Supplementary file 1. Table. Baseline characteristics for HEAVY and LIGHT.

Variables	LIGHT n=50	HEAVY n=50
Physical activity (IPAQ short version), n (%)		
Low	28 (56)	25 (50)
Moderate	16 (32)	20 (40)
High	6 (12)	5 (10)
Education (%)		
Academic education	24 (48)	26 (52)
Vocational education	12 (24)	16 (32)
Unskilled labour/no education	14 (28)	8 (16)
Employment status (%)		
Full-time	26 (52)	28 (56)
Part-time/flex job/sick listed	6 (12)	12 (24)
Student	12 (24)	5 (10)
Unemployed/retired	6 (12)	5 (10)
Previous shoulder treatment? (yes %)	31 (62)	37 (74)
Physiotherapy exercise	26 (52)	30 (60)
Physiotherapy passive treatment	18 (36)	26 (52)
Chiropractic	8 (16)	8 (16)
Analgesic medication (prescribed)	11 (22)	8 (16)
Other treatment	15 (30)	12 (24)
steroid injection	9 (18)	9 (18)
surgery	1 (2)	5 (10)
<u>Secondary self-reported outcomes</u>		
WOSI Physical symptoms (scale 0–1000)	469.0 (182.1)	474.8 (179.6)
WOSI Sports/recreation/work (scale 0–400)	211.2 (96.1)	192.2 (95.8)
WOSI Lifestyle (scale 0–400)	185.9 (91.9)	174.6 (77.3)
WOSI Emotions (scale 0–300)	205.4 (62.9)	200.6 (51.5)
Shoulder pain past seven days (scale 0-10)		
Lowest rating	2.4 (2.2)	2.4 (1.9)
Highest rating	6.5 (2.7)	6.0 (2.2)
Average rating	4.1 (2.2)	3.9 (2.1)
Discomfort due to mechanical shoulder symptoms past seven days (scale 0-10)		
Lowest rating	2.4 (2.3)	2.5 (2.1)
Highest rating	4.9 (2.4)	4.4 (2.7)
Average rating	3.7 (2.1)	3.2 (2.1)
Patient-Specific Functional Scale (scale 0-10)	3.9 (2.1)	3.9 (1.7)
Checklist Individual Strength (scale 8-56)	37.2 (9.8)	36.9 (11.4)
COOP/WONCA (scale 6-30)	15.0 (3.5)	13.9 (3.7)
Tampa Scale of Kinesiophobia (scale 11-44)	23.4 (5.2)	22.1 (5.8)
EQ-5D-5L (scale <0-1)	0.67 (0.16)	0.72 (0.11)
EQ-VAS (scale 0-100)	58.9 (21.2)	70.4 (16.5)
<u>Secondary objective outcomes</u>		
Range of motion (˚)		
Internal rotation passive	71 (19)	68 (17)
Internal rotation active	68 (19)	65 (18)
External rotation passive	100 (23)	105 (23)
External rotation active	97 (22)	102 (21)
Isometric shoulder torque strength (Nm/kg)		
Scaption	0.45 (0.20)	0.45 (0.22)
Internal rotation	0.33 (0.15)	0.33 (0.16)
External rotation	0.25 (0.10)	0.24 (0.11)
Proprioception in flexion (error ˚)		
Low range,	5.1 (3.2)	4.4 (2.6)

Mid-range	4.1 (2.5)	3.9 (2.2)
Shoulder instability and laxity tests (positive %)		
Shoulder flexion test‡, positive = yes	33 (66)	30 (60)
Shoulder rotation test, positive >180˚	20 (40)	20 (40)
Apprehension test†, positive = yes	36 (72)	38 (76)
Relocation test*, positive = yes	27 (54)	29 (58)
Release test*, positive = yes	24 (48)	21 (42)
Load and shift anterior ‡, positive 2-3	38 (76)	44 (88)
Load and shift posterior ‡, positive 2-3	19 (38)	14 (28)
Sulcus sign, positive >1 cm	43 (86)	46 (92)
Gagey‡, positive >105˚	36 (72)	44 (88)
Rotés Queról‡, positive > 90˚	26 (52)	22 (44)

Continuous data are presented as mean or median with 95% confidence interval (CI), and categorical variables are presented as frequency % (95% CI).

