# Land-based Exercise for Osteoarthritis of the Knee: A Metaanalysis of Randomized Controlled Trials

### MARLENE FRANSEN and SARA McCONNELL

**ABSTRACT. Objective.** To determine if clinical guidelines recommending therapeutic exercise for people with knee osteoarthritis (OA) are supported by rigorous scientific evidence. To explore whether the magnitude of treatment benefit reported in randomized controlled trials (RCT) is associated with exercise dosage or study methodology.

*Methods.* We conducted a metaanalysis of RCT comparing some form of land-based therapeutic exercise with a nonexercise group using pain and self-reported physical function outcomes.

**Results.** The 32 included studies provided data on almost 3800 participants. Metaanalysis revealed a beneficial treatment effect: standardized mean difference (SMD) 0.40 [95% confidence interval (CI) 0.30 to 0.50] for knee pain; SMD 0.37 (95% CI 0.25 to 0.49) for physical function. While the pooled beneficial effects of the 9 RCT evaluating exercise programs providing fewer than 12 direct supervision occasions or the 9 RCT judged to have a low risk of bias remained significant and clinically relevant, the magnitude of treatment benefit pooled from these RCT was significantly smaller than the comparator group (12 or more supervision occasions, moderate to high risk of bias, respectively). The mode of treatment delivery (individual treatments, exercise classes, home program) was not significantly associated with the magnitude of treatment benefit.

*Conclusion.* There is evidence that land-based therapeutic exercise has at least short-term benefit in terms of reduced knee pain and physical disability for people with knee OA. The magnitude of the treatment effect was significantly associated with the number of direct supervision occasions provided and study methodology (assessor blinding, adequate allocation concealment). (J Rheumatol First Release May 15 2009; doi:10.3899/jrheum.090058)

Key Indexing Terms: OSTEOARTHRITIS RANDOMIZED CONTROLLED TRIAL

KNEE

EXERCISE METAANALYSIS

People with symptomatic knee osteoarthritis (OA) complain of deep, aching pain. In early disease, pain is intermittent and mostly associated with joint use. For many people, symptomatic disease will progress. The knee pain becomes chronic, often also present at rest and during the night. The daily functional activities required to remain independent become increasingly more difficult. In fact, knee OA is responsible for more disability in walking, stair climbing, and housekeeping in noninstitutionalized people aged 50 years and over than any other disease<sup>1,2</sup>. Ultimately, chronic knee OA can also lead to reduced physical fitness and increased risk of cardiovascular comorbidity<sup>3,4</sup>.

Altered biomechanics and reduced neuromuscular control, resulting in localized stress on the articular cartilage

From The George Institute, University of Sydney, Camperdown, Australia; and the Department of Medicine, St. Joseph's Health Care Centre, Toronto, Ontario, Canada.

Address reprint requests to Dr. M. Fransen, The George Institute, PO Box M201 Missenden Road, Camperdown NSW 2050, Australia. Accepted for publication January 30, 2009.

and increased joint loading rate, have important roles in both the initiation and progression of knee OA<sup>5-10</sup>. Currently, there is no known cure for OA. However, disease-related factors, such as impaired muscle function and reduced fitness, are potentially amenable to exercise. Most international clinical guidelines advocate therapeutic exercise as the first line of effective disease management for people with knee OA<sup>11,12</sup>. The objective of our metaanalysis was to determine whether these recommendations are supported by a high level of scientific evidence. If the majority of clinical trials conducted to date have been small or poorly conducted, the magnitude of treatment benefit reported may be overestimated. We also wished to explore whether the reported magnitude of treatment benefit was associated with aspects of exercise program dosage or delivery mode. Do we have evidence of what constitutes an optimal graded exercise program for people with knee OA?

#### MATERIALS AND METHODS

Five electronic databases were searched: Medline, EMBASE, CINAHL, PEDro (Physiotherapy Evidence Database), and the Cochrane Central Register of Controlled Trials (CENTRAL; The Cochrane Library) up to December 2007. The search strategies were specific to each database (Appendix: Medline search strategy).

Reports of randomized or quasirandomized controlled trials (RCT),

Dr. Fransen was supported by a National Health and Medical Research Council Career Development award.

M. Fransen, MPH, PhD, Senior Research Fellow, The George Institute, University of Sydney; S. McConnell, MSc, Senior Physiotherapist, Department of Medicine, St. Joseph's Health Care Centre.

published in the English language, comparing some form of land-based therapeutic exercise with a nonexercise group among people with either an established diagnosis of knee OA according to accepted criteria or selfreporting knee OA on the basis of chronic joint pain (without radiographic confirmation) were considered for inclusion. The exercise program could include any land-based (in contrast to water-based), but not perioperative, therapeutic regime aiming to relieve the symptoms of knee OA. The comparator (control) group could include any nonexercise intervention, including no treatment or waiting list. The trial needed to include an assessment of knee pain or self-reported physical function. For our metaanalysis, data from the outcomes assessment conducted immediately after completion of the treatment program (or the most immediate assessment post-treatment) have been used.

The 2 authors (SM, MF) independently screened retrieved clinical studies for inclusion, extracted data from all included studies, and conducted the methodological quality assessment. If agreement was not achieved at any stage, a third review author (Dr. Mary Bell) adjudicated.

*Methodological quality assessment.* The quality of each included RCT was evaluated according to 3 criteria: (1) blinding of outcomes assessment; (2) appropriate handling of withdrawals and dropouts; (3) adequate allocation concealment.

An overall assessment was then assigned: low risk of bias (all 3 criteria met); moderate risk of bias (1 or 2 criteria met); or high risk of bias (no criteria met).

