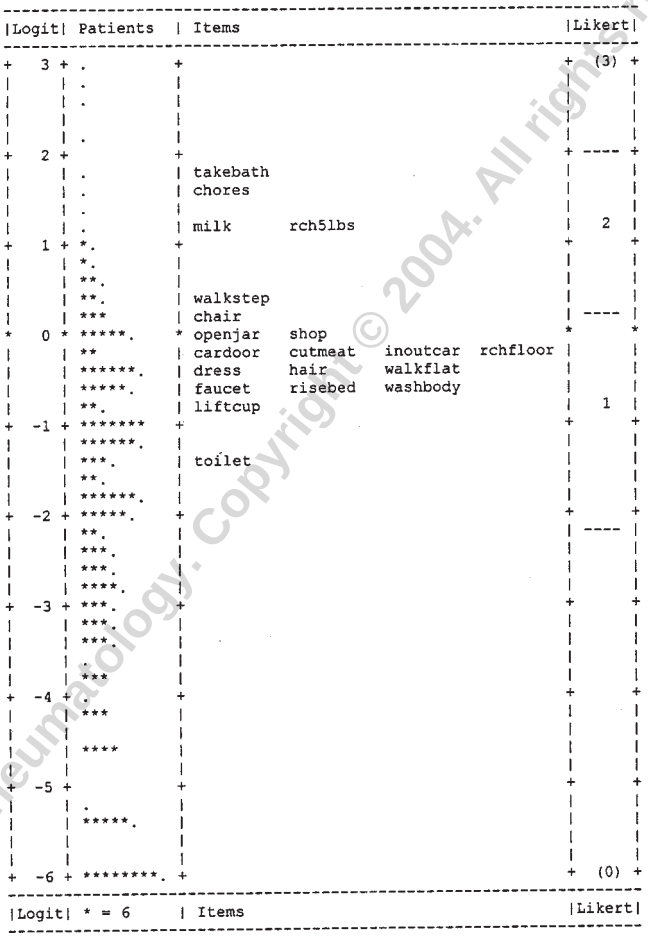

Figure 1. Distribution of RA patients, functional status scores, item difficulties, and distances between response categories, on the Rasch logit scale. Data are from Time 1.

fitting a Rasch model[18]. If there is meaningful covariation in the residuals, the Rasch model is inappropriate. The PCA found 3 components, which accounted for a total of 19.85% of the variance where a meaningful PCA would be expected to account for 75% of the residual variance. Therefore, since the PCA found no meaningful structure in the residuals, it is fair to infer the presence of a single underlying variable: functional ability. If the Rasch model fits, the Rasch estimates of each subject's functional ability are interval-scale measures of overall functional ability and can be compared meaningfully from subject to subject, as well as within the same subject over time.

This study uses data from a national sample of RA patients and a Rasch model to ascertain how self-reports of functional ability limitations in ADL change over time for men and women with RA. We have reported that the HAQ items fit the Rasch measurement model[19].

For each year of data, the Rasch model was fit to the HAQ data using the FACETS program[20]. The Rasch estimates of reliability for the 20-item HAQ ranged from 0.91 to 0.99, consistent with previous applications of the Rasch model to the HAQ and other sets of ADL and IADL items[19]. The outfit mean-square is an outlier-sensitive mean-square fit statistic; it has an expectation of 1. Values substantially less than 1 indicate dependency in the data, while values substantially greater than 1 indicate the presence of outliers. A value greater than 2.0 indicates a potential outlier. The infit mean-square statistic is similar to outfit mean-squares, except a value substantially greater than 1, or larger than 2.0 indicates noise. All items had

infit and outfit mean-squares of at least 0.7 and less than 2.0. "Take a bath" had the highest statistics, with outfit and infit mean-squares of 1.9. "Getting in and out of a car" had the lowest statistics, with infit and outfit mean-squares of 0.7. Infit and outfit values of 0.7 and 1.9 are acceptable values, therefore all items may be used.

