A case-control study to investigate the relation between low and moderate levels of physical activity and osteoarthritis of the knee using data collected as part of the Allied Dunbar National Fitness Survey

A J Sutton, K R Muir, S Mockett, P Fentem

Abstract

Background—**Physical activity is being recommended as an intervention for seemingly almost universal improvements to health. A potential concern with this recommendation for increased exercise is that some believe increased levels of activity may lead to increased incidence of osteoarthritis of the knee (knee OA), as a result of accelerated "wear and tear" of the major joints.**

Objective—**To investigate the hypothesis that the occurrence of knee OA may be related to the duration of participation in some forms of sport and active recreation.**

Methods—**The relation between habitual exercise, reported by a cross section of people surveyed in England, and self reported knee OA was investigated. Data were derived from the Allied Dunbar National Fitness Survey (1990–91). A matched retrospective case-control design was used and a new exposure classification system which categorised different grades of activities for different time periods for each subject's lifetime participation in regular physical activities was developed. Additional data on knee injuries sustained and bodily composition were also included in a multivariate analysis.**

Results—**From 4316 people originally interviewed, 216 eligible cases (66 men, 150 women) were identified (mean age 57.1). Each case was matched to four controls. When habitual sport/exercise participation were examined during a subject's life, only exposure to regular long walks and being physically active between the ages of 20 and 24 suggested any association with developing knee OA later in life. The only strong association found was a greatly increased risk of knee OA having previously sustained a knee injury (p<0.01, odds ratio 8.0 (95% confidence interval 2.0 to 32.0)).**

Conclusions—**There was little evidence to suggest that increased levels of regular physical activity throughout life lead to an increased risk of knee OA later in life. Previous knee injury was associated with an increased risk of knee OA. Additionally, most injuries were caused through participation in physical activities. Hence,** when deciding on participation in activities, it is worth taking the likelihood of joint injury into consideration, as the chance of injury is greater in some activities than others.**

(*Ann Rheum Dis* 2001;**60**:756–764)

Exercise remains popular throughout the industrialised world. It has become clear that regular exercise can improve general health and possibly increase longevity.[1] [2] Previous studies have shown strong associations between increased physical activity and decreased risk of cardiovascular disease, assistance in weight reduction, reduction in blood pressure, and improved mood.[3] One potential conflict to this impression is that many believe that there is still truth in the "wear and tear" model of osteoarthritis (OA)—that is, an accelerated wearing out of joints owing to increased exposure to levels of high activity.[4–6] Additionally, it has been suggested that increased knee injury rates, or a combination of these, may cause a higher incidence of knee OA in people who were very active (compared with average population levels) earlier in their lives.[4] If this joint wearing model is true, and whole populations take up the recommendation for increased exercise, then there is a potential for considerable increases in joint wearing diseases as the population ages, in addition to disease caused by injuries related to exercise.

Several studies have investigated the associated risks of contracting OA in elite sportsmen and women, some of which have reported evidence suggesting that high levels of physical activity lead to an increased risk.[5–7] In comparison with this is the study carried out by White *et al*.[8] This examination of middle aged women specialist teachers of physical education who had undertaken habitual physical activity over many years demonstrated a lower prevalence of knee OA.

The effects of moderate and low levels of physical activity experienced in the general population are possibly more controversial as studies have presented somewhat inconsistent results, suggesting harmful effects only in differing subgroups of the general population.[9–11] Indeed, it cannot be ruled out that a "U" shaped relation exists, where very high and low levels of activity are detrimental

Department of Epidemiology and Public Health, University of Leicester, 22–28 Princess Road West, Leicester LE1 6TP, UK
A J Sutton

Division of Public Health Medicine and Epidemiology, University of Nottingham, University Hospital Queen's Medical Centre, Nottingham NG7 2UH, UK
K R Muir

Division of Physiotherapy Education, School of Community Health Science, Faculty of Medicine and Health Sciences, University of Nottingham, UK
S Mockett

Division of Stroke Medicine, University of Nottingham, Nottingham City Hospital (NHS) Trust, UK
P Fentem

Correspondence to:
Mr Sutton
ajs22@le.ac.uk

to the risk of OA, but moderate levels are beneficial. Hence, to assess the likely impact of realistic levels of physical exercise attainable by normal people on the risk of OA later in life, we chose to investigate the relation between self reported levels of activity throughout life and self reported OA in the community using data from the Allied Dunbar National Fitness Survey (ADNFS)[12] through a retrospective case-control study.

It has been suggested that if a protective effect of exercise on knee OA does exist, a possible explanation would be that it was because of activities strengthening quadriceps muscles, which in turn may offer increased support of the knee joint, and in this way provide some protection against knee OA.[13–18] This is, in part, the basis of the use of physiotherapy regimens for rehabilitating people with existing knee OA. For this reason exposure to activities was graded according to their potential to strengthen quadriceps muscles.