*Data analysis*. Standardized mean differences (SMD) and a random-effects model were used to combine treatment effects, extrapolated from mean change scores and related standard deviations (Review Manager 5, The Cochrane Collaboration). Change scores were used, as many of the studies were small and demonstrated marked differences at baseline between the allocation groups. Authors were contacted if these data could not be extrapolated from the published manuscript.

Sensitivity analyses were conducted according to treatment delivery mode (individual, class-based, home programs) and the number of directly supervised contact occasions. Sensitivity analyses were also conducted on critical aspects of study methodology (blinding of outcomes assessment, statistical analysis method, allocation concealment) and an overall assessment of bias risk. Differences in treatment effect between the stratifications were tested with chi-squared distribution.

## RESULTS

Characteristics of the included RCT. Of the 85 retrieved clinical trails identified from the literature searches, 32 RCT met the inclusion criteria<sup>13-45</sup>. Four of the included studies recruited people with hip or knee OA18,20,23,44. Contact with the authors provided data specific for participants with knee OA. Of the 32 included studies, 1 large study had 2 clearly different land-based exercise groups (aerobic walking and resistance training) and was treated as 2 trials, with the sample size of the control group being equally divided between the 2 exercise groups<sup>17</sup>. Two of the included studies included both a hydrotherapy and a land-based exercise allocation<sup>18,20</sup>. Only the land-based arm was evaluated in our review. Two studies allocated participants to 2<sup>21</sup> or 3<sup>24</sup> forms of muscle strengthening. The mean effect of the exercise allocations were combined and compared with the control group.

One study<sup>25</sup> had treatment allocations combining exercise with ultrasound or hyaluron, another study had 3 active treatment allocations, 2 of which included a weight reduction program<sup>30</sup>, and a third study had 4 allocations, 2 involving a spouse-assisted coping strategy intervention<sup>27</sup>. Only the allocation to exercise alone was considered in our systematic review. One study<sup>45</sup> included participants without knee pain. Data were provided by the author on the 37 participants with knee pain and confirmed knee OA.

The pain outcome measure for 1 study<sup>35</sup> was not included as all participants were required to take daily nonsteroidal antiinflammatory drugs (NSAID), a study design factor we considered would unfairly attenuate any pain-relieving benefit attributable to the exercise program. Two studies did not provide self-reported physical function as an outcome measure<sup>27,40</sup>.

Most studies recruited 50 to 150 participants. However, 11 (36%) studies recruited fewer than 25 participants in 1 or both allocation groups<sup>13,14,18,21,27,32,37-40,45</sup>, while 1 study recruited more than 750 participants<sup>41</sup>.

Participant characteristics. Sample recruitment varied widely, with studies recruiting exclusively community volunteers<sup>15,17,20,26,33,34,36</sup>, specialist rheumatology or orthopedic clinic patients<sup>18,38,39,42</sup>, a mix of community volunteers and specialist clinic patients<sup>14,27,32,35</sup>, general physician referrals<sup>22,41,44</sup>, or patients on the physiotherapy waiting list<sup>16,19</sup>. Approximately 50% of participants in 1 study reported a symptom duration of less than a year<sup>44</sup>, while other studies reported a mean symptom duration of more than 10 years<sup>29,32</sup>. Most studies stated that the American College of Rheumatology diagnosis criteria were used for study inclusion. However, "knee pain in the past week"33 and patellofemoral knee pain<sup>36</sup> were sufficient in 2 studies. Two other studies required at least Kellgren and Lawrence Grade III radiographic disease for study participation<sup>37,42</sup>. Studies ranged from those excluding people taking NSAID<sup>14</sup> to others including only people currently taking NSAID at least twice a week<sup>28</sup>. One study recruited only overweight or obese participants [body mass index (BMI)  $\ge 28 \text{ kg/m}^2$ ]<sup>30</sup>, resulting in a sample with a mean BMI of  $34 \text{ kg/m}^2$ . This range of recruitment strategies and inclusion criteria resulted in a wide variability in baseline radiographic and symptomatic disease severity between studies, when reported.

*Exercise programs.* A wide range of therapeutic exercise programs were assessed. All included studies were categorized into 1 of 3 groups according to the treatment delivery mode: individual treatments, class-based programs, or "home" programs. However, it should be noted that many "home" programs incorporated home visits by a trained nurse or community physiotherapist. Also, most individual treatments or class-based programs included provision of a home exercise program.

Treatment content varied from mostly aerobic walking programs<sup>17,28,30,32,40</sup> to very complex, comprehensive programs including manual therapy, upper limb and/or truncal muscle strengthening, and balance coordination<sup>15,16,34,37,44</sup>, in addition to the more usual lower limb muscle strengthening. Two studies evaluated tai chi classes<sup>20,39</sup>.

Apart from delivery mode and content, treatment "dosage" varied widely between studies. The total number of monitored exercise sessions provided ranged from none<sup>35,40</sup> to 36 or more<sup>14,17,27,30,32,34</sup>. Monitored treatment sessions, in either individual or class-based format, ranged from 30 minutes<sup>14,22,29,44</sup> to 90 minutes<sup>28</sup> per session. The total treatment duration ranged from 1 month<sup>16</sup> to 6 months<sup>30,33,45</sup> and 2 years<sup>41</sup>. Treatment intensity ranged from "maximum effort" muscle strengthening<sup>21,38</sup> to low intensity aerobic walking<sup>14,30,40</sup>.

*Methodological quality.* According to our criteria of methodological quality assessment (see Materials and Methods), only 9 (28%) studies could be considered low risk of bias from the published report. A further 14 (44%) were categorized as at moderate risk of bias, while the remaining 9 (28%) had a high risk of bias.

