MEDICAL CARE
Volume 39, Number 5, pp 413–424


# Economic Evaluation of Aquatic Exercise for Persons With Osteoarthritis

Donald L. Patrick, PhD, MSPH,*§ Scott D. Ramsey, MD, PhD,†§ Anna C. Spencer, MPH,* Susan Kinne, PhD,* Basia Belza, PhD, RN,‡ and Tari D. Topolski, PhD*

**Objectives. To estimate cost and outcomes of the Arthritis Foundation aquatic exercise classes from the societal perspective.**

**Design. Randomized trial of 20-week aquatic classes. Cost per quality-adjusted life year (QALY) gained was estimated using trial data. Sample size was based on 80% power to reject the null hypothesis that the cost/QALY gained would not exceed $50,000.**

**Subjects and Methods. Recruited 249 adults from Washington State aged 55 to 75 with a doctor-confirmed diagnosis of osteoarthritis to participate in aquatic classes. The Quality of Well-Being Scale (QWB) and Current Health Desirability Rating (CHDR) were used for economic evaluation, supplemented by the arthritis-specific Health Assessment Questionnaire (HAQ), Center for Epidemiologic Studies-Depression Scale (CES-D), and Perceived Quality of Life Scale (PQOL) collected at baseline and postclass. Outcome results applied to life expectancy tables were used to estimate QALYs. Use of health care facilities was assessed from diaries/questionnaires and Medicare reimbursement rates used to estimate costs. Nonparametric bootstrap sampling of costs/QALY ratios established the 95% CI around the estimates.**

**Results. Aquatic exercisers reported equal (QWB) or better (CHDR, HAQ, PQOL) health-related quality of life compared with controls. Outcomes improved with regular class attendance. Costs/QALY gained discounted at 3% were $205,186 using the QWB and $32,643 using the CHRD.**

**Conclusion. Aquatic exercise exceeded $50,000 per QALY gained using the community-weighted outcome but fell below this arbitrary budget constraint when using the participant-weighted measure. Confidence intervals around these ratios suggested wide variability of cost effectiveness of aquatic exercise.**

**Key words: Cost-effectiveness; aquatic exercise; osteoarthritis; quality of life (Med Care 2001;39:413–424)**

Osteoarthritis is the most common form of arthritis, causing pain and discomfort, limiting independence, and reducing quality of life in 21 million affected persons in the United States.[1,2,3] Arthritis ranks first or second as a cause of long-term disability, work disability, restricted activity

*From the Department of Health Services, University of Washington, Seattle.

†From the Department of Medicine, University of Washington, Seattle.

‡From the Department of Biobehavioral Nursing and Health Systems, University of Washington, Seattle.

§From the Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington.

This work performed at the University of Washington was funded by the Centers for Disease Control and Prevention, grant number, U48/CCU00954.

Previous versions of these analyses were presented at meetings of the Society for Medical Decision Making and American Public Health Association.

Address correspondence and reprint requests to: Donald L. Patrick, University of Washington, Department of Health Services, 1959 NE Pacific Street, Box 357660, Seattle, WA 98195-7660. E-Mail: donald@u.washington.edu

Received January 10, 2000; initial review completed March 27, 2000; accepted December 22, 2000.

days, medical visits, and prescription and nonprescription drug use among persons with musculoskeletal conditions.[3] The impact of osteoarthritis on the national economy is well documented. Estimates of direct and indirect medical costs range from \$54.6 billion to \$65 billion yearly.[1,3] Several nonpharmacologic interventions have been reported to be cost-effective or even cost-saving in the treatment of osteoarthritis, including patient education, self-help courses, and joint replacement of the hip and knee.[4–8]

Prevention of osteoarthritis, including secondary prevention, is challenging.[7] The National Arthritis Action Plan emphasizes physical activity as one of several important interventions for minimizing pain and disability.[9] Clinical trials have shown that exercise therapies that emphasize muscle strengthening and range of motion may reduce discomfort of osteoarthritis, increase functional ability, and combat depression.[10–15] Moreover, aquatic therapy has been shown to increase strength and range of motion in affected joints, improve functional ability, and decrease joint pain.[13,16] Although the efficacy of exercise has been demonstrated in clinical trials, the cost-effectiveness of exercise programs, including aquatic therapy for people with osteoarthritis, has not been established. Because exercise programs entail a substantial commitment on the part of the participant, can be costly, and could potentially affect millions of persons with osteoarthritis, evaluating the cost-effectiveness of these programs is important.

The University of Washington, in collaboration with the Washington State Chapter of the Arthritis Foundation, designed a naturalistic, community-based, randomized-controlled trial to evaluate the impact of warm water exercise on health outcomes and medical care costs for persons with osteoarthritis. The study included a community preference-weighted outcome measure (the Quality of Well-Being Scale [QWB]) and a participant-weighted outcome measure (Current Health Desirability Rating [CHDR]) so that the cost-utility of warm water exercise could be estimated.