*Relocation and release tests were only performed on patients with a positive apprehension test.

† Of 74 patients with a positive apprehension test, 40 patients reported pain, 20 patients reported apprehension, and 14 patients reported both pain and apprehension.

‡ Of 100 patients, 17 patients were unable to complete the shoulder flexion test, 8 the Gagey test, 6 the Rotés Queról test, 1 the load and shift anterior test, and 3 the load and shift posterior test, resulting in the test scores for these patients were interpreted as negative.

Abbreviations: CIS, Checklist Individual Strength; COOP/WONCA, Dartmouth Primary Care Cooperative Research Network/World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians; EQ-5D-5L, European Quality of life - 5 Dimensions – Five-Level; IPAQ, International Physical Activity Questionnaire; NPRS, Numeric Pain Rating Scale; VAS, Visual Analogue Scale; WOSI, Western Ontario Shoulder Instability Index.

Supplementary file 2. Table. Self-reported pain medication and other treatments received during the 16-week intervention (as observed)

	LIGHT	**HEAVY**	**Between-Group Risk Difference or Median Difference ([a]) with 95% CI (unadjusted)**
Pain medication for index shoulder			
Week 1			
Number of participants (n)	48	48	
Use of pain medication (n (%))	11 (22.9)	9 (18.8)	-4 (-20, 12)
No pills	38 (79)	39 (81)	2 (-14, 18)
< 1 pills/day	3 (6)	6 (13)	8 (-4, 20)
1-4 pills/day	4 (8)	2 (4)	-4 (-14, 5)
> 4 pills/day	3 (6)	1 (2)	-4 (-12, 4)
Type of pain medication			
Paracetamol	10 (20.8)	7 (14.6)	-6 (-21, 9)
NSAIDs	6 (12.5)	2 (4.2)	-8 (-19, 3)
Other	1 (2.1)	2 (4.2)	2 (-5, 9)
Days of pain medication use (median, 95% CI)*	4 (2.7, 7)	2 (1, 2.9)	**-2 (-3.5, -0.5)**
Week 16			
Number of participants (n)	42	41	
Use of pain medication (n (%))	5 (11.9)	4 (9.8)	-2 (-16, 11)
No pills	38 (90)	37 (90.2)	-2 (-13, 12)
< 1 pills/day	2 (5)	1 (2.4)	-2 (-10, 6)
1-4 pills/day	2 (5)	2 (5)	0 (-9, 9)
> 4 pills/day	0 (0)	1 (2.4)	2 (-2, 7)
Type of pain medication			
Paracetamol	3 (7)	4 (10)	3 (-9, 15)
NSAIDs	2 (5)	1 (2)	-2 (-10, 6)
Other	2 (5)	2 (5)	0 (-9, 9)
Days of pain medication use (median, 95% CI)*	7 (1, 7)	4 (1, 7)	-3 (-16, 10)
Other treatments received			
Number of participants	49	48	
Sought GP during the 16-week intervention period (n (%)) because of the affected shoulder	7 (14.3)	5 (10.4)	-4 (-17, 9)
Received other treatment(s) for the affected shoulder (n (%))	14 (28.6)	11 (22.9)	-6 (-23, 12)
No other treatment	35 (71.4)	37 (77.1)	6 (-12, 23)
Physiotherapy (exercises)	0 (0.0)	2 (4.2)	4 (-1, 10)
Physiotherapy (stretch, ultrasound, massage)	4 (8.2)	2 (4.2)	-4 (-14, 6)
Professional massage	5 (10.2)	5 (10.4)	0 (-12, 12)
Chiropractor	3 (6.1)	1 (2.1)	-4 (-12, 4)
Acupuncture	1 (2.0)	2 (4.2)	2 (-5, 9)
Pain medication prescribed by GP	2 (4.1)	1 (2.1)	-2 (-9, 5)
Other	8 (16.3)	3 (6.3)	-10 (-22, 2)

Abbreviations: CI, confidence interval; GP, General practitioner; NSAIDs, Non-steroidal anti-inflammatory drugs.
* Based on the proportion of patients reporting use of pain medication.
Statistically significant results (p<0.05) are marked with bold.