Statistical analyses. The Rasch measure of functional ability was the main dependent variable in this study, with patients measured yearly over a 6-year period. Subjects with higher scores signify more limitations in functional ability. To estimate rates of change in functional ability over time, we fit linear, quadratic, and cubic mixed effects models using Proc Mixed in SAS[21]. Mixed effects models contain both fixed effects and random effects. Fixed effects assess the influence of the usual grouping variables such as sex or race or experimental factors that apply to the entire population. Random effects are associated with individuals drawn at random from a population, and allow a different rate of change for each patient over time, rather than assuming each person's change in functional ability is the same from year to year[22]. Mixed effects models allow fixed effects and random effects to be included in the same model. In this study, age, duration of illness, and sex were considered fixed effects, while age and duration of illness were also considered random. Modeling a single effect such as duration of illness as both fixed and random fits an overall fixed effect as well as an individual adjustment to the fixed effect. Both the fixed and random portions of the model contain an intercept, the point at which the function crosses the y-axis, reflecting the level of functional ability when initially diagnosed.

Because men and women report different levels of functional ability limitations and are thought to change at different rates over time[1,6,7], sex was an important covariate in this analysis. Also, since functional ability decreases over time, age at each interview as well as duration of illness were also included. Duration of illness was defined as the year of the interview minus the year of diagnosis. To test for nonlinear effects, quadratic and cubic terms were used in the statistical models for both age and duration of illness. And because of reports that men and women change at different rates, all interaction terms were also included in both models. Since there were only 22 non-white women and 5 non-white men, race was not used as a covariate.

Because age and duration of illness are correlated both statistically and conceptually, we tested separate models and then compared whether the model for age or duration of illness was better. The model for age estimates the effects of age, age squared, age cubed, sex, and the interactions between these effects. Nonsignificant terms are successively removed, starting with the largest p-values in the highest order interactions, until all terms remaining are significant at $p \leq 0.05$ level of significance. A similar model was tested for duration of illness, duration of illness squared, duration of illness cubed, sex, and all the interactions betweens these effects. The interaction between sex and duration of illness reveals whether the rate of change for RA disability is different for men and women. Again, nonsignificant terms were successively removed until all remaining terms were significant. The best model containing age was then compared to the best model containing duration of illness, using Akaike's Information Criterion (AIC). AIC is a fit index reflecting goodness of fit, which rewards parsimony by penalizing for overparamaterization[23]. The model with the smaller AIC is better.

RESULTS

Table 1 contains summary statistics describing the 551 women and 149 men who participated in the study, as well as the 210 women and 78 men who were excluded because of missing data. Data were missing on those 288 patients since they were lost to followup, died, or dropped out before the fifth year of the study. Most participants identified themselves as white, and the average age of the 551 women was 54.3 years at the fifth year, ranging from 29 to 68. Of these 551 women, 332 had complete data, that is, had completed all 6 interviews between the fifth and tenth years of the study. The average age for men was 55.0, also with a range of 29 to 68. Of these 210 men, 73 had complete data. At the fifth year, the mean duration of illness was 14.2 years for women and 14.1 years for men, with maximum values of 35 and 46 years of illness, respectively. Table 2 displays the number of participants that completed all the interviews and any fewer number of interviews by sex. All interviews were completed by 332 women and 73 men, while 218 women and 76 men completed only 5 or fewer interviews. Therefore, 551 women and 149 men were included in the analyses.

Table 3 displays the means and standard deviations of Rasch person scores by sex and duration of illness at Time 1 (year 5), with higher scores reflecting higher functional limitations. All the means for person measures have negative signs, suggesting a higher level of functional ability than might be expected from items with a mean of zero. The standard deviations for the person scores are large, however, relative to the means, suggesting a great deal of spread in levels of functional limitations. Neither the skewness, which varies from –0.976 to –0.206 for the 6 time points, nor the kurtosis, which varies from –1.312 and –0.051, is extreme. Functional limitations scores for women are consistently higher than for men, suggesting a higher level of functional limitations among women.