## Methods

### Sample Size Estimation

The sample size for the trial was chosen to have 90% power to detect a 0.038 ($\alpha = 0.05$) mean difference in QWB scores between the intervention and control groups. This effect size was shown to be clinically important in previous studies that used the QWB for assessing outcomes of patients with osteoarthritis and rheumatoid arthritis.[17,18,19] The projected age and gender distribution of participants was based on information provided by the Washington State Chapter of the Arthritis Foundation. Accounting for a possible attrition rate of approximately 30% per arm, the recruitment goal was set at 125 participants in each group. Using cost data from a previous study of osteoarthritis[4] and a method for determining sample size needs for cost-effectiveness studies,[20] it was determined that 71 per group would achieve 80% power to reject the null hypothesis that the undiscounted incremental cost-effectiveness of the arthritis exercise program was more than \$50,000 per QALY, a commonly used but arbitrary budget constraint.

### Recruitment and Subject Eligibility

Between March 1997 and December 1997, 249 participants were recruited throughout the state of Washington using direct invitation letters to Arthritis Foundation members, notices in their newsletter, network television coverage of the study, physician referrals, public service announcements, and newspaper advertisements.[21] Interested adults who called a toll free number were screened for eligibility using criteria established by the Arthritis Foundation and study investigators. Eligibility criteria were as follows:

1) Clinically confirmed diagnosis of osteoarthritis from a physician.
2) Aged 55 to 75.
3) Not currently exercising, defined as engaging in an average of less than 60 minutes of exercise per week during the last month.
4) Permission by the subject's primary physician to participate in the aquatic class.
5) Not currently enrolled in another medical study.
6) Not scheduled for joint replacement surgery during the study period.
7) Living in an area where Arthritis Foundation aquatic programs were offered.
8) Willingness to be randomized and to commit to the 5-month study period.

The Human Subjects Division of the University of Washington approved all procedures.

### Randomization

Participants were randomized to the treatment or control group using a stratified randomization process.[22] Previous research suggested that outcomes of exercise might be significantly different for men and women with osteoarthritis, so stratification based on sex was used to obtain a sufficient number of men in the treatment and control groups. Randomization within each stratum took place once a week after recruitment began, which evenly distributed early and late recruits between groups to control for differences that might exist between earlier and later recruits.

### Aquatic Exercise Protocol

After randomization, persons in the treatment group were directed to an Arthritis Foundation certified aquatic class. The aquatic program, run by certified instructors, is held in pools with a temperature of 85° F to 92° F. Participants engage in gentle upper- and lower-body activities to help increase joint flexibility and range of motion, and maintain muscle strength. Treatment group participants were asked to attend class at least twice weekly for the 20-week study period. The number of classes offered per week varied from 2 to 7, class length ranged from 45 minutes to 1 hour, and class size ranged from 6 to 40 persons with an average of 16. Participants in the control group were asked to follow their usual pattern of activities and to abstain from new exercise programs for the duration of the study, after which they were invited to attend an Arthritis Foundation aquatic class free of charge.

Participants in both groups completed questionnaires at baseline and 20 weeks (postclass). Additionally, all participants were asked to complete a weekly postcard diary for 20 weeks that provided data on use of health care facilities and class attendance. Participants were paid $10 for completing each questionnaire. Age, gender, marital status, work status, education, race, and grouped household income of participants were reported on the baseline questionnaire.

### Outcome Measures

Outcome measures included 2 preference-weighted health status measures to estimate cost-utility ratios (1 community and 1 participant-specific)[23] and 3 nonpreference-weighted measures to examine other outcomes of equal interest.[24] The preference-weighted outcomes were the QWB,[25] a generic measure of health status with community-derived preference weights for use in cost-effectiveness analyses and previously used in evaluating other types of arthritis therapy,[18] and the single-item CHDR, in which participants were asked to rate the desirability of their current health on a 0 to 100 scale, ranging from 0 for "least desirable" to 100 for "most desirable." This item, adapted from a desirability-rating question used in a previous study,[26] is specific to the participants with osteoarthritis. Nonpreference-weighted outcomes included assessments of arthritis-specific function (Health Assessment Questionnaire [HAQ]),[27] perceived quality of life (Perceived Quality of Life Scale [PQOL]),[28,29] and depressive symptoms (Center for Epidemiological Studies-Depression Scale [CED-S).[30]

### Use and Cost of Arthritis-Specific Services

Using the weekly postcard diary, participants recorded use of primary care physicians, arthritis specialists, podiatrists, chiropractors, surgeons, physical and occupational therapists, home health care nurses, acupuncture therapy, massage therapy, and 'other' arthritis-related care (ie, wax therapy, hot gloves, transcutaneous electrial nerve stimulator [TENS] unit). Costs for medical care goods and services were based on the 1997 Medicare reimbursement rates for the state of Washington. The health care provider visits were matched to Current Procedural Terminology codes of mid-level complexity[31] and then to corresponding Medicare reimbursement rates. Nontraditional health care, such as massage and acupuncture, was assigned average costs ascertained from a survey of 10 Seattle area providers.

Participants were asked to report use of arthritis-specific drugs, aids, and devices on questionnaires given at the beginning and end of the study period. Medications itemized included over the counter pain relievers, nonsteroidal anti-inflammatory drugs (NSAIDs), steroids, anti-rheumatic medications, muscle relaxants, anti-depressants or sleep aids, and other prescription drugs. Medication prices were based on average wholesale price from the *1997 Drug Topics Red Book*.[32] Average expenditures over each time period were based on the amount taken during the past week. Costs of aids and devices—such as canes, walkers, wheelchairs, special utensils, removable splints, bathroom aids, and appliances for grip and reach—were an average obtained from a survey of 10 Seattle area special equipment retail stores. Arthritis-related household or chore worker help was recorded on the weekly postcard. The estimated cost of a chore worker visit was the average derived from 10 local housekeeping agencies.