Supplementary file 3. Table. Outcomes at 16-week follow-up for HEAVY and LIGHT presented for sensitivity analysis using intention-to-treat baseline value varied forward (ITT BVCF) and the per protocol (PP) analysis (multiple data imputation) for patients having satisfactory adherence to the interventions.

	Total no. of assessments (LIGHT/HEAVY) *	**Mean at 16 weeks in LIGHT** (95% CI) n (ITT) = 50 n (per protocol) = 33	**Mean at 16 weeks in HEAVY** (95% CI) n (ITT) = 50 n (per protocol) = 34	**Between-Group difference at 16 weeks (crude)** (95% CI)	**Between-Group difference at 16 weeks (adjusted)** † (95% CI)
Primary outcome measure					
WOSI total (scale 0-2100)					
ITT BVCF	96/97	834.4 (722.3, 946.5)	616.1 (495.2, 737.1)	**-218.3 (-383.3, -53.3)**	**-198.7 (-335.3, -62.1)**
PP	66/68	725.1 (597.8, 852.4)	475.0 (350.6, 599.4)	**-250.0 (-424.7, -75.3)**	**-250.7 (-323.4, -178.0)**
Secondary self-reported outcomes					
WOSI Physical symptoms (scale 0–1000)					
ITT BVCF	96/97	360.2 (302.8, 417.6)	283.5 (228.7, 338.3)	-76.7 (-155.1, 1.7)	**-78.9 (-142.8, -15.0)**
PP	66/68	307.5 (239.9, 375.1)	216.5 (161.0, 272.0)	**-91.0 (-176.6, -5.4)**	**-96.8 (-140.3, -53.3)**
WOSI Sports/recreation/work (scale 0–400)					
ITT BVCF	96/97	158.8 (131.9, 185.6)	110.7 (82.4, 138.9)	**-48.1 (-86.6, -9.6)**	**-39.4 (-70.6, -8.1)**
PP	66/68	140.2 (108.9, 171.6)	83.7 (54.4, 113.0)	**-56.5 (-98.6, -14.4)**	**-50.3 (-83.0, -17.5)**
WOSI Lifestyle (scale 0–400)					
ITT BVCF	96/97	141.0 (116.4, 165.7)	97.5 (72.0, 123.0)	**-43.5 (-78.6, -8.5)**	**-36.9 (-59.7, -14.1)**
PP	66/68	118.2 (89.1, 147.3)	74.2 (47.6, 100.7)	**-44.0 (-82.7, -5.4)**	**-49.7 (-65.1, -34.3)**
WOSI Emotions (scale 0–300)					
ITT BVCF	96/97	169.3 (150.2, 188.4)	121.7 (99.6, 143.8)	-50.0 (-79.9, -20.0)	-46.0 (-67.4, -24.6)
PP	66/68	159.2 (135.4, 182.8)	100.6 (76.3, 125.0)	-58.5 (-91.8, -25.2)	-56.2 (-71.6, -40.8)
Shoulder pain last 7 days (scale 0-10)					
Lowest rating					
ITT BVCF	95/97	1.3 (0.8, 1.8)	1.1 (0.6, 1.6)	-0.3 (-1.1, 0.4)	-0.4 (-1.1, 0.3)
PP	66/68	0.8 (0.4, 1.3)	0.4 (0.2, 0.7)	-0.4 (-0.9, 0.1)	-0.5 (-1.0, 0.0)
Highest rating					
ITT BVCF	95/97	4.0 (3.3, 4.8)	2.8 (2.1, 3.5)	**-1.4 (-2.4, -0.3)**	**-1.1 (-2.0, -0.2)**
PP	66/68	3.2 (2.5, 3.9)	1.8 (1.2, 2.4)	**-1.4 (-2.2, -0.5)**	**-1.2 (-1.8, -0.6)**
Average rating					
ITT BVCF	95/97	2.3 (1.7, 2.8)	1.7 (1.1, 2.3)	-0.7 (-1.5, 0.2)	-0.6 (-1.6, 0.3)
PP	66/68	1.6 (1.1, 2.1)	0.9 (0.5, 1.2)	-0.8 (-1.4, -0.1)	-0.8 (-1.2, -0.3)