Table 1. Summary statistics by sex of participants included and not included in the analyses.

	Included in Analyses		Not Included in Analyses*	
	Women	Men	Women	Men
N	551	149	210	78
White	493**	134	167**	70
Black	29**	8	19**	5
Age at year 5	54.3 (9.9)†	55.0 (9.1)†	55.0 (10.6)†	55.8 (8.6)†
Duration of RA at year 5	14.2 (8.7)†	14.1 (9.2)†	14.4 (8.9)†	13.2 (7.7)†
Total no. of joint flares past month	1.00 (1.3)†	0.88 (1.4)†	1.01 (1.2)†	1.15 (1.5)†
Total no. of joint flares past year	1.92 (1.5)†	1.79 (1.6)†	1.88 (1.4)†	1.97 (1.6)†

* Lost to followup, died, or dropped out of the study before year 5. ** Cases included in the analyses and not included in analyses are significantly different, $p < 0.05$. † Mean (standard deviation).

Table 2. Number of interviews completed, by sex.

	All 6 years	5 years	4 years	3 years	2 years	1 year
Women	332	45	33	26	53	61
Men	73	12	15	8	14	27

Table 3. Means and SD of Rasch person scores at Time 1 by sex and duration of illness. The reader may use column 4 of Figure 1 to relate these Rasch scores to a corresponding Likert scale category.

	Duration of Illness, yrs	N	Mean	SD
Women	4–7	137	–2.49	1.88
	8–11	126	–1.88	1.78
	12–18	145	–1.83	1.99
	> 18	143	–1.33	2.12
Men	4–7	39	–3.76	2.50
	8–11	34	–3.42	2.37
	12–18	36	–2.44	1.99
	> 18	40	–2.62	2.29

Estimates from the best-fitting statistical model, the model featuring duration of illness, are shown in Table 4. The mixed effects modeling began with the Rasch score as the dependent variable and duration of illness, sex, and all interactions as the fixed effects. Duration of illness was also fit as random, providing each patient an individual intercept and slope. Models were also fit using age in place of duration of illness. These models were reduced as previously described, and compared using AIC. According to AIC, the model that best predicted functional limitations included an intercept, duration of illness, and sex as fixed effects, with no significant interactions. Those estimates are shown in Table 4. The model also included duration of illness, as well as an intercept, as random effects. The effect of sex on functional limitations shows an estimated effect of –1.23 units for men and zero for women, suggesting that women will score 1.23 Rasch units higher than men on average, indicating worse functional limitations for women given the same duration of illness. The model also suggests that functional limitations advance over time at about the same rate for men and women, since there were no significant interactions between sex and the duration of illness terms. There is a significant, positive effect for duration of illness, which would imply that functional limitations increase with duration of illness. In addition, neither the squared nor the cubic terms were statistically significant, suggesting no evidence of nonlinear changes in functional limitations associated with duration of illness.

Table 4. Fixed effects estimates for duration of illness and sex as predictors of limitation in functional ability.

Effect	Estimate	p
Intercept	–2.5702	< 0.0001
Duration of illness	0.0532	< 0.0001
Men	–1.2311	< 0.0001
Women	0	

The linear effect of duration of illness is seen in Figure 2, with duration of illness on the x-axis and estimated effects in Rasch score units on the y-axis. The left graph shows the mean functional limitations scores for men and women predicted by the fixed effects part of the model. The right side plots the estimates of mean functional limitations predicted by the fixed and random effects together for men and women. The sample size is somewhat small for men (n = 149), which adds to the variation for the fixed and random graph. Both graphs show a higher level of functional limitations for women than men, and also show that the rate of change is the same for men and women. In Figure 2, the horizontal line at –2 logits represents the level of functional limitation, which coincides with the transition from category 0 to category 1 on the Likert Scale, as shown in Figure 1, or having some difficulty on HAQ items. The level of –1 logit corresponds to the midpoint of category 1 on the Likert Scale, having some difficulty with 6 or more of the HAQ items. Women reach the –2 logits level of functional limitations by a duration of illness of about 10 years, a level reached by men at about 34 years' duration.

