SUPPLEMENTAL MATERIAL

Table S1

Between group differences in effect of meniscal surgery and supervised exercise therapy and education at 12 months for patients with mechanical symptoms at baseline.

	No. of patients	Mean improvement in surgery group	Mean improvement in exercise therapy group	Between group difference in mean improvement (crude)[1]	Between group difference in mean improvement (adjusted*)
	Surgery group/ Exercise group	(95% CI);	(95% CI);	(95% CI);	(95% CI);
KOOS scores ‡					
$KOOS_4$	26/28	16.9 (9.4; 24.4)	18.4 (12.5; 24.3)	-1.4 (-10.7; 7.8)	0.3 (-8.7; 9.25)
Pain	26/29	11.0 (3.1; 18.9)	14.4 (8.7; 20.0)	-3.4 (-12.7; 6.0)	1.2 (-7.3; 9.8)
Symptoms	26/29	15.2 (7.8; 22.7)	15.4 (8.5; 22.3)	-0.2 (-10.1; 9.8)	3.7 (-4.9; 12.2)
ADL	27/29	10.4 (4.1; 16.7)	12.0 (7.1; 16.8)	-1.5 (-9.3; 6.2)	1.3 (-5.7; 8.3)
Sport/Rec	27/28	22.0 (11.6; 32.4)	23.8 (15.5; 32.0)	-1.7 (-14.7; 11.2)	1.7 (-10.6; 14.1)
QOL	26/28	18.0 (9.1; 26.9)	17.6 (9.4; 25.9)	0.4 (-11.5; 12.2)	1.1 (-10.3; 12.5)
WOMET scores §					
Total scores[2]	23/24	24.7 (14.3; 35.2)	24.5 (16.7; 32.4)	0.2 (-12.5; 12.8)	4.4 (-6.9; 15.7)
Symptoms	23/24	19.1 (8.3; 29.8)	23.8 (15.2; 32.5)	-4.8 (-18.1; 8.6)	1.8 (-9.4; 13.1)
Sport/Rec/work/lifestyle	23/24	31.3 (18.0; 44.7)	23.0 (13.5; 32.4)	8.3 (-7.4; 24.1)	11.3 (-3.1; 25.7)
Emotions	23/24	32.9 (20.1; 45.6)	28.7 (18.0; 39.4)	4.2 (-12.0; 20.3)	6.3 (-8.4; 20.9)

Table 1:

All estimates are presented as mean differences with corresponding 95% confidence intervals (95% CI).

[1] Negative values denotes a higher improvement in favor of the exercise therapy group.

‡ The Knee Injury and Osteoarthritis Outcome Score (KOOS) includes subscales for pain, symptoms, function in daily living, function in sport and recreation, and quality of life, with scores ranging from 0 (worst) to 100 (best). $KOOS_4$ is the mean score of four of five of the KOOS subscale scores (i.e., pain, symptoms, function in sport and recreation, and quality of life). Improvements of 10 points or more are considered clinically relevant.

§ Western Ontario Meniscal Evaluation Tool (WOMET) includes subscales of physical symptoms, disabilities due to sports, recreation, work and lifestyle, and emotions. Results were converted to scores from 0 to100, with lower scores indicating worse quality of life.

[2] For the total score scale improvements of 15.5 points or more are considered clinically relevant.

***** Adjusted for the randomization stratification factors (center and sex) and age.

Damsted C, *et al. Br J Sports Med* 2023;0:1–8. doi: 10.1136/bjsports-2022-106207

Table S2

Patients response to the degree of mechanical symptoms on the original 0-4 point scale.

	Baseline		3 month follow-up		6 month follow-up		12 month follow-up	
	Exercise therapy	Meniscal surgery	Exercise therapy	Meniscal surgery	Exercise therapy	Meniscal surgery	Exercise therapy	Meniscal surgery
Knee symptoms during the last week: Does your knee catch or hang up when moving?								
0: Never	-	-	8	16	9	13	9	17
1: Rarely	13	9	8	6	9	2	13	7
2: Sometimes	9	19	5	3	4	1	4	1
3: Often	7	4	2	0	1	0	2	1
4: Always	1	1	0	0	0	0	1	0

Statistical analyses of the secondary outcomes

The between-group difference in change on the $KOOS_4$ and the 5 KOOS-subscales, and on the WOMET was analyzed using a linear mixed model with time (baseline, 3, 6, and 12 months), treatment arm (surgery or exercise therapy) and the interaction between treatment arm and time as fixed effects constraining the difference between the arms to 0 at baseline.[35] The model was adjusted for the randomization stratification factors (center and sex) and age. To accommodate within-person measurement dependence, a patient-specific intercept and slope were added as random effects. A 95% CI excluding differences greater than 10 KOOS-units [27] and 15.5 WOMET-units [30] between treatment arms was interpreted as no clinical meaningful difference.

To assess the assumptions for model validity, the two types of outcomes were checked as below. In case of continuous outcomes, we created scatter plots of the residuals versus time and two-dimensional scatterplots of the BLUPs (Best Linear Unbiased Prediction) of the random effects. In case of binary outcomes, only the latter was used. All scatterplots were stratified by treatment.

These plots indicated distributions compatible with the assumption of normality and did not indicate the existence of outliers. In general, model checking in this context is challenging due to the limited sample size. Hence an explicit check of the linearity of age was not performed.

Damsted C, *et al. Br J Sports Med* 2023;0:1–8. doi: 10.1136/bjsports-2022-106207