Discomfort due to mechanical symptoms last 7 days (scale 0-10)					
Lowest rating					
ITT BVCF	95/97	1.3 (0.9, 1.6)	1.1 (0.7, 1.6)	-0.2 (-0.8, 0.5)	-0.2 (-0.9, 0.4)
PP	66/68	1.1 (0.7, 1.6)	0.6 (0.3, 1.0)	-0.5 (-1.1, 0.1)	-0.5 (-1.0, 0.0)
Highest rating					
ITT BVCF	95/97	3.0 (2.4, 3.6)	2.2 (1.6, 2.8)	-0.9 (-1.8, 0.0)	**-0.7 (-1.3, -0.05)**
PP	66/68	2.6 (1.9, 3.3)	1.6 (1.0, 2.3)	**-1.0 (-1.9, -0.002)**	**-0.8 (-1.6, -0.02)**
Average rating					
ITT BVCF	95/97	1.9 (1.5, 2.4)	1.5 (1.0, 2.1)	-0.6 (-1.3, 0.2)	-0.4 (-1.0, 0.2)
PP	66/68	1.6 (1.0, 2.2)	0.9 (0.4, 1.4)	-0.7 (-1.4, 0.1)	-0.5 (-1.3, 0.2)
Patient-Specific Functional Scale (scale 0-10)					
ITT BVCF	95/97	5.4 (4.7, 6.1)	5.7 (5.0, 6.5)	0.3 (-0.8, 1.3)	0.3 (-0.7, 1.4)
PP	66/68	6.1 (5.4, 6.9)	6.0 (5.1, 6.9)	-0.1 (-1.3, 1.0)	0.2 (-0.8, 1.2)
Checklist Individual Strength (scale 8-56)					
ITT BVCF	95/97	33.1 (29.6, 36.6)	30.1 (26.8, 33.4)	-3.0 (-7.8, 1.8)	-2.8 (-6.4, 0.8)
PP	66/68	29.2 (25.0, 33.5)	26.4 (23.3, 29.5)	-2.8 (-8.0, 2.4)	-1.9 (-6.3, 2.5)
COOP/WONCA (scale 6-30)					
ITT BVCF	94/97	14.3 (13.1, 15.5)	12.9 (11.7, 14.1)	-1.4 (-3.0, 0.3)	-0.6 (-1.7, 0.4)
PP	66/68	13.1 (11.7, 14.5)	11.4 (10.5, 12.4)	-1.6 (-3.3, 0.04)	-1.1 (-2.5, 0.4)
Tampa Scale of Kinesiophobia, (scale 11-44)					
ITT BVCF	94/97	22.4 (20.7, 24.1)	20.4 (18.8, 22.0)	-2.1 (-4.4, 0.3)	**-1.1 (-2.1, -0.02)**
PP	66/68	21.5 (19.6, 23.4)	19.6 (17.9, 21.3)	-1.9 (-4.4, 0.6)	**-1.6 (-2.4, -0.9)**
EQ-5D-5L, (scale <0-1)					
ITT BVCF	94/97	0.74 (0.70, 0.78)	0.79 (0.76, 0.83)	0.05 (-0.004, 0.1)	0.03 (-0.008, 0.06)
PP	66/68	0.78 (0.74, 0.81)	0.83 (0.80, 0.87)	**0.06 (0.02, 0.10)**	**0.06 (0.01, 0.10)**
EQ-VAS (scale 0-100)					
ITT BVCF	94/97	68.8 (63.7, 73.9)	74.4 (69.7, 79.2)	5.6 (-1.4, 12.6)	-0.4 (-7.2, 6.3)
PP	66/68	73.4 (68.2, 78.6)	80.6 (76.9, 84.4)	**7.2 (0.8, 13.6)**	2.4 (-4.2, 8.9)
<u>Secondary objective outcomes</u>					
Range of motion (˚)					
Internal rotation passive					
ITT BVCF	87/90	71.5 (66.6, 76.4)	69.6 (65.1, 74.1)	-1.9 (-8.5, 4.8)	0.4 (-2.1, 3.0)
PP	63/66	75.9 (70.0, 81.8)	68.1 (62.5, 73.8)	-7.7 (-15.6, 0.1)	-3.0 (-6.6, 0.5)
Internal rotation active					
ITT BVCF	87/90	67.9 (63.4, 72.5)	70.2 (66.2, 74.2)	2.3 (-3.8, 8.4)	**4.2 (0.4, 8.0)**
PP	63/66	71.0 (65.6, 76.3)	69.7 (64.9, 74.4)	-1.3 (-8.2, 5.6)	2.7 (-4.0, 9.3)
External rotation passive					
ITT BVCF	87/90	103.7 (96.2, 111.2)	107.2 (100.8, 113.5)	3.5 (-6.3, 13.3)	0.4 (-9.5, 10.3)
PP	63/66	110.2 (104.0, 116.3)	108.3 (100.9, 115.7)	-1.5 (-13.2, 10.2)	-1.8 (-10.8, 7.1)