Fixed effects estimates of the best-fitting model for age as a predictor of functional limitations are shown in Table 5. This model had a higher AIC, 10,480, than the AIC for duration of illness, 10,429, and therefore does not fit the data as well. However, since there was not much difference in the AIC of both models, we find it important to present the results of this model as well. In this model, the effect of sex on functional limitations shows men with a functional limitations level 1.25 units lower than women of the same age, similar to the estimate of –1.23 from the duration of illness model. Table 5 also shows that the cubic term for age is statistically significant. Even though the linear and quadratic effects for age are not statistically significant, it is customary to retain a nonsignificant lower-order effect if a higher-order effect is statistically significant. The nonlinear effects of age can be seen in Figure 3. The cubic effect of age is seen as functional limitation increases, levels off, then increases again as age increases. It is interesting that women reach the –2 logit level of functional limitations at about age 47, whereas men do not approach it until after age 70.

DISCUSSION

Our analysis of functional ability among patients with RA builds on previous studies by using the Rasch model to estimate functional limitation levels and mixed effects modeling to assess changes in functional limitations over time. Further, it uses a sample of patients drawn from a national random sample of rheumatology practices, which may be more representative of RA patients than previously studied patients who were recruited from university prac-

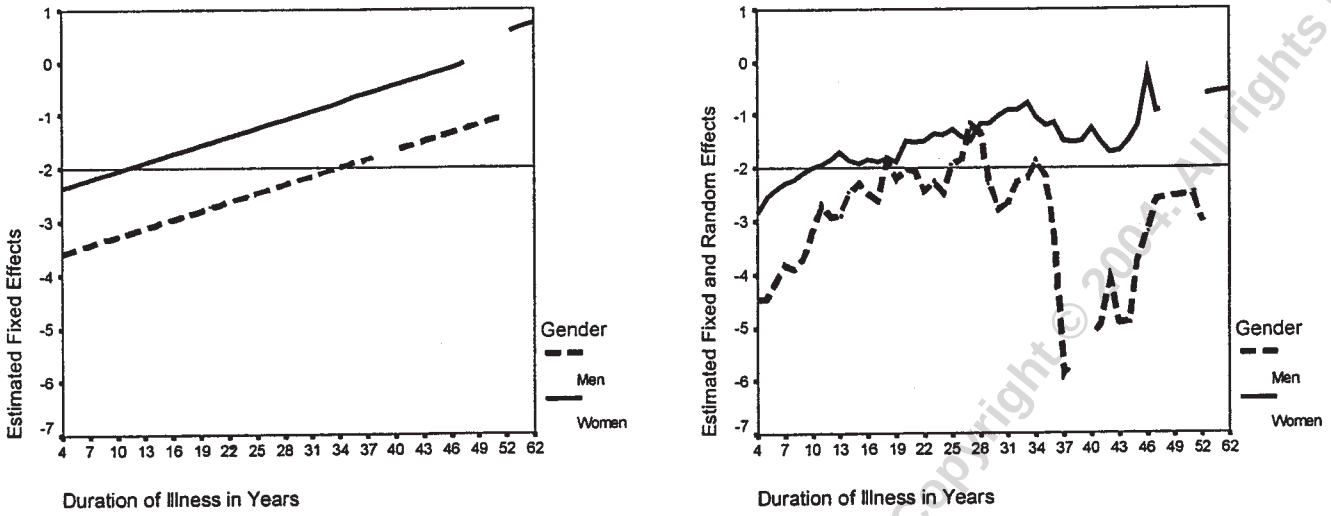

Figure 2. Estimated effects by duration of illness for men and women, fixed effects on the left, fixed effects plus random effects on the right.