*Magnitude of treatment effect.* All included RCT (Figures 1 and 2): The 32 included RCT provided data on almost 3800 participants. Combining results demonstrated a statistically significant benefit (SMD) using a random-effects model of 0.40 (95% CI 0.30 to 0.50) for knee pain and 0.37 (95% CI 0.25 to 0.49) for self-reported physical function. Both these effect sizes would be considered small<sup>46</sup>. Between-study heterogeneity was marked:  $I^2 = 47\%$  and 62% for pain and physical function, respectively.

*Sensitivity analyses (Table 1).* Treatment delivery mode: All included RCT were stratified to exercise programs delivered individually to the patient<sup>15,16,21,24,25,29,36,38,44</sup>, provided in a class-based format<sup>14,17,18,20,26-28,32,34,37,39,42</sup>, and exercise programs mostly undertaken by the patient at home<sup>13,22,33,35,41,43</sup>. All 3 forms of treatment delivery achieved significant treatment benefits in terms of pain and

|  | Exercise |      | Control |        |       | Std. Mean Difference |                                       |
|--|----------|------|---------|--------|-------|----------------------|---------------------------------------|
| Study or Subgroup  | Mean     | SD   | Total   | Mean   | SD    | Total                | IV, Random, 95% Cl                    |
| Minor 1989   | -0.76    | 1.7  | 49      | -0.31  | 1.6   | 19                   |                                       |
| Kovar 1992   | -1.38    | 1.99 | 47      | -0.1   | 2.31  | 45                   |                                       |
| Schilke 1996   | -6.1     | 4.9  | 10      | 0.4    | 6.7   | 10                   | <  .                                  |
| Etttinger 1997b  | -0.25    | 0.72 | 146     | 0      | 0.61  | 75                   |                                       |
| Ettinger 1997a   | -0.32    | 0.6  | 144     | 0      | 0.61  | 75                   |                                       |
| Bautch 1997  | -1.4     | 2.32 | 15      | 1.03   | 1.55  | 15                   |                                       |
| van Baar 1998  | -27.4    | 28.7 | 54      | -11.7  | 28.5  | 59                   |                                       |
| Rogind 1998  | -3       | 3.9  | 11      | -0.1   | 6.7   | 12                   | · · · · · · · · · · · · · · · · · · · |
| Peloquin 1999  | -1.44    | 2    | 59      | -0.59  | 2.2   | 65                   |                                       |
| O'Reilly 1999  | -1.45    | 3.5  | 108     | -0.42  | 2.8   | 72                   |                                       |
| Maurer 1999  | -43.54   | 80.3 | 49      | -28.49 | 80.3  | 49                   |                                       |
| Deyle 2000   | -129.63  | 91   | 33      | -33.83 | 111.5 | 36                   |                                       |
| Hopman-Rock 2000   | -0.7     | 24.1 | 45      | 4      | 21.2  | . 37                 |                                       |
| Fransen 2001   | -10.6    | 19.5 | 83      | 1.5    | 19.4  | 43                   | <b>_</b>                              |
| Baker 2001   | -79      | 88   | 22      | -20    | 93    | 22                   |                                       |
| Topp 2002  | -1.53    | 3.2  | 67      | 0.02   | 3.2   | 35                   |                                       |
| Thomas 2002  | -1.27    | 3.6  | 467     | -0.46  | 3.6   | 316                  | -                                     |
| Gur 2002   | -20.9    | 8.3  | 17      | 0.7    | 4.6   | 6                    | ←                                     |
| Huang 2003   | -1.6     | 1.5  | 99      | -0.4   | 1.6   | 33                   |                                       |
| Song 2003  | -2.45    | 3.9  | 22      | 0.61   | 5.1   | 21                   | •                                     |
| Quilty 2003  | -7.93    | 27.5 | 43      | -2.59  | 22    | 44                   |                                       |
| Talbot 2003  | -0.12    | 0.99 | 17      | 0      | 1.67  | 17                   |                                       |
| Foley 2003   | -1.19    | 2.94 | 21      | -0.05  | 2.55  | 20                   |                                       |
| Messier 2004   | -0.42    | 3.5  | 80      | -1.06  | 3.4   | 78                   |                                       |
| Keefe 2004   | -0.7     | 1.69 | 16      | 0.03   | 1.27  | 18                   |                                       |
| Hughes 2004  | -1       | 3.9  | 68      | -0.3   | 3.9   | 43                   |                                       |
| Thorstensson 2005  | -1.8     | 14   | 30      | 0.3    | 15    | 31                   |                                       |
| Bennell 2005   | -2.2     | 1.7  | 73      | -2     | 2.1   | 67                   |                                       |
| Huang 2005a  | -1.2     | 1.6  | 30      | -0.5   | 1.7   | 32                   |                                       |
| Hay 2006   | -1.56    | 3.4  | 93      | -0.41  | 2.8   | 89                   |                                       |
| Mikesky 2006   | -1.6     | 5.51 | 15      | 0.36   | 3.44  | 22                   |                                       |
| Fransen 2007   | -1.67    | 3.28 | 41      | -0.5   | 2.37  | 36                   |                                       |
| Total (95% CI)   |          |      | 2074    |        |       | 1542                 |                                       |
| Heterogeneity: Tau <sup>2</sup> = 0.03; Chi <sup>2</sup> = 58.23, df = 31 (P = 0.002); $I^2 = 47\%$<br>Test for overall effect: Z = 7.67 (P < 0.00001) |          |      |         |        |       | -2 -1 0 1 2          |                                       |

*Figure 1.* All randomized controlled trials: knee pain outcome. Mean change, standard deviation, sample size for treatment and control groups.