### Exercise Intervention-Related Costs

Recruitment costs included production of direct mail flyers, recruitment letters to active Arthritis Foundation members, television and newspaper coverage, and public service announcements.[21] Participant costs included class fees, transportation costs, and time costs associated with traveling to and from and participating in the classes. Class fees for the 20 weeks ranged from $0 to $262 per person (these costs were waived for participants but included in the analysis). Class fees included the cost of renting the pool and the cost of the instructor. Transportation costs were based on the mileage between participant homes and the aquatic facilities where they attended classes multiplied by $0.31/mile. Distances were obtained from a World Wide Web directory that computes travel distances between destinations.[33]

The time cost associated with travel and participating in the exercise classes was based on the median personal income in 1997 for the modal participant (women, white, age 55–68) enrolled in the study.[34] No widely agreed upon valuation of time exists for retired persons, and thus we included modal values in our analyses.

### Primary Analysis

After randomization but before the intervention, treatment group characteristics were compared using a generalized linear model analysis of variance. Analysis of covariance (SPSS GLM[35]) was used to compare changes in outcomes between treatment and control groups over time. An intent-to-treat approach was used to compare treatment and control groups; that is, participants assigned to the aquatic therapy arm who did not attend some or all classes were still included in the intervention group for analysis. A generalized linear model was used to compare QWB and Health Desirability measures for adherers, nonadherers, and controls.

Cost estimates were based on annualized projections from the 20-week observation period. Recruitment costs were the same for treatment and control participants. Although we observed one class for which participants were not charged, the average class fee was applied to all participants in the treatment group, to incorporate the societal perspective and the costs of providing the intervention.

Reference case cost-utility results were computed using the societal perspective and following the recommendations of the Panel on Cost-Effectiveness in Health and Medicine.[23,36] Quality-adjusted life years (QALYs) for the intervention and control groups were calculated by multiplying the mean postclass QWB score and CHDR by expected years of remaining life for the participants. Expected years of remaining life, based on the ages of participants, were obtained from life tables.[37] Future costs and QALYs were discounted at 3%.[36]

### Uncertainty Analysis

The bias-corrected nonparametric bootstrap method was used to derive a 95% CI for the incremental cost-effectiveness ratio, using QALY results derived from the QWB and CHDR.[38,39] One thousand samples with replacement were drawn for the bootstrap estimates. Estimates were located on a scatterplot indicating different levels of cost and effectiveness.

## Results

Table 1 displays characteristics of participants by randomization status. The sample contained a significantly smaller proportion of males than expected (13.7% vs. 20%; $\chi^2_1 = 4.94$ $P < 0.05$); however, the proportion of males in the treatment and control groups was not significantly different.

No significant differences were found between the groups on baseline demographic characteristics, health status, or utilization measures, except for number of medications used weekly. One participant's medication count was more than 2 standard deviations above the mean. When this score was excluded from the analysis, the difference was no longer significant. Based on these

TABLE 1. Sample Characteristics of Participants by Randomization Status

| | Entire Group N = 249 | Control N = 124 | Treatment N = 125 |
|---|---|---|---|
| Age (mean, y) | 65.7 | 66.1 | 65.7 |
| Gender (% female) | 86.3 | 87.1 | 85.3 |
| Income | | | |
| Below $10,000 (%) | 8.9 | 11.7 | 6.1 |
| Between $10–$39,999 (%) | 56.2 | 55.9 | 56.5 |
| Above $40,000 (%) | 34.9 | 32.5 | 37.3 |
| Insurance: | | | |
| None (%) | 1.2 | 0.8 | 1.6 |
| Medicaid (%) | 1.2 | 1.6 | 0.8 |
| Medicare (%) | 55.7 | 58.9 | 52.5 |
| Private (%) | 41.9 | 38.7 | 45.1 |
| Martial status (% married) | 59.8 | 60.0 | 59.7 |
| Work status (% unemployed) | 85.1 | 84.0 | 86.3 |
| At least high school education | 69.1 | 72.8 | 65.3 |
| Live alone (%) | 32.1 | 33.6 | 31.1 |
| Race (% white) | 94.0 | 96.0 | 92.0 |

analyses, it was concluded that the randomization process was successful.

Twenty-one participants (16.8%) in the treatment group and 3 (3%) in the control group did not complete the study. No significant differences were found in the characteristics of the dropouts and completers at baseline. Eighty-nine percent (n = 222/249) of the sample returned questionnaires at both time points and 90.3% (n = 225/249) returned all 20 of the weekly diaries. QWB scores could be calculated for all participants for both baseline and post class observations. Information on the CHDR scale was available from 216 (86%) participants used in the analyses. Complete (20 weeks) diary data were available for 92.3% (96/104) of treatment group participants and 96.6% (117/121) of the control group.