Table 5. Estimates for age and sex as predictors of limitation in functional ability.

Effect	Estimate	p
Intercept	–1.9083	< 0.0001
Age	0.0178	0.0593
$(Age)^2$	–0.0004	0.4475
$(Age)^3$	0.0001	0.0086
Men	–1.2452	< 0.0001
Women	0	

tices. Our results support some of what has been reported in earlier studies, and is contrary if not contradictory to findings from other studies. In any case, our findings provide a clearer picture of the progression of functional limitations among these patients.

The minimum clinically meaningful difference in HAQ scores is of some dispute in the literature. Bruce and Fries discuss this discrepancy in figures and cite a range from 0.10 to 0.22 Likert units as being reported in the literature[14]. However, since these figures are based on arithmetic manipulation of ordinal data, they cannot be easily compared to the logit scores. Felson, *et al* report that in the American College of Rheumatology outcome measures for RA, a change of 20% is clinically meaningful[24]. Calculating 20% of the least-square mean of women, since they are the reference group in the mixed models, would be the clinically meaningful difference in our sample. The least-square mean of women in the mixed model that included duration was –1.6975. The absolute value of 20% of this mean is 0.338. The mixed model found that men had 1.23 logits more functional ability than women; this is more than 3 times the

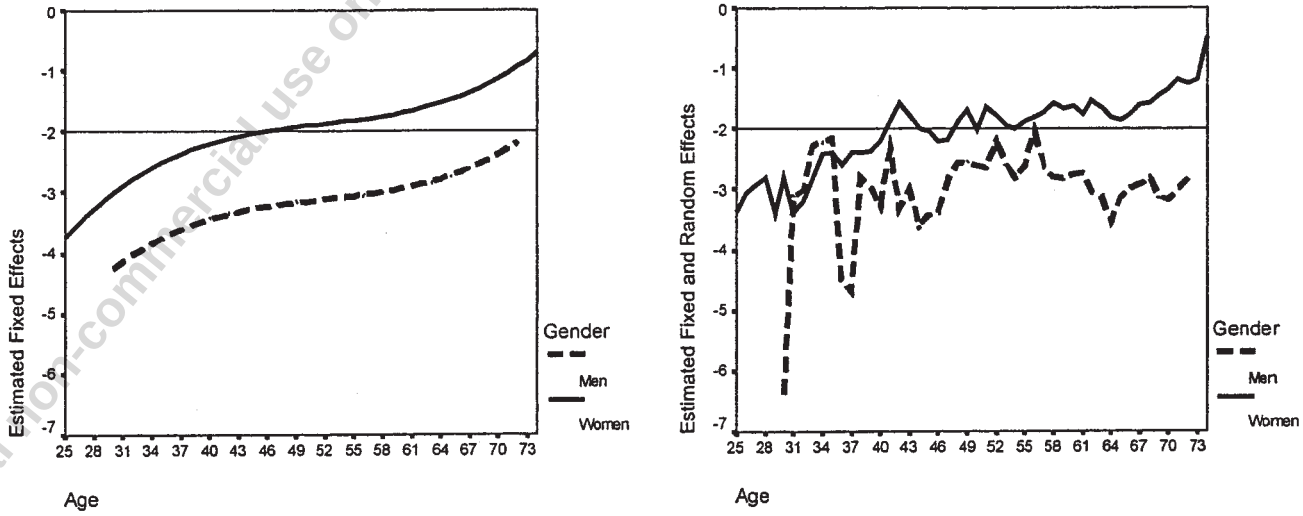

Figure 3. Estimated effects by age for men and women, fixed effects on the left, fixed effects plus random effects on the right.

minimum clinically meaningful difference of 0.338. Therefore, women not only have statistically significantly more functional limitation than men, but also clinically significantly more functional limitation.