External rotation active					
ITT BVCF	87/90	99.2 (92.7, 105.7)	106.4 (100.9, 111.9)	7.2 (-1.4, 15.7)	3.9 (-3.2, 10.9)
PP	63/66	104.9 (99.8, 110.1)	106.7 (100.0, 113.5)	1.8 (-6.0, 9.6)	1.8 (-8.1, 11.6)
Isometric shoulder torque strength (Nm/kg)					
Scaption					
ITT BVCF	87/90	0.46 (0.41, 0.52)	0.51 (0.44, 0.57)	0.04 (-0.04, 1.3)	**0.04 (0.02, 0.07)**
PP	63/66	0.51 (0.44, 0.59)	0.56 (0.47, 0.64)	0.04 (-0.07, 0.16)	0.07 (-0.002, 0.13)
Internal rotation					
ITT BVCF	87/90	0.36 (0.31, 0.40)	0.35 (0.30, 0.40)	-0.01 (-0.07, 0.06)	0.002 (-0.02, 0.02)
PP	63/66	0.39 (0.33, 0.45)	0.38 (0.32, 0.45)	-0.01 (-0.10, 0.08)	0.01 (-0.03, 0.04)
External rotation					
ITT BVCF	87/90	0.25 (0.22, 0.27)	0.26 (0.23, 0.30)	0.02 (-0.03, 0.06)	0.02 (-0.01, 0.06)
PP	63/66	0.26 (0.23, 0.29)	0.29 (0.24, 0.34)	0.03 (-0.03, 0.08)	0.03 (-0.02, 0.08)
Proprioception in flexion (error ˚)					
Low range					
ITT BVCF	87/90	4.6 (3.8, 5.4)	4.8 (3.9, 5.7)	0.2 (-0.9, 1.4)	0.7 (-0.4, 1.7)
PP	63/66	4.5 (3.5, 5.5)	5.1 (3.8, 6.5)	0.6 (-1.0, 2.2)	1.1 (-0.5, 2.7)
Mid-range					
ITT BVCF	86/90	3.6 (3.0, 4.1)	4.2 (3.3, 5.1)	0.6 (-0.4, 1.7)	0.7 (-0.5, 1.8)
PP	63/66	3.2 (2.4, 3.9)	4.4 (3.1, 5.6)	1.2 (-0.2, 2.6)	1.1 (-1.0, 3.3)
Shoulder instability and laxity tests (positive %) §					
Shoulder flexion test, positive = yes					
ITT BVCF	87/90	72 (58, 83)	62 (48, 74)	0.60 (0.12, 3.09)	0.63 (0.27, 1.47)
PP	63/66	79 (65, 93)	61 (44, 78)	0.42 (0.14, 1.26)	0.38 (0.06, 2.31)
Shoulder rotation test, positive >180˚					
ITT BVCF	87/90	54 (40, 67)	42 (29, 56)	0.62 (0.28, 1.36)	0.45 (0.16, 1.24)
PP	63/66	62 (45, 79)	40 (23, 57)	0.41 (0.15, 1.15)	**0.35 (0.14, 0.90)**
Apprehension test, positive = yes					
ITT BVCF	87/90	72 (58, 83)	64 (50, 76)	0.69 (0.30, 1.61)	0.53 (0.27, 1.02)
PP	63/66	64 (47, 80)	65 (49, 82)	1.08 (0.39, 2.95)	0.75 (0.38, 1.48)
Relocation test ‖, positive = yes					
ITT BVCF	87/90	56 (42, 69)	44 (31, 58)	0.62 (0.28, 1.36)	0.49 (0.32, 0.77)
PP	63/66	52 (34, 70)	48 (30, 65)	0.84 (0.31, 2.22)	0.71 (0.41, 1.25)