|   | E       | cercise | Control |        |       | Std. Mean Difference |                    |  |
|---|---------|---------|---------|--------|-------|----------------------|--------------------|--|
| Study or Subgroup   | Mean    | SD      | Total   | Mean   | SD    | Total                | IV, Random, 95% Cl |  |
| Minor 1989  | -0.89   | 2.5     | 49      | 0.33   | 2.5   | 19                   |                    |  |
| Kovar 1992  | -2.4    | 2.27    | 47      | 0.24   | 2.49  | 45                   |                    |  |
| Schilke 1996  | -3.66   | 3.3     | 10      | -0.42  | 3.5   | 10                   |                    |  |
| Etttinger 1997b   | -0.16   | 0.48    | 146     | 0      | 0.48  | 75                   |                    |  |
| Bautch 1997   | -2.82   | 7.78    | 15      | -3.49  | 8.17  | 15                   |                    |  |
| Ettinger 1997a  | -0.18   | 0.48    | 144     | 0      | 0.48  | 75                   |                    |  |
| Rogind 1998   | -3      | 3.3     | 11      | -2     | 5.3   | 12                   |                    |  |
| van Baar 1998   | -1.3    | 5.7     | 54      | -0.5   | 5.6   | 59                   |                    |  |
| Peloquin 1999   | -1.5    | 2.4     | 59      | -0.54  | 2.6   | 65                   |                    |  |
| O'Reilly 1999   | -3.55   | 12.5    | 108     | -0.01  | 11.5  | 72                   |                    |  |
| Maurer 1999   | -106.9  | 390.1   | 49      | -88.3  | 390.1 | 49                   |                    |  |
| Petrella 2000   | -0.43   | 0.9     | 91      | -0.24  | 0.8   | 88                   |                    |  |
| Hopman-Rock 2000  | -0.8    | 4.6     | 37      | -1.7   | 5.2   | 34                   | - <u>+</u>         |  |
| Deyle 2000  | -402.51 | 339.56  | 33      | -98.17 | 393.9 | 36                   | ·                  |  |
| Fransen 2001  | -7.7    | 19.9    | 83      | 0.1    | 20.5  | 43                   |                    |  |
| Baker 2001  | -272    | 295     | 22      | -119   | 323   | 22                   |                    |  |
| Thomas 2002   | -2.59   | 10.5    | 467     | -0.02  | 10.5  | 316                  | -                  |  |
| Gur 2002  | -13.8   | 4.1     | 17      | 1      | 2.5   | 6                    | •                  |  |
| Topp 2002   | -4.16   | 10.9    | 67      | 0.17   | 10.9  | 35                   |                    |  |
| Foley 2003  | -2.81   | 7.89    | 21      | 2.1    | 8.1   | 20                   |                    |  |
| Song 2003   | -11.09  | 12      | 22      | -1.33  | 10.6  | 21                   |                    |  |
| Huang 2003  | -2      | 1.6     | 99      | -0.4   | 1.7   | 33                   | ·                  |  |
| Quilty 2003   | -0.86   | 7.3     | 43      | -0.27  | 7.6   | 44                   |                    |  |
| Hughes 2004   | -3.8    | 11.9    | 68      | -2.7   | 13.9  | 43                   |                    |  |
| Messier 2004  | -3.07   | 11.6    | 80      | -3.4   | 11.5  | 78                   | - <b>+</b> -       |  |
| Bennell 2005  | -7.8    | 8.7     | 73      | -8.2   | 10    | 67                   |                    |  |
| Huang 2005a   | -1.5    | 1.4     | 30      | -0.5   | 1.7   | 32                   |                    |  |
| Thorstensson 2005   | -2      | 12      | 30      | 0.6    | 18    | 31                   |                    |  |
| Mikesky 2006  | -0.2    | 11.58   | 15      | 1.93   | 9.11  | 22                   |                    |  |
| Hay 2006  | -4.79   | 10.8    | 95      | -0.8   | 8.5   | 90                   |                    |  |
| Fransen 2007  | -5.04   | 10.25   | 41      | 2.07   | 9.06  | 36                   |                    |  |
| Total (95% CI)  |         |         | 2126    |        |       | 1593                 | ♦ 1                |  |
| Heterogeneity: Tau <sup>2</sup> = 0.06; Chi <sup>2</sup> = 79.12, df = 30 (P < 0.00001); l <sup>2</sup> = 62% $\frac{1}{-2}$ -1 0 1 2 |         |         |         |        |       |                      |                    |  |
| Test for overall effect: $Z = 6.08$ (P < 0.00001)<br>Favours exercise Favours control   |         |         |         |        |       |                      |                    |  |

Figure 2. All randomized controlled trials: physical function outcome. Mean change, standard deviation, sample size for treatment and control groups.

physical function. The difference in mean effect size among these 3 categories did not reach statistical significance for knee pain or physical function (Table 1).

Supervision occasions: All included RCT were stratified according to the number of directly supervised sessions provided (in clinics or as home visits): those providing fewer than 12 occasions<sup>15,16,22,23,33,35,36,40,41,45</sup>, and the others providing 12 or more occasions. Both categories achieved significant treatment benefits in terms of pain and physical function. However, programs providing fewer than 12 direct supervision occasions demonstrated only small mean benefits for pain and for physical function. Studies evaluating programs providing 12 or more direct supervision occasions demonstrated moderate mean effect sizes for pain and physical function. The difference in mean treatment effect between the 2 categories of exercise programs was significant for both outcome measures (Table 1).

Blinding of outcomes assessment: All included studies were stratified according to whether the blinding of outcomes assessment was reported (blinded) or those RCT with uncertain blinding or where blinding of outcomes assessment was not part of the study design (unblinded). Just over half (56%) of the 32 studies clearly stated that the outcomes assessor was blinded to group allocation<sup>15-18,20,22,23,25,29,30,34-37,39,41,44,45</sup>.