RA functional limitation estimated from duration of the patient's illness described by the graphs in Figure 2 progresses in a linear manner and reaches a level where women experience some difficulty with the harder day-to-day activities at about 18 years' duration. This finding runs counter to the generally held belief that RA patients experience serious declines in functional ability in the early years of arthritis, with tapering levels of functional limitations as disease duration increases[6,7]. The study also finds that duration of illness is a better predictor of functional limitations than age, a finding that runs counter to Wolfe[3], who concluded that "the data of this study suggest duration of disease is only weakly related to HAQ score...", perhaps due to the small sample of 50 patients in the longitudinal analysis that led to this conclusion. Our results concur with Leigh[1] and Wolfe[3], who also found that functional limitations are slow to develop. More longitudinal studies with extensive periods of observation are needed to determine whether it is, in fact, slow to develop and to identify factors aside from disease duration, age, and sex that influence longterm outcomes among arthritis patients.

Our results also reveal a significant sex effect on functional limitation levels. The graph in Figure 2 shows men lagging 1.23 Rasch units behind women in overall levels of functional limitations, but limitations progress at the same rate for men and women. Typically, the age of onset of RA is between the ages of 20 and 45[25], with peak incidence between the fourth and sixth decade[26], which suggests that many men with RA will have difficulty by their early seventies, and women by their early sixties. The finding of worse functional limitations in women is consistent with the work of Sherrer, *et al*[6] and Wiles, *et al*[7]. However, Leigh, *et al*[1] found sex variation in the rate of change, a finding inconsistent with our observation that women and men progress at about the same rate. These inconsistencies among studies regarding effect of sex point to the need for more focused longitudinal studies on sex differences and the underlying mechanisms that account for these differences, if they exist. The question remains whether there is some underlying biological difference based on sex or whether the greater limitations women experience are associated with social roles or willingness to report activity limitations, or possibly because of bias in the HAQ activities, which may be more sensitive to functional limitations in women than in men.

Women have more distress over health problems and are more likely to report chronic diseases than men. They are also more likely to rate their health as being slightly worse than men[27,28]. This may account for some of the difference we found between men and women and their functional limitations. However, there is more than 3 times the minimal clinically meaningful difference between sexes, as found in the mixed model including duration of disease.

Biases can play a role in interpreting the results of a study. Selection bias in this study was minimized by the random selection of rheumatologists. However, there may be a sampling bias, as those RA patients who saw a rheumatologist were the only patients eligible for our study. Measurement bias was minimized by all patients being interviewed using the same protocol and interviewed by trained interviewers following a script. What role might attrition bias have had on these findings? As noted, this study dealt with the last 6 years of followup in a 10-year study, because the 20-item HAQ was available only for those years. This study effectively began at year 5, with 700 patients. There were complete data on 402 patients for all 6 years of followup. However, our mixed model was able to accommodate whatever responses were available from all 700 patients, with 332 women responding to all 6 interviews and 209 women to 5 or fewer interviews (Table 2). It also appears from Table 1 that the 700 patients participating in this study were similar to the 288 patients who were lost to followup by year 5. There were 552 (55.9%) patients who had either refused to continue the study, died, or were lost to followup. Although employing a different model than Leigh, *et al*[1], Reisine, *et al*[29] studied the difference between the participants who remained in the study the entire 10 years and those that did not. They found that participants were more educated, female, had moderate to high levels of social support, had fewer joint groups with flares, and were employed. Disease stage, level of pain, HAQ score, and duration of RA were not significantly different between participants completing the study and those not completing the study.