Group means, standard deviations, and univariate results from the analysis of covariance, controlling for individual differences in the outcome measures, are presented in Table 2. The QWB score for the treatment group improved only slightly between Baseline and Post Class, while the mean score remained the same for the control group; however, this difference between groups was not statistically significant. A significant mean difference between groups was found, however, for the participant-specific CHDR score with the treatment group reporting a higher mean rating after class, whereas the mean rating for the control group declined. Treatment group scores were also significantly improved for the disability measure of the HAQ and the physical domain score of the PQOL.

## Adherence to Exercise Protocol

Adherence to the exercise protocol, as assessed by report of class attendance from the weekly postcard, indicated that 36 (29%) of the Treatment Group participants attended classes a minimum of twice per week for at least 16 weeks. Contrast analyses showed that the mean QWB score for adherers (mean = 0.613) was significantly higher ($P = 0.017$) than for the controls (mean = 0.599) or for the nonadherers (mean = 0.602; [$P = 0.01$]) after controlling for individual differences in age, gender, number of medical conditions, and this measure at baseline. For the CHDR, there was insufficient power to detect a significant difference between the means for the three groups. A multivariate general linear model was computed on baseline: after class difference scores with age, gender, and number of medical conditions entered as covariates to assess whether there were significant differences in the means on the other out-

TABLE 2. Health Outcome Measures at Baseline and Postclass Comparing Treatment and Control Groups

| | Treatment Group | | | Control Group | | | |
|---|---|---|---|---|---|---|---|
| Health Outcome | Mean | SD | (n) | Mean | SD | (n) | *P*-value* |
| **Preference-Weighted** | | | | | | | |
| QWB† | | | | | | | |
| Baseline | 0.597 | 0.068 | 125 | 0.599 | 0.065 | 124 | |
| Postclass | 0.606 | 0.069 | 101 | 0.599 | 0.079 | 121 | 0.495 |
| Health Desirability | | | | | | | |
| Baseline | 62.32 | 16.95 | 123 | 62.52 | 16.78 | 120 | |
| Postclass | 65.80 | 15.61 | 100 | 61.78 | 15.91 | 116 | 0.035 |
| **Non-Preference Weighted** | | | | | | | |
| HAQ‡ disability | | | | | | | |
| Baseline | 1.035 | 0.535 | 125 | 1.047 | 0.608 | 124 | |
| Postclass | 0.933 | 0.550 | 101 | 1.127 | 0.671 | 121 | 0.015 |
| HAQ‡ pain | | | | | | | |
| Baseline | 1.533 | 0.602 | 121 | 1.440 | 0.610 | 123 | |
| Postclass | 1.382 | 0.737 | 98 | 1.462 | 0.619 | 117 | 0.660 |
| PQOL§ physical | | | | | | | |
| Baseline | 5.741 | 1.622 | 122 | 5.919 | 1.729 | 124 | |
| Postclass | 6.396 | 1.697 | 101 | 5.790 | 1.752 | 121 | 0.007 |
| CES-D¶ | | | | | | | |
| Baseline | 7.261 | 5.308 | 123 | 7.715 | 4.995 | 120 | |
| Postclass | 6.956 | 4.729 | 101 | 8.092 | 6.005 | 113 | 0.096 |

*Analysis of covariance controlling for baseline health outcome measures.
†Quality of Well-Being Scale.
‡Health Assessment Questionnaire.
§Perceived Quality-of-Life Scale.
¶Center for Epidemiological Studies Depression Scale.

come measures. Because the multivariate analysis showed that there were significant differences in the means between the groups on the set of dependent variables (Wilk's Lambda = 0.587, $P = 0.000$), the univariate results were examined. Table 3 shows the group means, standard deviations, and univariate results for all outcome measures comparing controls with nonadherers and adherers. The univariate results showed that there were significant differences between the adherers and both the nonadherers and the controls on the QWB. Also, a significant difference between the adherers and the controls was noted on the arthritis specific measure of the HAQ functional disability index. No other significant differences were noted.

## Cost Results

All costs were annualized in 1997 dollars (Table 4). Recruitment and promotion costs averaged $56 per person. The average annual cost of the exercise classes applied to each treatment group participant was $291. Time costs of attending class for the treatment group averaged $9.68 per class or $894 on an annual basis. The average annual direct medical and direct nonmedical costs (including time costs) for the treatment group were $1,945 (SD = $1,690; Range = $13,835) and $1,688 (SD = $599; Range = $3,127), respectively. Average direct medical costs for the control group were $2,922 (SD = $2,750; Range = $16,463). Average direct nonmedical costs for the control group were $260 (SD = $628; Range = $3,990).

## Reference Case Results

The results of the incremental cost-effectiveness of the exercise program using costs and utility weights from the QWB are reported in Table 5. Undiscounted lifetime osteoarthritis-related costs from the societal perspective were $8,328 higher