Methods

DATASET

The ADNFS was carried out in 1990–91.[12] The survey focused on the adult population (age 16 and over) of England, and subjects were recruited from all 14 health regions. The electoral register was used to select 200 addresses at random from each of 30 parliamentary constituencies (chosen from a total of 523), giving a total of 6000 addresses. One adult from each of these addresses was selected at random for inclusion in the survey. A total of 4316 subjects responded, with subjects being lost to the survey for a variety of reasons, including unoccupied address. Each respondent was interviewed at their home by specially trained interviewers, who completed a questionnaire about the respondent's health, lifestyle, and physical activity. From these volunteers, 55% were also examined at a mobile laboratory, where, among other things, data on body dimension, composition, and fitness were recorded.

The survey questionnaire was extensive and comprehensively covered details of sport and exercise participation from the age of 14 to the date of interview. The same instrument as that employed in the ADNFS was used to measure levels of habitual physical activity in the Health Survey for England[19] for four consecutive years, when the observed prevalence was consistent. No specific measure of validity was undertaken at the time of the survey in 1990. Reliable objective measures on which to base measures of validity were not available at this time. CSA accelerometers with and without heart rate monitoring and analysis provided a robust external criterion measure. However, analysis of covariance using the ADNFS data for maximum aerobic capacity (VO_2MAX) on the activity summary adjusted for age, body mass index (BMI), smoking, and social class showed significant associations for both men and women, with correlations of 0.68 and 0.63, respectively.

Data on suspected confounding factors were also available; this included a comprehensive section on knee injury. In addition, the data obtained from examination at a mobile laboratory were available for a proportion of the subjects, which provided physiological measurements on subjects at the time of interview, allowing their BMI to be calculated. A more detailed account of this fitness survey is available elsewhere.[12]

CASE SELECTION

No specific question was available which allowed direct identification of subjects who had knee OA. Instead, we defined cases using logical criteria acknowledging that knee pain is the malady,[20] and following as closely as we could the self reporting American College of Rheumatology criteria,[21] as has been done previously.[22] To be defined as a case a subject had to:

1 Respond "yes" to the question "Have you ever had arthritis?" and to reply positively when asked if they still had it.
2 Respond "yes" to the question "Do you suffer from any recurrent or continuous pain, swelling or stiffness in your knees?"
3 Report age of onset of arthritis as over 40 years.
4 Not report any suffering from pain/swelling/stiffness in both wrists and both hands/fingers.

Criteria 3 and 4 were applied to exclude patients who were likely to have inflammatory arthritis. The reported age of onset of arthritis was taken to be the diagnosis date.

CONTROL SELECTION

Controls were matched by both sex and age. It was possible to pick multiple controls for each case owing to the size of the survey; (exactly) four controls were randomly selected for each case.[23] Anyone who was "disease free" at the diagnosis age of a particular case was eligible for selection as one of their controls. In this way age matching could always be done to exactly the age of the case. This means that a case could be a control for someone who was diagnosed at a younger age than themselves. Additionally, controls were selected with replacement, which implies a subject could be used as a control for multiple cases[24]; hence subjects could potentially be included several times in the dataset, both as a case and multiple controls. Importantly, only exposure/confounder information (see below) for controls, up until the age of the case diagnosis to which they were matched, was used in the analysis.

For example, in matching controls to a case whose diagnosis date was 50 years, all people disease free at the age of 50, including those who developed knee OA after the age of 50, were eligible to be selected. If a person aged 70 was selected as a control, then they were matched at age 50 and the exposure information between the ages of 50 and 70 ignored for both case and control. It is then possible that this control was also selected as a control for another case whose diagnosis date was different, say 60 years; in this instance the exposure between the ages of 60 and 70 is ignored. It is also possible that this subject developed knee

Box 1   Exposure variables
Activity grade 1 (probably strengthen/maintain quadriceps)
Exposure 5–14 years before diagnosis
 1 Binary indicator for some regular participation
 2 Number of activities participated in regularly

Exposure 15–24 years before diagnosis
 3 Binary indicator for some regular participation
 4 Number of activities participated in regularly

Activity grade 2 (possibly strengthen/maintain quadriceps)
Exposure 5–14 years before diagnosis
 5 Binary indicator for some regular participation
 6 Number of activities participated in regularly

Exposure 15–24 years before diagnosis
 7 Binary indicator for some regular participation
 8 Number of activities participated in regularly

Activity grade 3 (generally not strengthen/maintain quadriceps)
Exposure 5–14 years before diagnosis
 9 Binary indicator for some regular participation
10 Number of activities participated in regularly

Exposure 15–24 years before diagnosis
11 Binary indicator for some regular participation
12 Number of activities participated in regularly