Finally, because of the sample and methods, our study offers a baseline or benchmark for comparison of future studies of functional limitations over time, especially to determine whether or not new treatments alter or extend the trajectories of functional limitations in patients in general and for women in particular. Participants in The National Rheumatoid Arthritis Study are relatively young, ranging in age from 21 to 65 years at the beginning of the study. The followup period spans 10 years, allowing the study of patients between the ages of 21 and 74. Functional limitations increase in the general population with age, and people with RA are likely to display more functional limitations beyond age 75, a limitation of this study, unless followup is resumed. One advantage of using the Rasch model is that as more ADL/IADL or HAQ research is reported on the Rasch metric, comparisons can be made across studies and across populations.

ACKNOWLEDGMENT

We thank Dr. Naomi Rothfield for her helpful comments on the manuscript. We are also grateful for the helpful comments of the anonymous reviewers.

REFERENCES

1. Leigh JP, Fries JF, Parikh N. Severity of disability and duration of disease in rheumatoid arthritis. J Rheumatol 1992;19:1906-11.
2. Wolfe F, Hawley DJ. The longterm outcomes of rheumatoid arthritis: work disability: a prospective 18 year study of 823 patients. J Rheumatol 1998;25:2108-17.
3. Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. Arthritis Rheum 2000;43:2751-61.
4. Schuna AA, Megeff C. New drugs for the treatment of rheumatoid arthritis. Am J Health Syst Pharm 2000;57:225-34.
5. Taylor PC. Anti-tumor necrosis factor therapies. Curr Opin Rheumatol 2001;13:164-9.
6. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. Arthritis Rheum 1986;29:494-500.
7. Wiles N, Dunn G, Barrett E, Silman A, Symmons D. Associations between demographic and disease-related variables and disability over the first five years of inflammatory polyarthritis: a longitudinal analysis using generalized estimating equations. J Clin Epidemiol 2000;53:988-96.
8. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857-67.
9. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: MESA Press; 1980.
10. Reisine S, Fifield J. Expanding the definition of disability: implications for planning, policy and research. Milbank Q 1992;70:491-509.
11. Rodnan G. Primer on the rheumatic diseases. 8th ed. Atlanta: Arthritis Foundation; 1983.
12. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
13. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10:20-30.
14. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. Health Qual Life Outcomes 2003;1:20-5.
15. Ramey DR, Raynauld JP, Fries JF. The Health Assessment Questionnaire 1992: status and review. Arthritis Care Res 1992;5:119-29.
16. Baker F. Item response theory: parameter estimation techniques. New York: Marcel Dekker; 1992.
17. Andrich D. A rating formulation for ordered response categories. Psychometrika 1978;43:561-73.
18. Wright B. Comparing Rasch measurement and factor analysis. Structural Equation Modeling 1996;3:3-24.
19. Sheehan TJ, DeChello LM, Garcia R, Fifield J, Rothfield N, Reisine S. Measuring disability: application of the Rasch model to activities of daily living (ADL/IADL). J Outcome Meas 2000;5:839-63.
20. Linacre JM. Facets. Chicago: MESA Press; 1999.
21. SAS Institute. SAS. Cary, NC: SAS Institute; 1999.
22. Pinheiro JC, Bates DM. Mixed-effects models in S and S-PLUS. New York: Springer-Verlag; 2000.
23. Akaike H. A new look at the statistical model identification. IEEE Trans Automatic Control 1974;19:716-23.
24. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727-35.
25. American College of Rheumatology. Rheumatoid arthritis. Internet [cited April 28, 2004]. www.rheumatology.org/public/factsheets/ra.asp?aud=pat
26. Klippel J, Weyand CM, Wortman RL, editors. Primer on the rheumatic diseases. Atlanta: Arthritis Foundation; 1997.
27. Verbrugge LM. Gender and health: an update on hypotheses and evidence. J Health Soc Behav 1985;26:156-82.
28. McDonough P, Walters V. Gender and health: reassessing patterns and explanations. Soc Sci Med 2001;52:547-59.
29. Reisine S, Fifield F, Winkelman DK. Characteristics of rheumatoid arthritis patients: who participates in long-term research and who drops out? Arthritis Care Res 2000;13:3-10.