Release test ‖, positive = yes					
ITT BVCF	87/90	50 (36, 64)	34 (22, 48)	0.52 (0.23, 1.15)	0.49 (0.20, 1.19)
PP	63/66	45 (27, 63)	42 (25, 59)	0.86 (0.32, 2.32)	0.79 (0.29, 2.13)
Load and shift anterior, positive 2-3					
ITT BVCF	87/90	68 (54, 79)	68 (54, 79)	1 (0.43, 2.32)	0.65 (0.29, 1.45)
PP	63/66	68 (51, 85)	65 (49, 82)	0.89 (0.31, 2.53)	0.60 (0.22, 1.64)
Load and shift posterior, positive 2-3					
ITT BVCF	87/90	30 (19, 44)	20 (11, 34)	0.58 (0.23, 1.46)	0.67 (0.22, 2.03)
PP	63/66	30 (14, 46)	15 (3, 28)	0.43 (0.13, 1.44)	0.46 (0.10, 2.13)
Sulcus sign, positive >1 cm					
ITT BVCF	87/90	84 (71, 92)	86 (73, 93)	1.17 (0.39, 3.51)	1.11 (0.44, 2.79)
PP	63/66	85 (72, 97)	85 (72, 97)	0.99 (0.26, 3.83)	1.08 (0.30, 4.00)
Gagey, positive >105˚					
ITT BVCF	87/90	88 (76, 95)	92 (80, 97)	1.57 (0.41, 5.93)	0.70 (0.39, 1.28)
PP	63/66	93 (84, 103)	91 (80, 101)	0.68 (0.10, 4.57)	0.36 (0.10, 1.36)
Rotés Queról, positive > 90˚					
ITT BVCF	87/90	60 (46, 73)	52 (38, 66)	0.72 (0.33, 1.60)	0.73 (0.27, 1.96)
PP	63/66	52 (34, 70)	51 (33, 68)	0.94 (0.35, 2.50)	0.62 (0.18, 2.10)
Global Perceived Effect ‡ §					
(% rated important effect postintervention)					
Physical symptoms					
PP	33/34	54.5 (37.4, 70.7)	79.4 (62.3, 90.0)	**OR 3.2 (1.1, 9.4)**	**OR 3.6 (1.7, 7.6)**
Sports/recreation/work					
PP	33/34	48.5 (32.0, 65.4)	64.7 (47.2, 79.0)	OR 1.9 (0.7, 5.2)	**OR 2.0 (1.0, 4.1)**
Lifestyle					
PP	33/34	51.5 (34.6, 68.0)	64.7 (47.2, 79.0)	OR 1.7 (0.6, 4.6)	OR 1.8 (0.7, 4.7)
Emotions					
PP	33/34	45.5 (29.3, 62.6)	64.7 (47.2, 79.0)	OR 2.2 (0.8, 5.9)	OR 2.2 (0.7, 6.3)

* For BVCF, there were 100 possible assessments for each group (50 at baseline and 50 at 16 weeks follow-up), except for Global Perceived Effect, which had 50 possible assessments for each group. For per protocol, there were 66 and 68 possible assessments for LIGHT and HEAVY, respectively, except for Global Perceived Effect which had 33 (LIGHT) and 34 (HEAVY) possible assessments.

† The results are adjusted for baseline score, age, sex, and the clustering of physiotherapy clinic

‡ No data imputation

§ Proportions of a positive test in % (95% CI) and odds ratio (OR) for between-group differences with group LIGHT as reference.