Long walks
Exposure 5–14 years before diagnosis
13 Binary indicator for some regular participation

Exposure 15–24 years before diagnosis
14 Binary indicator for some regular participation

Activity in young and early adulthood
Exposure in ages 14–19
15 Binary indicator for amount of exercise: a lot *v* moderate or very little or none at all
16 Binary indicator for physically active: very *v* fairly or not very or not at all

Exposure in ages 20–24
17 Binary indicator for amount of exercise: a lot *v* moderate or very little or none at all
18 Binary indicator for physically active: very *v* fairly or not very or not at all

(Note: see appendix for which exercises were classified in each activity grade)

OA, after the age of 60, say at age 67, and is included as a case, with exposure between the ages of 68 and 70 ignored. (See below for how exposure was calculated.)

EXPOSURE MEASURES
Individual data were available for each physical activity in which a volunteer reported participation for the duration of their lives starting from the age of 14, including habitual cycling and long walks. Explicitly, the questionnaire asked: "During these years from when your were 14 up to the present, have you, at any time, done any of these activities on a regular basis. Please don't include any activities you only did in school time, but do include any activities you may already have mentioned. By regular I mean at least once a week for a few months or more."

Subjects were then asked about 45 different activities (see appendix), in addition to habitual cycling and long walks (defined as walking for two miles or more, at least once a week). For each activity, the start and stopping age (if

stopped) was recorded, together with the number of years of non-regular participation (if any) within that period. From these data, summary exposure variables were calculated across all activities. (Note, habitual cycling was included in these, but long walks were considered separately.)

Each sport's potential for increasing quadriceps strength was graded on a three point scale by consensus of a panel of seven physiotherapists. A score of one indicated it would probably strengthen/maintain quadriceps muscles, a score of two indicated that it would possibly strengthen/maintain quadriceps muscles, and a score of three indicated that it generally would not strengthen maintain quadriceps muscles. A table describing the category assigned to each sport is given in the appendix (table 7).

In addition to this, to enable each person's exposure history to be summarised, two 10 year exposure "windows" were defined. These windows were defined to cover the periods 5–14 and 15–24 years before the age of diagnosis/matched age (for controls). A five year gap was left between diagnosis data and the start of exposure history to allow for error in the reported date of diagnosis, and to make sure disease onset did not influence the activities in which the subject participated. The number of activities participated in for each grade was calculated for each window. A person was deemed as exposed for a window and a given activity (and hence corresponding activity grade) if his/her history recorded participation in the activity at any time within the 10 year span of the window.

These calculations were relatively straightforward for most subjects, but when a subject reported a break from an activity for 10 years or more, it was impossible to ascertain whether this break covered the whole of the window of interest. For example, if a subject reported playing football from the age of 14 to 35, but reported not doing it regularly for 12 of those years, it was impossible to ascertain whether they did play regularly for any adjacent 10 year period that was enclosed within the ages 15 to 34. In such situations subjects were classified as exposed, but the potential impact of wrong classification was explored in a sensitivity analysis by making them unexposed.

Six variables were created reporting the number of activities participated in for each grade of activity and time window combination. A further six variables were created by dichotomising each of these six—defining them as one if at least one activity had been participated in for that window/grade and 0 otherwise. Binary variables were also defined for exposure to long walks, for each time window.

Subjects were asked two further questions relating to activity between two age bands: 14–19 and 20–24 years. These were (*i*) when you were aged (14–19 or 20–24) how much sport or exercise did you take part in? (*a*) a lot; (*b*) a moderate amount; (*c*) very little; and (*d*) none at all; and (*ii*) compared with other people of that age at that time, would you describe yourself as (*a*) very physically active; (*b*) fairly

Table 1 Proportion of cases and controls exposed to the number of graded exercises in each time period

| | Activity grade 1 (probably strengthen/maintain quadriceps) | | | | Activity grade 2 (possibly strengthen/maintain quadriceps) | | | | Activity grade 3 (generally not strengthen/maintain quadriceps) | | | |
| | Period 1-n (%) | | Period 2-n (%) | | Period 1-n (%) | | Period 2-n (%) | | Period 1-n (%) | | Period 2-n (%) | |
No of exercises	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
0	162 (75.0)	667 (77.2)	139 (64.4)	570 (66.0)	142 (65.7)	599 (69.3)	135 (62.5)	589 (68.2)	140 (64.8)	571 (66.1)	132 (61.1)	543 (62.8)
1	30 (13.9)	117 (13.5)	45 (20.8)	175 (20.3)	39 (18.1)	184 (21.3)	44 (20.4)	189 (21.9)	45 (20.8)	177 (20.5)	53 (24.5)	202 (23.4)
2	9 (4.2)	27 (3.1)	13 (6.0)	50 (5.8)	12 (5.6)	25 (2.9)	15 (6.9)	27 (3.1)	15 (6.9)	59 (6.8)	16 (7.4)	60 (6.9)
≥3	4 (1.9)	16 (1.8)	8 (3.7)	32 (3.7)	11 (5.1)	19 (2.2)	10 (4.6)	22 (2.5)	4 (1.9)	20 (2.3)	3 (1.4)	22 (2.5)
Missing	11 (5.1)	37 (4.3)	11 (5.1)	37 (4.3)	12 (5.6)	37 (4.3)	12 (5.6)	37 (4.3)	12 (5.6)	37 (4.3)	12 (5.6)	37 (4.3)