TABLE 3. Health Outcome Measures at Pre-class and Post-class Comparing Adherers, Non-Adherers and Controls Among Those Who Completed the Study

| Health Outcome | Controls (n = 121) Mean | Controls (n = 121) SD | Non-Adherers (n = 68) Mean | Non-Adherers (n = 68) SD | Adherers (n = 36) Mean | Adherers (n = 36) SD |
|---|---|---|---|---|---|---|
| **Preference-Weighted** | | | | | | |
| Pre-QWB | 0.599 | 0.08 | 0.606 | 0.08 | 0.581 | 0.07 |
| Post-QWB* | 0.599 | 0.08 | 0.602 | 0.08 | 0.613¶‖ | 0.06 |
| Pre-self-rated health | 3.02 | 0.80 | 3.06 | 0.75 | 2.94 | 0.75 |
| Post-self-rated health | 2.93 | 0.76 | 3.16 | 0.67 | 3.09 | 0.78 |
| Pre-health desirability | 62.31 | 16.72 | 62.86 | 17.74 | 62.64 | 15.70 |
| Post-health desirability | 61.64 | 15.95 | 65.47 | 15.63 | 66.39 | 15.79 |
| **Non-Preference Weighted** | | | | | | |
| Pre-HAQ disability | 1.05 | 0.60 | 1.00 | 0.54 | 1.00 | 0.50 |
| Post-HAQ† disability | 1.13 | 0.67 | 0.95 | 0.57 | 0.91¶ | 0.51 |
| Pre-HAQ pain | 1.45 | 0.61 | 1.46 | 0.65 | 1.60 | 0.55 |
| Post-HAQ† pain | 1.46 | 0.62 | 1.37 | 0.68 | 1.40 | 0.86 |
| Pre-PQOL physical | 5.92 | 1.73 | 5.83 | 1.65 | 5.90 | 1.53 |
| Post-PQOL‡ physical | 6.69 | 1.82 | 6.95 | 1.55 | 7.10 | 1.50 |
| Pre-CES-D | 7.73 | 5.03 | 6.81 | 4.84 | 7.72 | 6.15 |
| Post-CES-D§ | 8.09 | 6.00 | 7.11 | 4.70 | 6.68 | 4.84 |

*Quality of Well-Being Scale.
†Health Assessment Questionnaire.
‡Perceived Quality of Life Scale.
§Center for Epidemiological Studies Depression Scale.
¶Simple contrasts on baseline-postclass difference scores significantly different from control $P < 0.05$.
‖Simple contrasts on baseline-postclass difference scores significantly different from non-adherers $P < 0.05$.

for the aquatic exercise group than for the usual care control group. Using the 3% discount rate, lifetime incremental costs for the program were $11,363. Using the QWB with community preference weights, the undiscounted incremental cost-effectiveness for the exercise program was $69,400/QALY and the discounted rate $205,186/QALY gained. Using the bootstrap procedure, the 95% CI around the reference case estimate ranged from the aquatics class being dominant (ie, less expensive and more effective, Quadrant 2) to being dominated by the usual care (ie, more expensive and less effective, Quadrant 4) (Fig. 1). Overall, 24% of the bootstrap replicates showed the intervention being dominated by usual care. Eight percent showed the intervention to be cost-saving with higher QALYs.

Incremental cost-effectiveness results for the reference case using the CHDR weighted by participants is shown in Table 6. Using the participant-generated preference weights from the CHDR to derive QALYs, the discounted incremental cost/QALY gained for the exercise intervention was $32,643. The 95% CI ranged from being dominated to $498,700/QALY gained. Overall, as shown in Fig. 2, the bootstrap replicates showed aquatic exercise being dominated by usual care (Quadrant 4). Seven percent showed the intervention to be cost-saving with higher QALYs (Quadrant 2).

## Discussion

This randomized, community-based study of aquatics-based exercise program for persons with osteoarthritis did not demonstrate reduced costs and improved health outcomes compared with usual care. Direct medical care costs fell slightly for those in the intervention group, but were more than offset by the program costs and costs associated with taking the classes. As a result, the incremental cost-utility of the exercise program (using the QWB community-derived preference weights) was not favorable compared with other common health care interventions. Because the baseline change in utility was small for the intervention group, the confidence intervals derived around the mean incremental cost-utility value were extremely large. We thus cannot reject the

TABLE 4. Mean Direct and Nondirect Annualized Health Care Costs* for Treatment and Control Groups

| | Treatment (N = 104) | | Control (N = 121) | |
|---|---|---|---|---|
| | Cost ($) | Participant (SD) | Cost ($) | Participant (SD) |
| Direct Medical | | | | |
| Medical Care† | | | | |
| Primary care physician | 253 | (314) | 389 | (478) |
| Other physicians | 132 | (212) | 244 | (349) |
| Other clinicians | 225 | (1129) | 368 | (1027) |
| Aids and devices | 126 | (206) | 394 | (1039) |
| Other treatments | 277 | (304) | 293 | (626) |
| Medications | 932 | (772) | 1234 | (1028) |
| Direct Nonmedical | | | | |
| Class fees | 291 | (00) | | — |
| Recruitment | 56 | (00) | 56 | (00) |
| Household help | 44 | (63) | 181 | (612) |
| Time costs | 894 | (303) | 23 | (97) |
| Participants travel‡ | 404 | (383) | | — |
| Total Costs | 3,634 | | 3,182 | |

*Costs reported in 1997 US dollars.
†Average number of visits over 20-week study period annualized.
‡Average dollar cost per miles traveled over study period.

null hypothesis that the cost-effectiveness of the intervention exceeds the selected budget constraint of $50,000 per QALY gained.

Time costs were estimated from median personal income of the modal participant, and these costs added considerably to the marginal cost of taking the aquatics classes. Because many of the participants were retired and on fixed incomes not tied to employment, we may have overstated the true time cost of participating. Reducing or eliminating these costs, however, would not have allowed us to reject the null hypothesis and would have led at best to a conclusion of cost-minimization based on the QWB health benefit measure.