Both subgroups achieved significant treatment benefits. However, blinded outcomes assessment was clearly influential on the magnitude of treatment effect. Blinded studies demonstrated a small mean treatment effect for pain and for physical function. Unblinded studies demonstrated moderate treatment effects for pain and for physical function. The difference in mean treatment effect size between these 2 subgroups was significant for both outcome measures (Table 1).

The Journal of Rheumatology 2009; 36:6; doi:10.3899/jrheum.090058

| Outcome and Subgroup   | RCT, n | Participants, n | Effect Estimate SMD<br>(95% CI) | Chi-square for<br>Distribution (p) |
|------------------------|--------|-----------------|---------------------------------|------------------------------------|
| 1. Knee pain           | 32     | 3616            | 0.40 (0.30-0.50)                |                                    |
| Delivery mode          |        |                 |                                 |                                    |
| Individual             | 10     | 849             | 0.55 (0.29-0.81)                |                                    |
| Classes                | 17     | 1608            | 0.37 (0.24-0.51)                | NS                                 |
| Home program           | 6      | 1260            | 0.28 (0.16-0.39)                |                                    |
| Direct supervision     |        |                 |                                 |                                    |
| Less than 12 occasions | 9      | 1594            | 0.28 (0.16-0.40)                |                                    |
| 12 or more occasions   | 23     | 2022            | 0.46 (0.32-0.60)                | 4.82 (0.03)                        |
| Outcomes assessment    |        |                 |                                 |                                    |
| Blinded                | 18     | 2559            | 0.33 (0.22-0.43)                |                                    |
| Unblinded/uncertain    | 14     | 1057            | 0.53 (0.33-0.73)                | 5.64 (0.02)                        |
| Analysis method        |        |                 |                                 |                                    |
| Intention to treat     | 14     | 2394            | 0.36 (0.21-0.51)                |                                    |
| Efficacy               | 18     | 1222            | 0.45 (0.32-0.58)                | NS                                 |
| Allocation concealment |        |                 |                                 |                                    |
| Adequate               | 14     | 2599            | 0.33 (0.23-0.44)                |                                    |
| Unclear/inadequate     | 18     | 1202            | 0.49 (0.32-0.66)                | 3.94 (0.05)                        |
| Risk of estimate bias  |        |                 |                                 |                                    |
| Low                    | 10     | 2021            | 0.28 (0.15-0.42)                |                                    |
| Moderate               | 13     | 972             | 0.48 (0.29-0.67)                |                                    |
| High                   | 9      | 623             | 0.51 (0.30-0.72)                | 7.98 (0.02)                        |
| 2. Physical function   | 31     | 3820            | 0.36 (0.25-0.48)                |                                    |
| Delivery mode          |        |                 |                                 |                                    |
| Individual             | 10     | 849             | 0.52 (0.19-0.86)                |                                    |
| Classes                | 16     | 1563            | 0.35 (0.19-0.50)                |                                    |
| Home program           | 6      | 1408            | 0.28 (0.17-0.38)                | NS                                 |
| Direct supervision     |        |                 |                                 |                                    |
| Less than 12 occasions | 9      | 1731            | 0.23 (0.09-0.37)                |                                    |
| 12 or more occasions   | 22     | 1988            | 0.45 (0.29–0.62)                | 5.52 (0.02)                        |
| Outcomes assessment    |        |                 |                                 |                                    |
| Blinded                | 19     | 2730            | 0.28 (0.17-0.39)                |                                    |
| Unblinded/uncertain    | 12     | 989             | 0.55 (0.28–0.83)                | 7.21 (0.01)                        |
| Analysis method        |        | 0.554           |                                 |                                    |
| Intention to treat     | 15     | 2576            | 0.30 (0.16–0.45)                |                                    |
| Efficacy               | 16     | 1143            | 0.43 (0.23–0.64)                | NS                                 |
| Allocation concealment |        |                 |                                 |                                    |
| Adequate               | 15     | 2596            | 0.28 (0.18–0.38)                |                                    |
| Unclear/inadequate     | 16     | 1123            | 0.48 (0.23–0.73)                | 4.71 (0.03)                        |
| Risk of estimate bias  |        | 2024            | 0.05 (0.12, 0.20)               |                                    |
| Low                    | 10     | 2024            | 0.25 (0.13–0.38)                |                                    |
| Moderate               | 14     | 1140            | 0.39 (0.18–0.60)                | 0.00 (0.01)                        |
| High                   | 7      | 555             | 0.55 (0.21-0.90)                | 8.98 (0.01)                        |

RCT: randomized controlled trial; SMD: standardized mean difference.

Statistical analysis method: All included studies were stratified according to the method chosen to deal with study participants without followup data or dropouts: intention-totreat (all randomized participants) or efficacy analysis (only participants with followup data or only treatment completers). Both subgroups achieved significant treatment benefits. RCT using the more rigorous intention-to-treat analysis demonstrated a small mean effect for pain and for physical function, compared with the larger effects demonstrated by studies using efficacy analysis. However, the difference between the 2 subgroups did not reach statistical significance for either outcome measure (Table 1).

Fransen and McConnell: OA knee exercise

Allocation concealment: All included RCT were stratified according to the adequacy of allocation concealment: 15 RCT reporting randomization procedures providing adequate allocation concealment<sup>13,15-20,22,30,33,35,36,39,41,44</sup> and the others not reporting sufficient details of the randomization procedure to be certain that allocation was concealed.

Both study subgroups achieved significant treatment benefits. However, studies providing adequate allocation concealment reported small mean treatment effects for pain and for physical function. Studies not reporting sufficient detail for certain, adequate allocation concealment achieved moderate mean treatment effect sizes. The difference between the 2 study subgroups was significant for both outcome measures (Table 1).