physically active; (*c*) not very physically active; or (*d*) not at all physically active. For this analysis, subjects reporting the highest category (that is, a lot or very physically active) for these two questions for each age band were classed as exposed, hence dichotomising the exposure. Box 1 provides a summary of the exposure measures described above.

CONFOUNDING VARIABLES
It was possible to ascertain whether the subject had had a knee injury before the diagnosis/pseudo-diagnosis date. The cause of the injury had also been recorded, but there were too few injuries to make an adjustment and hence obtain risk estimates at this level. For the proportion of study subjects who had been to the mobile laboratory, it was possible to calculate a BMI value at the time of interview that was used as a proxy for BMI at the time of diagnosis.

STATISTICAL ANALYSIS
A multivariate analysis was undertaken using conditional logistic regression on matched sets of size five subjects (one case and four controls). The statistical package EGRET was used for these analyses.[25] Each of the 12 activity exposure variables was examined together with the two variables on habitual walking, and four dichotomised variables indicating general levels of physical activity in youth and early adulthood (box 1). Each variable was fitted individually while adjusting for known risk factors: knee injury, and then knee injury and BMI, where available. Exposure variables were then fitted simultaneously to try to ascertain the best model. A stepwise (forwards-backwards) variable selection procedure was adopted, initially retaining terms in the model if they were significant at the 10% level.

Results
EXPLORATORY ANALYSIS
From the 4316 people originally interviewed, 365 cases were identified using the two question case definition. Of these, 103 were aged less than 40, 38 had pain in both wrists and both hands/fingers, four reported no longer having arthritis, and four had a missing response to the "still got arthritis" question. All these people were excluded leaving 216 eligible cases consisting of 66 men and 150 women. For each case exactly four controls were selected, and thus the final study group comprised 1080 subjects. Owing to the sampling with replacement, the study group consisted of 636 people included once, 143 included twice, 38 three times and 11 four times. Hence the number of distinct individuals was 828. Of these, 57 were included as both cases and controls. The mean (SD) age for cases (and hence controls) was 57.1 (11.0) years (range 40–96).

Six cases reported knee injuries (3%) before the diagnosis date, four from sport and two from other causes (not recorded). For the controls, three (0.3%) knee injuries were reported, none of which were caused through participation in sports.

Table 2 Grade 1 activity (probably strengthen/maintain quadriceps) exposure (%) for both time periods split by case/control status and sex

	Time period 1 (5–14 years previous)				Time period 2 (15–24 years previous)			
	Number (%) of cases		Number (%) of controls		Number (%) of cases		Number (%) of controls	
No of activities	Male	Female	Male	Female	Male	Female	Male	Female
0	50 (75.8)	112 (74.7)	192 (72.7)	475 (79.2)	39 (59.1)	100 (66.7)	145 (54.9)	425 (70.8)
1	8 (12.1)	22 (14.7)	47 (17.8)	70 (11.7)	14 (21.2)	31 (20.7)	72 (27.3)	103 (17.2)
2	5 (7.6)	4 (2.7)	11 (4.2)	16 (2.7)	5 (7.6)	8 (5.3)	23 (8.7)	27 (4.5)
≥3	3 (4.5)	1 (0.7)	11 (4.2)	5 (0.8)	8 (12.0)	0 (0.0)	21 (8.0)	11 (1.9)
Missing	0 (0.0)	11 (7.3)	3 (1.1)	34 (5.7)	0 (0.0)	11 (7.3)	3 (1.1)	34 (5.7)

Data on BMI were limited to the proportion of cases and controls who had the laboratory assessment—108 (50%) cases and 379 (44%) controls had data available. (Note: For the purposes of this analysis we assume values of BMI are missing completely at random.) The mean (SD) BMI for cases was 28.4 (5.0) kg/m^2 and 26.5 (3.9) kg/m^2 for controls. Because this measurement of BMI was at the date of interview and not at the diagnosis/matching date the difference between these two dates was examined. The mean (SD) time between diagnosis/pseudo-diagnosis and BMI measurement for cases was 11.1 (8.6) years and for controls 13.1 (10.1) years.