‖ Relocation and release tests were only performed on patients with a positive apprehension test.

Abbreviations: CI, Confidence Interval; CIS, Checklist Individual Strength; COOP/WONCA, Dartmouth Primary Care Cooperative Research Network/World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians; EQ-5D-5L, European Quality of Life - Five Dimensions – Three Level; NPRS, Numeric Pain Rating Scale; VAS, Visual Analogue Scale; WOSI, Western Ontario Shoulder Instability Index.

Statistically significant results (p<0.05) are marked with bold.

Supplementary file 4. Table. Outcomes at 16-week follow-up for HEAVY and LIGHT.

	Total no. of assessments (LIGHT/HEAVY) *	**Mean at 16 weeks in LIGHT** (95% CI) n = 50	**Mean at 16 weeks in HEAVY** (95% CI) n = 50	**Between-Group difference at 16 weeks (crude risk ratio)** (95% CI)	**Between-Group difference at 16 weeks (adjusted risk ratio) †** (95% CI)
Shoulder instability and laxity tests (positive %) §					
Shoulder flexion test, positive = yes	87/90	78 (64, 91)	62 (47, 76)	0.79 (0.49, 1.29)	0.84 (0.62, 1.14)
Shoulder rotation test, positive >180˚	87/90	62 (47, 76)	42 (28, 56)	0.68 (0.38, 1.19)	**0.68 (0.49, 0.95)**
Apprehension test, positive = yes	87/90	70 (55, 85)	62 (48, 76)	0.89 (0.55, 1.44)	0.86 (0.69, 1.06)
Relocation test‖, positive = yes	87/90	55 (38, 72)	44 (30, 58)	0.81 (0.45, 1.45)	0.78 (0.60, 1.01)
Release test‖, positive = yes	87/90	50 (32, 68)	37 (23, 51)	0.74 (0.40, 1.38)	0.74 (0.48, 1.16)
Load and shift anterior, positive 2-3	87/90	68 (52, 84)	62 (47, 77)	0.91 (0.55, 1.51)	0.84 (0.66, 1.06)
Load and shift posterior, positive 2-3	87/90	28 (13, 44)	18 (7, 29)	0.65 (0.27, 1.57)	0.73 (0.32, 1.67)
Sulcus sign, positive >1 cm	87/90	84 (68, 93)	85 (70, 93)	1.00 (0.64, 1.55)	1.00 (0.82, 1.23)
Gagey, positive >105˚	87/90	92 (85, 100)	90 (78, 100)	0.97 (0.64, 1.47)	0.95 (0.84, 1.07)
Rotés Queról, positive > 90˚	87/90	63 (48, 77)	55 (41, 69)	0.88 (0.52, 1.47)	1.00 (0.78, 1.28)
Global Perceived Effect ‡ §					
(% rated important effect postintervention)					
Physical symptoms	45/47	44 (31, 59)	64 (49, 76)	1.44 (0.82, 2.53)	**1.46 (1.06, 2.01)**
Sports/recreation/work	45/47	38 (25, 53)	51 (37, 65)	1.35 (0.73, 2.52)	1.36 (0.90, 2.06)
Lifestyle	45/47	44 (31, 59)	55 (41, 69)	1.24 (0.69, 2.23)	1.26 (0.78, 2.03)
Emotions	45/47	40 (27, 55)	51 (37, 65)	1.28 (0.69, 2.35)	1.27 (0.70, 2.30)

* There were 100 possible assessments for each group (50 at baseline and 50 at 16 weeks follow-up), except for Global Perceived Effect which had 50 possible assessments for each group.

† The results are adjusted for baseline score of the variable of interest, age, sex, and the clustering around clinic.

‡ No data imputation

§ Proportions of positive test in % (95% CI) and risk ratio for between-group differences with group LIGHT as reference.

‖ Relocation and release tests were only performed on patients with a positive apprehension test.

Abbreviations: CI, Confidence Interval; CIS, Checklist Individual Strength; COOP/WONCA, Dartmouth Primary Care Cooperative Research Network/World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians; EQ-5D-5L, European Quality of Life - Five Dimensions – Three Level; NPRS, Numeric Pain Rating Scale; VAS, Visual Analogue Scale; WOSI, Western Ontario Shoulder Instability Index.