Overall estimate of bias risk: Only 9 (28%) studies could be considered low risk of bias from the published report<sup>15,17,18,20,22,30,36,41,44</sup>. All 3 study categories (low, medium, or high risk of bias) achieved significant mean treatment benefits in terms of pain and physical function. However, studies at low risk of bias demonstrated small mean treatment effects for pain and physical function. Studies at moderate or high risk of bias demonstrated mostly moderate mean treatment effects. The difference among the 3 study subgroups was significant for both outcome measures (Table 1).

## DISCUSSION

Our systematic review was restricted to RCT evaluating land-based therapeutic exercise for people with symptomatic knee OA in terms of self-reported knee pain and physical function. Overall, metaanalysis demonstrated that the evaluated exercise programs resulted in a mean treatment benefit for both knee pain (SMD 0.40, 95% CI 0.30-0.50) and physical function (SMD 0.37, 95% CI 0.25-0.49). These mean treatment benefits, extrapolated from 32 RCT recruiting almost 3800 participants, would be considered small. The magnitude of treatment benefit is, however, comparable to reported estimates for current simple analgesics and NSAID taken for knee pain<sup>47</sup>. If the metaanalysis is restricted to those 9 RCT evaluated as having a low risk of bias, land-based therapeutic exercise demonstrated smaller but still significant benefits in terms of knee pain (SMD 0.28, 95% CI 0.15–0.42) and physical function (SMD 0.25, 95% CI 0.13-0.38).

Due to marked heterogeneity between the content of evaluated exercise programs in the included RCT, sensitivity analyses could only meaningfully be conducted according to fairly crude exercise program characteristics: the mode of treatment delivery and the number of directly supervised treatment occasions. While these subgroups analyses are nonrandomized comparisons and should therefore be viewed as being exploratory, there were some interesting findings. While RCT assessing home programs demonstrated effects for pain and physical function that were consistently smaller than those using more closely supervised forms of treatment delivery (individual treatments or classbased programs), the difference between the various treatment delivery modes did not reach statistical significance. This nonsignificant finding is likely to reflect the incorporation of regular home or clinic visits by trained health professions into several of these "home" programs<sup>13,22,33,41</sup>. In fact, the magnitude of the treatment benefit demonstrated by the RCT included in this metaanalysis was significantly influenced by the number of directly supervised occasions provided within a program.

Exercise "dosage" is a factor of frequency, intensity, and

program duration and varies considerably between the studies included in this systematic review. Uncertainties in actual dosage arise due to the dependence of exercise intensity not only upon exercise prescription but also upon individual exertion. The influence of program duration upon dosage is difficult to quantify, with simple addition not providing a sufficiently physiological, plausible model. None of the included studies attempted to evaluate the influence of exercise dosage. Further, there were insufficient studies with comparable exercise program content to provide a meaningful subgroup analysis of the influence of exercise dosage on treatment effectiveness. Specific recommendations cannot, therefore, be made about optimal dosage (frequency, intensity, duration).

To reduce the risk of bias in a clinical trial, apart from adequate allocation concealment and limited loss to followup, blinding of therapists, study participants, and outcomes assessors is recommended. This approach provides the best protection that trial results will be free of selection, performance, attrition, and detection bias. Blinding of therapists and study participants is arguably impossible to achieve in studies evaluating exercise programs. Using "sham" exercise as the control intervention is fraught with ethical concerns (substantial wasted time for control participants attending an ineffective program) and is likely to be fairly transparent to the majority of people with knee OA. Therefore, a slight modification to the usual methodological criteria has been used in this systematic review. Within the blinding criteria, only blinding of outcomes assessment was required. It is of concern, therefore, that only 18 (56%) of the included RCT reported using blinded outcomes assessment; only 14 (43%) studies used an intention-to-treat analysis; and only 15 (47%) studies reported adequate allocation concealment. Not unexpectedly, the 9 studies evaluated as having a low risk of bias by fulfilling all 3 methodological quality criteria demonstrated significantly smaller mean effect sizes for pain and physical function compared with the studies evaluated as having moderate or high risk of bias (Table 1).

There are some important caveats to this review. The first concerns the responsiveness of self-reported pain and physical function. Many of the studies included in this systematic review included mostly participants with early or mild symptomatic disease. Although people with early disease frequently demonstrate reduced muscle strength and aerobic capacity compared with their age and sex peers without symptomatic OA, these physiological impairments are often not yet large enough to translate into reportable difficulties on simple questionnaires. This lack of reportable difficulties would considerably reduce the potential range of improvement possible (ceiling effect) on self-report questionnaires in people with early or mild disease. One of the potential benefits of exercise in people with early disease, such as increasing physiological reserve capacity, will not be identified by these questionnaires. Objective measures of physical performance not only strengthen the methodological quality of a study where masking to allocation is unattainable for the participant, but also potentially provide data better able to discriminate between people with early disease, where disease-related impairments have not yet developed into self-reported functional limitations or disability. Second, regular exercise provides general health benefits beyond reducing joint symptoms. This review is, therefore, likely to be underestimating the overall beneficial effect of exercise among people with knee OA.

Most people with knee OA have a pattern of chronic, fluctuating symptoms. Longterm adherence to exercise, or increased leisure-time physical activity, is required to maintain the benefits of exercise. Adherence to therapeutic exercise, however, usually requires the stimulus of regular supervision or monitoring<sup>48</sup>. Unfortunately, most individuals or healthcare systems do not have sufficient resources to allow ongoing unrestricted access to individually provided treatments for chronic musculoskeletal conditions. This systematic review could not establish a significant difference in treatment benefits, in terms of knee pain or physical function, between studies assessing individual treatments, classbased programs, or (usually closely individually monitored) home programs. It could, however, be argued that the classbased format potentially provides a cost-effective alternative that could be more regularly accessed by older people when introduced to community centers or gymnasiums; and that the social contact with peers, particularly those experiencing

#### APPENDIX. Medline search strategy.

exp osteoarthritis/
osteoarthr\$.tw.