Table 1 gives a breakdown of the percentages of cases/controls exposed to each grade of activity by the number of activities done in each grade. It can be seen that around 65% of cases and controls were totally unexposed for each grade except grade 1, time period one where approximately 75% of cases and controls were unexposed. The proportions participating in the number of events is again quite similar for each grade with approximately 20% of cases/controls doing one activity, 5% doing two, and <5% doing more than three.

Table 2 displays the distribution of the number of grade 1 (probably strengthen/maintain quadriceps) activities participated in split by case/control status and sex. The exposure for all subject groupings was quite low,

with most people participating in no exercise. In no group did more than 12% of people participate in more than three activities. Table 3 shows exposure to long walks. Cases rates are 7.5% and 11.7% higher than those of controls for time periods 1 and 2 respectively. Table 4 shows the self reported levels of activity for the ages 14–19 and 20–24. More people reported doing a lot of exercise/were very physically active for the ages 14–19 than for 20–24. Generally, cases appear to report higher levels of activity in early life than controls, but the numbers of subjects not completing the questionnaire were considerable. The statistical significance of these differences is considered below.

STATISTICAL MODELLING RESULTS (USING CONDITIONAL LOGISTIC REGRESSION)

Knee injury, when fitted singularly, was highly significant (p=0.003, OR=8.0 (95% CI 2.0 to 32.0), and remained so when all other variables were added to the model. BMI (fitted as a continuous variable) was also highly significant (p=0.002 OR=1.1 (95% CI 1.0 to 1.2). Because it was not possible to calculate the BMI value for half the subjects, inclusion of this term reduced the number of matched sets from 216 to 101, with many of these remaining sets having fewer than their original four controls. For this reason modelling proceeded both including and excluding BMI.

Table 5 shows the results of including each exercise exposure variable in the model individually. Owing to the small numbers of subjects reporting doing more than one activity for any grade and time window, analysis focused on the six binary variables. None of these were statistically significant at the 10%

Table 3 Percentage of cases and controls who regularly participated in long walks over the time periods given

	Value	No (%) of cases	No (%) of controls
Regular long walks: Time period 1 (5–14 years previous)	Exposed	53 (24.5)	147 (17.0)
	Unexposed	163 (75.5)	717 (83.0)
Regular long walks: Time period 2 (15–24 years previous)	Exposed	66 (30.6)	163 (18.9)
	Unexposed	150 (69.4)	701 (81.1)

Table 4 Responses to general exercise and activity level questions in adolescent and early adult life

Age (years)	Question	Value	No (%) of cases	No (%) of controls
14–19	Amount of sport and exercise	A lot	59 (27.3)	140 (16.2)
		Other	74 (34.3)	246 (28.5)
		Missing	83 (38.4)	478 (55.3)
	Physically active compared with others	Very	49 (22.7)	135 (15.6)
		Other	81 (37.5)	251 (29.1)
		Missing	86 (39.8)	478 (55.3)
20–24	Amount of sport and exercise	A lot	39 (18.1)	80 (9.3)
		Other	94 (43.5)	307 (35.5)
		Missing	83 (38.4)	477 (55.2)
	Physically active compared with others	Very	37 (17.1)	107 (12.4)
		Other	95 (44.0)	278 (32.2)
		Missing	84 (38.9)	479 (55.4)

Note. Missing is owing to no response to questionnaire.

Table 5 Univariate analysis for binary exercise exposure variables

Binary exposure variable	Odds ratio (95% CI)	p Value
Grade 1		
Time period 1	1.1 (0.7 to 1.6)	0.69
Time period 2	1.0 (0.7 to 1.5)	0.93
Grade 2		
Time period 1	1.2 (0.8 to 1.8)	0.40
Time period 2	1.3 (0.9 to 1.8)	0.18
Grade 3		
Time period 1	1.0 (0.7 to 1.4)	0.99
Time period 2	1.1 (0.8 to 1.5)	0.77
Walking		
Time period 1	1.7 (1.1 to 2.4)	0.01
Time period 2	2.1 (1.5 to 3.0)	<0.001
Lot of exercise		
Age 14 to 19	1.5 (1.0 to 2.4)	0.07
Age 20 to 24	1.8 (1.0 to 3.0)	0.03
Physically active		
Age 14 to 19	1.2 (0.8 to 1.9)	0.39
Age 20 to 24	1.0 (0.6 to 1.6)	1.00

Table 6   Model including walking and physically active term

Variable	Logistic model coefficient (ln (OR))	Odds ratio (95% CI)	p Value
Walking time period 2	0.416	1.52 (0.95 to 2.43)	0.083
Lot of exercise age 20–24	0.469	1.60 (0.94 to 2.73)	0.085
Knee injury	1.895	6.66 (1.29 to 34.46)	0.024

level, whether or not adjustment was made for knee injury or BMI.