Unlike the preference weights derived using the community-weighted QWB, preference weights elicited using the participant-specific CHDR improved significantly for the exercise group. As a result, the cost-utility outcome of the aquatics exercise program using the CHDR was more favorable; however, the 95% CI also does not allow us to reject the null hypothesis.

The nonpreference-weighted, arthritis-specific measures of health outcome indicated that aquatics exercise significantly reduced perceived disability specific to arthritis and significantly improved perceived quality of life specific to physical health. In light of the improvement observed in these specific domains, it is possible that the generic, community-weighted QWB was not sensitive enough to detect a significant difference at the sample size used.

TABLE 5. Lifetime Cost and QALY*Results for the Reference Case

| | Total Cost of Arthritis Care | | Total QALY based on QWB† | | Incremental Cost | | Incremental Cost«ALY Gained | |
|---|---|---|---|---|---|---|---|---|
| | Annualized | Lifetime | No Discount | 3%‡ | No Discount | 3%‡ | No Discount | 3%‡ |
| Aquatic class | $3,634 | $67,017 | 11.16 | 6.47 | $8,328 | $11,363 | $69,400 | $205,186 |
| Usual care | $3,182 | $59,689 | 11.04 | 6.40 | — | — | — | — |

*QALY indicates quality-adjusted life year.
†Quality of Well-Being Scale.
‡Based on 18.44 life years remaining.

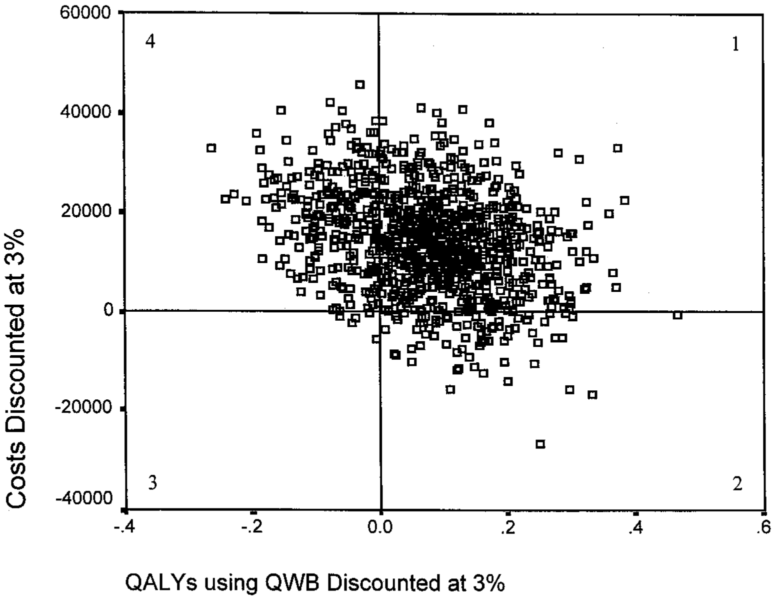


FIG. 1. Scatterplot of bootstrap replicates of costs per QALY gained or lost estimated using the Quality of Well-being Scale and Social Preference Weights.

Exercise participants in this study did not report improvements in pain. This finding contrasts with the cost-effectiveness modeling of the Arthritis Self-Help Course which showed cost savings and reductions in pain units,[7] an outcome measure not comparable to QALYs. Differences in outcome measures used and in costing methodologies make it difficult to compare these results.

A strength of this study was that it was conducted specifically as a randomized economic evaluation in a community setting. The design allowed us to compare the effectiveness of alternative measures of recruitment[21] and measure levels of adherence to the class over time in a naturalistic setting. Primary data collection on medical utilization and on outcomes was costly in time and resources, but it yielded high-quality data that have not been available from previous studies. Completion rates on the questionnaires and weekly diaries were excellent (90%), primarily because of close personal contact with each study participant in the treatment and delayed-treatment control group.

Less than a third of participants randomized to the treatment group met our criteria for adherence. Participation in regular exercise at recommended levels of frequency is still hard to motivate, even with economic incentives. Analyses

TABLE 6. Lifetime Cost and QALY* Results for Reference Case based on CHDR

| | Total Cost of Arthrtitis Care | | Total QALY Based on CHDR† | | Incremental Cost | | Incremental Cost«ALY Gained | |
|---|---|---|---|---|---|---|---|---|
| | Annualized | Lifetime | No Discount | 3%‡ | No Discount | 3%‡ | No Discount | 3%‡ |
| Aquatic class | $3,634 | $67,017 | 12.13 | 7.03 | $8,328 | $11,363 | $10,958 | $32,643 |
| Usual care | $3,182 | $59,689 | 11.37 | 6.59 | — | — | — | — |

*QALY indicates quality-adjusted life year.
†Current Health Desirability Rating.
‡Based on 18.44 life years remaining.