- 3. (degenerative adj2 arthritis).tw.
- 4. arthrosis.tw.
- 5. or/1-4
- 6. Knee/
- 7. exp Knee Joint/
- 8. knee\$.tw. 9. or/6-8
- 10. exp EXERCISE/
- 11. exp exertion/
- 12. exp Physical Fitness/
- 13. exp Exercise Test/
- 14. exp Exercise Tolerance/
- 15. exp Sports/
- 16. exp PLIABILITY/
- 17. exp Physical Endurance/
- 18. exertion\$.tw.
- 19. exercis\$.tw.
- 20. sport\$.tw.
- 21. ((physical or motion) adj5 (fitness or therap\$)).tw.
- 22. (physical\$ adj2 endur\$).tw.
- 23. ((strength\$ or isometric\$ or isotonic\$ or isokinetic\$ or aerobic\$ or endurance or weight\$) adj5 (exercis\$ or train\$)).tw.
- 24. exp physical therapy modalities/
- 25. physiotherap\$.tw.

similar disease-related symptoms, is highly likely to encourage treatment adherence.

There is evidence that land-based therapeutic exercise has a benefit in terms of reduced knee pain and disability for people with knee OA. This is supported by our systematic review that includes at least 2 individual controlled trials, each with sample sizes of at least 50 per group and satisfying methodological quality criteria<sup>49</sup>. People with painful knee OA can be reassured that therapeutic exercise (individual treatments, exercise classes, or monitored home program) provided by a physiotherapist or trained health professional has the real potential to reduce, at least for the short term, their knee pain and physical disability. The magnitude of treatment benefit is likely to be associated with the number of direct supervision occasions provided.

#### REFERENCES

- Davis MA, Ettinger WH, Neuhaus JM, Mallon KP. Knee osteoarthritis and physical functioning: evidence from the NHANES 1 epidemiologic followup study. J Rheumatol 1991;18:591-8.
- Guccione AA, Felson DT, Anderson JJ, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham study. Am J Public Health 1994;84:351-8.
- Minor MA, Hewett JE, Webel RR, Dreisinger TE, Kay DR. Exercise tolerance and disease related measures in patients with rheumatoid arthritis and osteoarthritis. J Rheumatol 1988;15:905-11.
- Philbin EF, Groff GD, Ries MD, Miller TE. Cardiovascular fitness and health in patients with end-stage osteoarthritis. Arthritis Rheum 1995;38:799-805.
- 26. manipulat\$.tw. 27. kinesiotherap\$.tw. 28. exp Rehabilitation/ 29. rehab\$.tw. 30. (skate\$ or skating).tw. 31. run\$.tw. 32. jog\$.tw. 33. treadmill\$.tw. 34. swim\$.tw. 35. bicycl\$.tw. 36. (cycle\$ or cycling).tw. 37. walk\$.tw. 38. (row or rows or rowing).tw. 39. muscle strength\$.tw. 40. or/10-39 41. randomized controlled trial.pt. 42. controlled clinical trial.pt. 43. randomized.ab. 44. placebo.ab. 45. drug therapy.fs. 46. randomly.ab. 47. trial.ab. 48. groups.ab. 49. 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 50. humans.sh. 51. 49 and 50 52. and/5,9,40,51

- 5. Cooper C. Occupational activity and the risk of osteoarthritis. J Rheumatol 1995;22:10-2.
- Kujala UM, Kettunen J, Paananen H, et al. Knee osteoarthritis in former runners, soccer players, weight lifters, and shooters. Arthritis Rheum 1995;38:539-46.
- McAlindon TE, Wilson PW, Aliabadi P, Weissman B, Felson DT. Level of physical activity and the risk of radiographic and symptomatic knee osteoarthritis in the elderly: The Framingham Study. Am J Med 1999;106:151-7.
- Rangger C, Klestil T, Gloetzer W, Kemmler G, Benedetto KP. Osteoarthritis after arthroscopic partial meniscectomy. Am J Sports Med 1995;23:240-4.
- 9. Slemenda C, Brandt KD, Heilman DK, et al. Quadriceps weakness and osteoarthritis of the knee. Ann Intern Med 1997;127:97-104.
- Zhang Y, Glynn RJ, Felson D. Musculoskeletal disease research: should we analyze the joint or the person? J Rheumatol 1996;23:1130-4.
- Roddy E, Zhang W, Doherty M, et al. Evidence-based recommendations for the role of excercise in the management of the hip or knee — The MOVE consensus. Rheumatology 2005;44:67-73.
- Zhang W, Moskositz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis. Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16:137-62.
- Baker KR, Nelson ME, Felson DT, Layne JE, Sarno R, Roubenoff R. The efficacy of home-based progressive strength training in older adults with knee osteoarthritis: a randomized controlled trial. J Rheumatol 2001;28:1655-65.
- Bautch J, Malone D, Vailas A. Effects of exercise on knee joints with osteoarthritis: a pilot study of biologic markers. Arthritis Care Res 1997;10:48-55.
- Bennell KL, Hinman RS, Metcalf BR, et al. Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. Ann Rheum Dis 2005;64:906-12.
- Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. Ann Intern Med 2000;132:173-81.
- 17. Ettinger WH, 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. The Fitness Arthritis and Seniors Trial (FAST). JAMA 1997;277:25-31.
- Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis — a randomised controlled trial comparing a gym based and a hydrotherapy based strengthening programme. Ann Rheum Dis 2003;62:1162-7.
- Fransen M, Crosbie J, Edmonds J. Physical therapy is effective for patients with osteoarthritis of the knee: a randomized controlled clinical trial. J Rheumatol 2001;28:156-64.
- Fransen M, Nairn L, Winstanley J, Lam P, Edmonds J. The Physical Activity for Osteoarthritis Management (PAFORM) study. A randomised controlled clinical trial evaluating hydrotherapy and Tai Chi classes. Arthritis Rheum 2007;57:407-14.
- Gur H, Cakin N, Akova B, Okay E, Kucukoglu S. Concentric versus combined concentric-eccentric isokinetic training: effects on functional capacity and symptoms in patients with osteoarthrosis of the knee. Arch Phys Med Rehabil 2002;83:308-16.
- 22. Hay EM, Foster NE, Thomas E, et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomized trial. BMJ 2006;333:995.
- 23. Hopman-Rock M, Westhoff M. The effects of a health educational