The variables for exposure to long walks for each time period were fitted. These were both statistically significant. For time period 1, the odds ratio (95% CI) was 1.7 (1.1 to 2.4) (p=0.01), and for time period 2 it was 2.1 (1.5 to 3.0) (p<0.001).

The general activity according to age was also examined. The difference for those who were physically active compared with others of the same age (14–19 and 20–24 years) were not significant for either group. However, the difference for those reporting doing a lot of exercise compared with controls aged 14–19 was marginally significant (p=0.07) with an increased risk of knee OA (OR=1.5 (95% CI 1.0 to 2.4) and for ages 20–24 doing lots of exercise was significant (p=0.03) with increased risk of knee OA (OR=1.8 (95% CI 1.0 to 3.0)).

Including several activity exposure terms in the model simultaneously was a problem owing to the high correlation between variables, especially between the same exposure measures over the two time periods. When model fitting using all subjects (excluding BMI but adjusting for knee injury), no two exposure terms retained their significance at the 5% level. However, with a 10% level inclusion cut off both walking in time period 2 and a lot of exercise at ages 20–24 were included (table 6).

These point estimates are consistent with those found when fitting the variables separately. When BMI was included only knee injury retained its significance at the 5% level. The point estimates remained consistent for all exposure variables previously fitted, suggesting that the lack of significance was possibly because of low power as a result of greatly reduced numbers.

This model suggests the possibility of an increased risk of OA later in life for people who are active in their early life, and for those who have participated in regular long walks in the past; though it should be emphasised that neither exercise term is significant at the 5% level.

Discussion

We chose to look at habitual physical activity and its effect on knee OA within a large random sample of people living in England. The results support the hypothesis that exercise, at the levels undertaken in the general population, does not markedly increase the risk of knee OA later in life. This finding is in agreement with previous studies on professional athletes, where associations with knee OA were observed,[5][7] because, generally, the impact loads sustained among this highly select group will be much higher than in the general population. However, a prior knee injury does seem to increase the risk of developing knee OA considerably. This result is in agreement with previous data only on knee injury.[20]

As the original ADNFS analysis reported[9] the levels of activity in the British population are quite low. There are many potential health benefits to be gained from increased regular exercise participation. This study found little evidence that exercise increases the risk of OA later in life, suggesting that the potential health benefits from exercise greatly outweigh any risk of OA incurred.

A limited number of previous studies have investigated the effect of habitual levels of exercise throughout life on knee OA in general populations. Studies based on the Framingham cohort have suggested that there is no increased risk,[9] except in the elderly.[10] Another recent study reported the converse of this, stating that high levels of physical activity may be a risk factor for (knee and hip) OA among men under the age of 50.[11] This study did have drawbacks, which brings its validity into question, including lack of information on knee injuries, shown to be important here, and exposure information limited only to walking and jogging.

An advantage of the ADNFS over previous studies in assessing the association between physical activity and knee OA is the extensive information available on recreational habitual physical activity. From this, a person's individual history of participation in physical activity can be derived for every year of their life from the age of 14 until the year of interview. Thus this provides a much more comprehensive exposure history than can be gained from the more typical "snapshot" data often collected relating to a single month/year.

To make the analysis feasible it was necessary to aggregate and classify the exposure information, owing to the large time scale covered, and the wide variety of activities in which people reported participation. The system developed here was new as previous studies in this area have not had such extensive data available.

Although the exposure information available to us is probably unique, we do, however, acknowledge a number of limitations of other aspects of the data. Firstly, the case definition used had limitations. No question was asked specifically about knee OA, but logical criteria were applied based on other information given in the questionnaire. Although it is difficult to validate the definition used, efforts were made to exclude people suspected of having inflammatory arthritis. No data were available on knee replacement surgery, and hence excluding people who said they no longer had knee OA might have excluded additional "genuine" cases, as may exclusion of people who report knee problems caused by other underlying conditions such as inflammatory back pain.

Additionally, the diagnosis date was self reported and so the potential for it to be wrongly recorded exists. Indeed, by looking at a histogram of this self diagnosis age, digit preference seemed to exist, ages divisible by five (that is, 50,

55, 60, 65, etc) being most popular. To take into account this apparent reduction in accuracy, exposure was censored for the five years before the reported diagnosis date by delaying the start of the first exposure window for duration.

The study did not include a formal assessment of the possible effects of lifetime work related physical activity. Evidence from other studies has suggested an increased risk of knee OA with high occupational physical activity.[26][27] It has also been found in some studies that high work related physical activity is associated with low participation in leisure physical activities. The closest surrogate measure we had for lifetime work related physical activity was a measure of how demanding their job was at the time of interview. We looked at this variable in a sample of our data and found some evidence to support the notion that high work related physical activity may be associated with low participation in leisure physical activities. However, such findings could certainly be prone to misclassification bias owing to possible selective changing employment status over time, and it was felt that using this proxy variable to adjust the analysis was potentially misleading. Therefore we suggest that further studies should model the affects of both leisure and occupational physical activity in prospectively collected datasets.