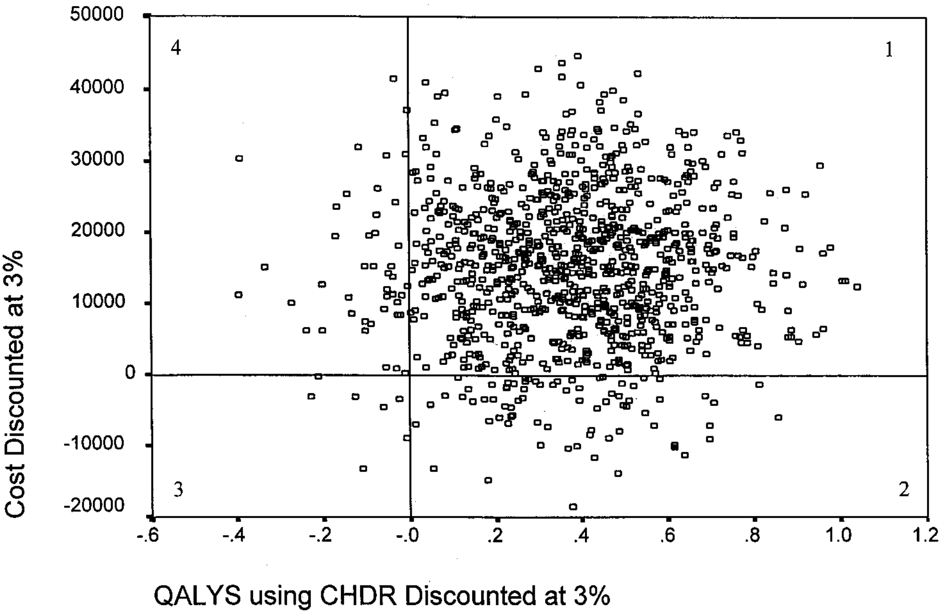


FIG. 2. Scatterplot of bootstrap replicates of costs per QALY gained or lost estimated using the Current Health Desirability Rating Scale and Participant Preference weights.

comparing adherers with nonadherers and control participants suggested that warm water exercise had a positive effect on health outcomes for those who regularly attended classes. Getting people to class regularly remains a significant barrier requiring motivational incentives. Incorporating incentives such as telephone follow-up, feedback, reinforcements, and rewards might augment and improve adherence though increase costs.[40,41]

All participants in this study were nonexercisers at baseline to permit estimation of the benefits of initiating regular exercise in this population. Excluding participants who were currently exercising limits the generalizability of the results. Participants in this study were highly educated, and although this is representative of membership in the Washington State Chapter of the Arthritis Foundation, the study sample was not representative of the population with osteoarthritis. This study was conducted in only one state. The Arthritis Foundation aquatic program, however, is standardized across the United States. This enhances generalizability of our results to other settings.

Projecting the results of an exercise intervention lasting 20 weeks to lifetime benefit does not model downstream adherence, length of exercise, and other possible scenarios. Moreover, it is highly unlikely that anyone will continue aquatic exercise throughout their lifetime. Our findings on adherence to exercise, however, appear realistic and the intent-to-treat analysis enhances the generalizability of the results. Our projections, and the use of stochastic analyses, helped to estimate cost and benefits similar to other economic evaluations of exercise and provides supportive evidence. Persons with osteoarthritis are higher users of health care, and are more likely to be over age 65. Preventive strategies will be needed to complement medical therapy as the proportion of the population over age 65, disproportionately affected by osteoarthritis, increases. Lower levels of adherence to exercise over the lifetime of persons with osteoarthritis would significantly reduce the potential for cost-effectiveness, and this remains to be evaluated in naturalistic settings.

## Acknowledgments

The Arthritis Foundation of Washington State provided assistance in recruitment. The authors thank Erin Maher for her participation in the design and early implementation of this study, Melinda T. Neri for assistance in recruitment and data collection, and Christine Fountain for research assistance.

## References


1. **Lawrence RC, Helmick CG, Arnett FC, et al.** Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998;41:778–799.

2. Centers for Disease Control and Prevention. Arthritis prevalence and activity limitations - United States, 1990. MMWR Morb Mortal Wkly Rep 1994;43(suppl 24):433–438.

3. **Rothfuss J, Mau W, Zeidler H, et al.** Socioeconomic evaluation of rheumatoid arthritis and osteoarthritis: a literature review. Arthritis Rheum 1997;26(suppl 5):771–779.

4. **Gabriel SE, Crowson CS, O'Fallon WM.** Costs of osteoarthritis: estimates from a geographically defined population. J Rheum 1995;43(suppl 22):23–25.

5. **Chang RW, Pellissier JM, Hazen GB.** A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. JAMA 1996;275:858–865.

6. **Kvien TK, Brors O, Staff PH, et al.** Improved cost-effectiveness ratio with a patient self-adjusted naproxen dosing regimen in osteoarthritis treatment. Scand J Rheumatol 1991;20(suppl 4):280–287.

7. **Kruger J, Helmick CG, Callahan LF, et al.** Cost-effectiveness of the arthritis self-help course. Arch Intern Med 1998;158:1245–1249.

8. **Lavernia CJ, Guzman JF, Gachupin-Garcia A.** Cost effectiveness and quality of life in knee arthroplasty. Clin Orthop 1997;345:134–139.

9. Arthritis Foundation, Association of State and Territorial Health Officials, Centers for Disease Control and Prevention. National Arthritis Action Plan: A Public Health Strategy. Atlanta, GA: Arthritis Foundation and National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention; 1999.