and exercise program for older adults with osteoarthritis of the hip or knee. J Rheumatol 2000;27:1947-54.

- 24. Huang M-H, Lin Y-S, Yang R-C, Lee C-L. A comparison of various therapeutic exercises on the functional status of patients with knee osteoarthritis. Semin Arthritis Rheum 2003;32:398-406.
- Huang M-H, Yang R-C, Lee C-L, Chen T-W, Wang M-C. Preliminary results of integrated therapy for patients with knee osteoarthritis. Arthritis Rheum 2005;53:812-20.
- Hughes SL, Seymour RB, Campbell R, Pollak N, Huber G, Sharma L. Impact of the Fit and Strong intervention on older adults with osteoarthritis. The Gerontologist 2004;44:217-28.
- Keefe FJ, Blumenthal J, Baucom D, et al. Effects of spouse-assisted coping skills and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. Pain 2004;110:539-49.
- Kovar PA, Allegrante JP, MacKenzie CR, Peterson MG, Gutin B, Charlson ME. Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. Ann Intern Med 1992;116:529-34.
- Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. Arch Phys Med Rehabil 1999;80:1293-9.
- Messier SP, Loeser RF, Miller GD, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: The Arthritis, Diet and Activity Promotion Trial (ADAPT). Arthritis Rheum 2004;50:1501-10.
- Milkesky AE. Weight training does not increase strength but may slow progression in OA patients. Arthritis Rheum 2006;55:690-9.
- Minor MA, Hewett JE, Webel RR, Anderson SK, Kay DR. Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. Arthritis Rheum 1989;32:1396-405.
- O'Reilly SC, Muir KR, Doherty M. Effectiveness of home exercise on pain and disability from osteoarthritis of the knee: a randomised controlled trial. Ann Rheum Dis 1999;58:15-9.
- Peloquin L, Bravo G, Gauthier P, Lacombe G, Billiard J-S. Effects of a cross-training exercise program in persons with osteoarthritis of the knee. A randomized controlled trial. J Clin Rheumatol 1999;5:126-36.
- Petrella RJ, Bartha C. Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial. J Rheumatol 2000;27:2215-21.
- 36. Quilty B, Tucker M, Campbell R, Dieppe P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. J Rheumatol 2003;30:1311-7.
- Rogind H, Bibow-Nielsen B, Jensen B, Moller HC, Frimodt-Moller H, Bliddal H. The effects of a physical training program on patients with osteoarthritis of the knees. Arch Phys Med Rehabil 1998;79:1421-7.
- Schilke JM, Johnson GO, Housh TJ, O'Dell JR. Effects of muscle-strength training on the functional status of patients with osteoarthritis of the knee joint. Nursing Res 1996;45:68-72.
- 39. Song R, Lee E-O, Lam P, Bae S-C. Effects of Tai Chi exercise on pain, balance, muscle strength, and perceived difficulties in physical functioning in older women with osteoarthritis: a randomized clinical trial. J Rheumatol 2003;30:2039-44.
- 40. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc 2003;51:387-92.
- Thomas KS, Muir KR, Doherty M, Jones A, O'Reilly SC, Bassey EJ. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. BMJ 2002;325:752-7.
- 42. Thorstensson CA, Roos EM, Petersson IF, Ekdahl C. Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis: a randomized controlled trial. BMC Musculoskel

The Journal of Rheumatology 2009; 36:6; doi:10.3899/jrheum.090058

Disord 2005;6:27.

- 43. Topp R, Woolley S, Hornyak J, Khuder S, Kahaleh B. The effect of dynamic versus isometric resistance training on pain and functioning among adults with osteoarthritis of the knee. Arch Phys Med Rehabil 2002;83:1187-95.
- 44. van Baar ME, Dekker J, Oostendorp RA, et al. The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized clinical trial. J Rheumatol 1998;25:2432-9.
- Mikesky AE, Mazzuca SA, Brandt KD, Perkins SM, Damush T, Lane KA. Effects of strength training on the incidence and progression of knee osteoarthritis. Arthritis Rheum 2006;55:690-9.
- Cohen J. Statistical power analysis for the behavioural sciences, rev. ed. New York: New York Academic; 1977.
- Bjordal JM, Ljunggren AE, Klovning A, Slordal L. Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: meta-analysis of randomised placebo controlled trials. BMJ 2004;329:1317-20.
- Woodard CM, Berry MJ. Enhancing adherence to prescribed exercise: structured behavioral interventions in clinical exercise programs. J Cardiopulm Rehabil 2001;21:201-9.
- 49. Tugwell P, Shea B. Evidence-based rheumatology. London: BMJ Books; 2004.