Statistically significant results ($p<0.05$) are marked with bold.

Supplementary file 5. Table. Adverse Events (specific, serious or minor, and withdrawals due to adverse events), and crude difference between risks and medians were calculated with 95% Confidence Intervals based on the "per protocol' data while still respecting the original group allocation, from baseline to 16-week follow-up for HEAVY vs LIGHT in patients with hypermobility spectrum disorder and shoulder complaints

Adverse events	**LIGHT (n = 33)**	**HEAVY (n = 34)**	**Between-Group Risk Difference or Median difference with 95% CI (unadjusted)**
Number of patients reporting serious adverse events*	0 (0)	0 (0)	0 (0, 0)
Number of patients reporting minor adverse events (n (%))	20 (61)	21 (62)	1 (-22, 24)
Index shoulder			
- Muscle soreness	15 (45)	17 (50)	5 (-19, 28)
- Shoulder is locked	2 (6)	1 (3)	-3 (-13, 7)
- Subluxation	3 (9)	1 (3)	-6 (-17, 5)
- Dislocation	0 (0)	1 (3)	3 (-3, 9)
- Persistent worsening of symptoms	6 (18)	5 (15)	-3 (-21, 14)
Other sites than index shoulder			
- Headache	6 (18)	11 (32)	14 (-6, 35)
"Other" minor events related to index shoulder or other sites	14 (42)	13 (38)	-4 (-28, 19)
Number of minor adverse events (median, (95% CI) †	1 (0, 4.3)	2 (0, 3.2)	1 (-0.6, 2.6)
Index shoulder			
- Muscle soreness	0 (0, 1)	0.5 (0, 2)	0.5 (-0.04, 1.04)
- Shoulder is locked	0 (0, 0)	0 (0, 0)	0 (0, 0)
- Subluxation	0 (0, 0)	0 (0, 0)	0 (0, 0)
- Dislocation	0 (0, 0)	0 (0, 0)	0 (0, 0)
- Persistent worsening of symptoms	0 (0, 0)	0 (0, 0)	0 (0, 0)
Other sites than index shoulder			
- Headache	0 (0, 0)	0 (0, 0.2)	0 (0, 0)
"Other" minor events related to index shoulder or other sites	0 (0, 1)	0 (0, 1)	0 (0, 0)

This table includes all adverse events that occurred during the 16-week study period, but which did not necessarily have a causal relationship with the treatment administered. *Serious adverse events are unexpected but cover death, life-threatening events, disability, and permanent damage.

† For each patient, each adverse event could count 0 to 16 times corresponding with the 16-week intervention period.

Supplementary file 6. Table. The adjusted (age, sex, clustering around clinic) risk difference using margins after fitting a logistic regression model for adverse events with 95% Confidence Intervals based on the 'as observed' data while still respecting the original group allocation, from baseline to 16-week follow-up for HEAVY vs LIGHT in patients with hypermobility spectrum disorder

Adverse events	**LIGHT (n = 46)**	**HEAVY (n = 45)**	**Model-based standardisation:** Between-group risk differences adjusted for age, sex, and cluster by clinics
Number of patients reporting minor adverse events (n (%))	24 (52)	29 (64)	12 (-6, 29)
Index shoulder			
- Muscle soreness	17 (37)	25 (56)	18 (-5, 42)
- Shoulder is locked	3 (4)	2 (4)	-2 (-10, 5)
- Subluxation	3 (7)	1 (2)	-5 (-10, 0)
- Dislocation	0 (0)	1 (2)	Not possible due to no events in the LIGHT group
- Persistent worsening of symptoms	8 (17)	8 (18)	0 (-20, 21)
Other sites than index shoulder			
- Headache	9 (20)	18 (40)	20 (-2, 42)
"Other" minor events related to index shoulder or other sites	18 (39)	19 (42)	2 (-27, 31)

This table includes all adverse events that occurred during the 16-week study period but did not necessarily have a causal relationship with the treatment administered.