The grading of recreational physical activities into three categories was carried out based on the perception of the ability of each activity to exert a strengthening effect on quadriceps muscles. This was done to investigate the possibility that stronger muscles may offer improved support of the knee joint and in this way protect from OA of the knee. Our data offered no support for such a hypothesis. These gradings were made by considering each activity being performed at some arbitrary "typical" level. Without any notion of intensity for an individual subject, a dilution of exposure is inevitable. The second limitation was lack of information on the average time spent on each activity per week/month, again preventing a more detailed classification system. We are currently preparing a manuscript reporting the association between quadriceps strength and the activity grades used here. If one wished to consider grading activities on a more general criteria than strengthening quadriceps muscles, say a three group measure for physical intensity, then, we believe, the grouping of activities would change, but only slightly.

The fact that BMI was only available for approximately 45% of participants, and for those there was often a considerable lag between its measurement and disease diagnosis, limited the multivariate modelling that could be done while controlling for this risk factor. The lack of data on lifetime work related physical activity is also acknowledged. However, because few associations were found between leisure time physical activity and knee OA, the potential for these confounders to raise associations artificially is minimal. Further, it is unlikely that the effect of uncontrolled confounders exactly cancels out any differential risk between the exposed and the unexposed. However, we acknowledge that it is impossible to predict the precise effects of an unmeasured confounder and due caution is needed in interpreting the results.

The decision was made to concentrate on two exposure windows—namely, 5–14 and 15–24 years before the diagnosis/pseudo-diagnosis date. The main reason for this was to simplify the exposure classification. Combining exposures from different activities over time is problematic. Limiting attention to windows avoids the need for calculating cumulative exposures. With no information on intensity or frequency, the only available measure of the amount of exercise was the number of sports or activities participated in within each window. (A calculation based on the number of years of participation within each window was avoided for simplicity.) This is recognised as being crude and for this reason (together with the fact that few subjects did more than one exercise in a given grade) we chose to dichotomise variables, simply indicating exposed/unexposed.

The correlation between the different grades of activity and the different time windows was high, suggesting that those who do exercise of one grade at one time are more likely to do exercise at the other time period and other grades in both time periods. The distribution of people over the two time periods for each grade was so similar that analysis of just one of the time periods would have been sufficient. This suggests that people's activity routines do not change greatly over time. This correlation also makes the investigation of individual activities difficult, as one would need to control for every other activity. There was a concern that because many grade 1 activities are team sports, such as football, rugby, then the exposure of women would be low. Overall, the exposure of women was slightly lower than men, but participation was still broadly comparable.

Studies in elite athletes suggest that at most a modest increase in risk, if any, would be expected.[6] This present study was large with 80% power to detect an odds ratios of 1.5, implying that these small effects could have been detected in our analysis. Our data do not support the suggestion that increased use of the knee joint through moderate sporting and exercise participation wears out the joint and therefore increases the risk of knee OA.

Exercise in levels observed from a random sample of the English population did not increase the risk of knee OA in a predictable way using our exercise classification.

Although two exposure variables remained significant at the 10% level, these associations could be spurious, or owing to confounding, possibly because of factors such as levels of occupational physical activity. There seems to be no reason why going on regular long walks should do more harm than those activities classified as grade 1. Similarly, the possible harmful effect of reporting high levels of activity as a youth and young adult is difficult to explain. Possibly, slight injuries and insults to the knee are more serious in youth, when a person is still

physically developing. These were either too slight or too far in the past for people to recall, and the chance of these occurring is increased in people more active during this time.

The strengths of this study are its size and the detailed exercise classification questionnaire used by the ADNFS, which gave an opportunity to explore the hypothesis investigated here. There is scope for further work in this area using both pre-existing datasets and specifically designed further studies. On the basis of our results and those of others, such studies would also need to be large as the magnitude of risk is likely to be small.

The implications for knee injury, the risk of which itself is increased by sports/activity participation is, however, more serious, with an estimated sixfold increase in risk having had a previous injury after the age of 14. Most of these injuries were reported to have been caused through sports. For this reason it is important that care is taken when participating—many injuries can be prevented by warming up sufficiently or using the correct equipment.