10. **Minor MA, Hewett JE, Webel RR, et al.** Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. Arthritis Rheum 1989;32(suppl 11):1396–1405.

11. **Schilke JM, Johnson JO, Housh TJ, et al.** Effects of muscle-strength training on the functional status of patients with osteoarthritis of the knee joint. Nurs Res 1996;45(suppl 2):68–72.

12. **Hopman-Rock M, Kraaimaat FW, Bijlsma JWJ.** Physical activity, physical disability and osteoarthritis pain in older adults. J Aging Phys Activ 1996;4:324–337.

13. **Templeton MS, Booth DL, O'Kelly WD.** Effects of aquatic therapy on joint flexibility and functional ability in subjects with rheumatic disease. J Orthop Sports Phys Ther 1996;23(suppl 6):376–381.

14. **Messier SP, Thompson CD, Ettinger WH Jr.** Effects of longer-term aerobic or weight training regimens on gait in an older, osteoarthritis population. J App Biomechanics 1997;13:205–225.

15. **Ettinger WH Jr, Burns R, Messier SP, et al.** A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. JAMA 1997;277:25–31.

16. **Suomi R, Lindauer S.** Effectiveness of Arthritis Foundation aquatic program on strength and range of motion in women with arthritis. J Aging Phys Activ 1997;5:341–351.

17. **Kaplan RM, Anderson JP, Wu AW, et al.** The quality of well-being scale: applications in AIDS, cystic fibrosis, and arthritis. Med Care 1989;27:S27-S43.

18. **Thompson MS, Read JL, Hutchings C, et al.** The cost effectiveness of auranofin: results of a randomized clinical trial. J Rheumatol 1988;15:35–42.

19. **Bombardier C, Ware J, Russell I, et al.** Auranofin therapy and quality of life in patients with rheumatoid arthritis. Am J Educ 1986;81:565–577.

20. **O'Brien BJ, Drummond MF, Labelle RJ, et al.** In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. Med Care 1994;32(suppl 2):150–163.

21. **Spencer AS, Kinne S, Belza BL, et al.** Recruiting adults with osteoarthritis into an aquatics exercise class: Strategies for a statewide intervention. Arth Care Res 1998;11(suppl 6):455–462.

22. *Microsoft® Access 97* [computer program]. Seattle, WA: Microsoft; 1996.

23. **Gold MR, Siegel JE, Russell LB, et al., ed.** Cost-effectiveness in Health and Medicine. New York, NY: Oxford University Press, 1996.

24. **Patrick DL, Erickson P.** Health Status and Health Policy. New York, NY: Oxford University Press; 1993.

25. **Kaplan RM, Anderson JP.** A general health policy model: update and applications. Health Serv Res 1988;23(suppl 2):203–235.

26. **Tsevat J, Cook EF, Green ML, et al.** Health values of the seriously ill. Ann Intern Med 1995;122(suppl 7):514–520.

27. **Fries JF, Spitz P, Young DL.** The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheum 1982;9(suppl 5):789–793.

28. **Patrick DL, Danis ML, Southerland LI, et al.** Quality of life following intensive care. J Gen Intern Med 1988;3(suppl 3):218–223.

29. **Patrick DL, Kinne S, Engelberg RA, et al.** Functional status and perceived quality of life in adults with and without chronic conditions. J Clin Epidemiol 2000;53(suppl 8):779–785.

30. **Kohout FJ, Berkman LF, Evans DA, et al.** Two shorter forms of the CES-D depression symptoms index. J Aging Health 1993;5(suppl 2):179–193.

31. American Medical Association. Physicians' Current Procedural Terminology: CPT 96. Chicago, IL: American Medical Association; 1995.

32. Medical Economics Company. 1997 Drug Topics Red Book. Montvale, NJ: Medical Economics Company; 1997.

33. Zip2 Corporation. Zip2 Yellow Pages, 1999. Available at: http://www.zip2.com. Accessed March, 31 1999.

34. US. Bureau of the Census. Statistical Abstract of the United States: 1999. Washington, DC: US Bureau of the Census; 1999.

35. SPSS 8.0 Base Manual for Windows: User's Guide. Chicago, IL: SPSS; 1998.

36. **Weinstein MC, Siegel JE, Gold MR, et al.** Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA 1996;276(suppl 15):1253–1258.

37. National Center for Health Statistics. Vital Statistics of the United States, 1995. Preprint of Vol II, mortality, part A, Sec 6, life tables. Hyattsville, MD: National Center for Health Statistics; 1998.

38. **Chaudhary MA, Stearns SC.** Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. Stat Med 1996;14:1447–1458.

39. **Drummond MF, O'Brien B, Stoddart GL, et al.** Methods for the Economic Evaluation of Health Care Programmes. 2nd ed. Oxford, UK: Oxford University Press; 1997.

40. **King AC, Taylor CB, Haskell WL, et al.** Identifying strategies for increasing employee physical activity levels: findings from the Stanford/Lockheed exercise survey. Health Educ Q 1990;17:269–285.

41. **Dishman R.** Determinants of participation in physical activity. In: Bouchard C, Shephard R, Stephens T, Sutton J, McPherson BD, eds. Exercise, Fitness, and Health: A consensus of current knowledge. Champaign, IL: Human Kinetics Books; 1990:78–101.