Appendix

Table 7   List of exercises: the knee strengthening/conditioning potential grades given to each sport

Grade 1 Vigorous sports/activities Probably strengthen/condition/ maintain quadriceps muscles	Grade 2 Moderately vigorous sports/activities Possibly strengthen/condition/maintain quadriceps muscles	Grade 3 Gentler sports/activities Probably not strengthen/condition/ maintain quadriceps muscles
Squash	Swimming	Shooting
Rugby	Exercises	Golf
Hockey	Cricket	Yoga
Basketball	Keep fit	Social dancing
Boxing	Rambling	Ten pin bowling
Weight lifting	Rounders	Fishing
Athletics	Gymnastics	Motor sports
Climbing	Aerobics	Bowls
Skiing	Hiking	Snooker
Roller skating	Horse riding	Skittles
Tennis	Canoeing	
Badminton	Sailing	
Netball		
Martial arts		
Rowing		
Table tennis		
Football		
Volleyball		
Weight training		
Jogging/running		
Ice skating		

1 U.S. Department of Health and Human Services. *Physical activity and health*. A Report Of The Surgeon General, Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1996

2 Bouchard C, Shephard RJ, Stephens T. *Physical activity, fitness, and health: international proceedings and consensus statement*. Champaign, IL, US: Human Kinetics Publishers, 1994.

3 Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, *et al*. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995;273:402–7.

4 Lane NE. Exercise: a cause of osteoarthritis. J Rheumatol 1995;22(suppl 43):3–6.

5 Kujala UM, Kaprio J, Sarna S. Osteoarthritis or weight bearing joints of lower limbs in former elite male athletes. BMJ 1994;308:231–4.

6 West RR. Valuation of life in long run health care programmes. BMJ 1985;291:1139–42.

7 Spector TD, Harris PA, Hart DJ, Cicuttini FM, Nandra D, Etherington J, *et al*. Risk of osteoarthritis associated with long-term weight-bearing sports: a radiologic survey of the hips and knees in female ex-athletes and population controls. Arthritis Rheum 1996;39:988–95.

8 White JA, Wright V, Hudson AM. Relationship between habitual physical activity and osteoarthrosis in ageing women. Public Health 1993;107:459–70.

9 Hannan MT, Felson DT, Anderson JJ, Naimark A. Habitual physical activity is not associated with knee osteoarthritis: the Framingham Study. J Rheumatol 1993;20:704–9.

10 McAlindon TE, Wilson PWF, 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.

11 Cheng Y, Macera CA, Davis DR, Ainsworth BE, Troped PJ, Blair SN. Physical activity and self-reported, physician-diagnosed osteoarthritis: is physical activity a risk factor? J Clin Epidemiol 2000;53:315–22.

12 Activity and health research, Allied Dunbar National Fitness Survey: a report on activity patterns and fitness levels: main findings. London, England: Sports Council and Health Education Authority, 1992.

13 Chamberlain MA, Care G, Harfield B. Physiotherapy in osteoarthritis of the knees: a controlled trial of hospital versus home exercises. Int Rehabil Med 1982;4:101–6.

14 Fisher NM, Pendergast DR, Gresham G, Calkins E. Muscle rehabilitation: its effect on muscular and functional performance of patients with knee osteoarthritis. Arch Phys Med Rehabil 1991;72:367–74.

15 Marks R. The effect of isometric quadriceps strength training in mid-range for osteoarthritis of the knee. Arthritis Care Res 1993;6:52–6.

16 Fisher NM, Gresham G, Abrams M, Hicks J, Horrigan D, Pendergast DR. Quantitative effects of physical therapy on muscular and functional performance in subjects with osteoarthritis of the knees. Archiv Phys Med Rehabil 1993;74:840–7.

17 O'Reilly SC, Jones A, Muir KR, Doherty M. Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. Ann Rheum Dis 1998;57:588–94.

18 Ettinger WH, Burns R, Messier SP, Applegate W, Rejeski WJ, Morgan T, *et al*. A randomized trial comparing aerobic exercise with a health education program in older adults with knee osteoarthritis. The fitness arthritis and seniors trial (FAST). JAMA 1997;277:25–31.

19 Office of Population Censuses and Surveys Social Survey Division. Health survey for England, 1993 [computer file]. 2nd ed. Colchester, Essex: The Data Archive [distributor], October 1997.

20 Hadler NM. Knee pain is the malady—not osteoarthritis. Ann Intern Med 1992;116:598–600.

21 Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al.* Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.
22 O'Reilly SC, Muir KR, Doherty M. Screening for pain in knee osteoarthritis: which question? Ann Rheum Dis 1996;55:931–3.
23 Ury HK. Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data. Biometrics 1975;31:643–9.
24 Wacholder S, Silverman D, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. III. Design options. Am J Epidemiol 1991;135:1042–50.
25 Cytel Software Corporation. *Egret reference manual 1985–1990.* 1990.
26 Cooper C, Coggon D. Physical activity and knee osteoarthritis. Lancet 1999;353:2177–8.
27 Maetzel A, Mäkelä M, Hawker G, Bombardier C. Osteoarthritis of the hip and knee and mechanical occupational exposure: a systematic overview of the evidence. J Rheumatol 1997;24:1599–607.