BJO

■ **SHOULDER & ELBOW**

Effectiveness of interventions for the management of primary frozen shoulder

A SYSTEMATIC REVIEW OF RANDOMIZED TRIALS

S. S. Rex,
L. Kottam,
C. McDaid,
S. Brealey,
J. Dias,
C. E. Hewitt,
A. Keding,
S. E. Lamb,
K. Wright,
A. Rangan

From York Trials Unit, Department of Health Sciences, University of York, York, UK

Aims

This systematic review places a recently completed multicentre randomized controlled trial (RCT), UK FROST, in the context of existing randomized evidence for the management of primary frozen shoulder. UK FROST compared the effectiveness of pre-specified physiotherapy techniques with a steroid injection (PTSI), manipulation under anaesthesia (MUA) with a steroid injection, and arthroscopic capsular release (ACR). This review updates a 2012 review focusing on the effectiveness of MUA, ACR, hydrodilatation, and PTSI.

Methods

MEDLINE, Embase, PEDro, Science Citation Index, Clinicaltrials.gov, CENTRAL, and the World Health Organization (WHO) International Clinical Trials Registry were searched up to December 2018. Reference lists of included studies were screened. No language restrictions applied. Eligible studies were RCTs comparing the effectiveness of MUA, ACR, PTSI, and hydrodilatation against each other, or supportive care or no treatment, for the management of primary frozen shoulder.

Results

Nine RCTs were included. The primary outcome of patient-reported shoulder function at long-term follow-up (> 6 months and ≤ 12 months) was reported for five treatment comparisons across four studies. Standardized mean differences (SMD) were: ACR versus MUA: 0.21 (95% confidence interval (CI) 0.00 to 0.42), ACR versus supportive care: -0.13 (95% CI -1.10 to 0.83), and ACR versus PTSI: 0.33 (95% CI 0.07 to 0.59) and 0.25 (95% CI -0.34 to 0.85), all favouring ACR; MUA versus supportive care: 0 (95% CI -0.44 to 0.44) not favouring either; and MUA versus PTSI: 0.12 (95% CI -0.14 to 0.37) favouring MUA. None of these differences met the threshold of clinical significance agreed for the UK FROST and most confidence intervals included zero.

Conclusion

The findings from a recent multicentre RCT provided the strongest evidence that, when compared with each other, neither PTSI, MUA, nor ACR are clinically superior. Evidence from smaller RCTs did not change this conclusion. The effectiveness of hydrodilatation based on four RCTs was inconclusive and there remains an evidence gap.

Correspondence should be sent to Dr Lucksy Kottam; email: lucksy.kottam@nhs.net

Introduction

Primary frozen shoulder, or adhesive capsulitis, is a painful stiffness of the shoulder (glenohumeral) joint that commonly affects people in the sixth decade of life.[1,2] The exact cause remains unknown. Multip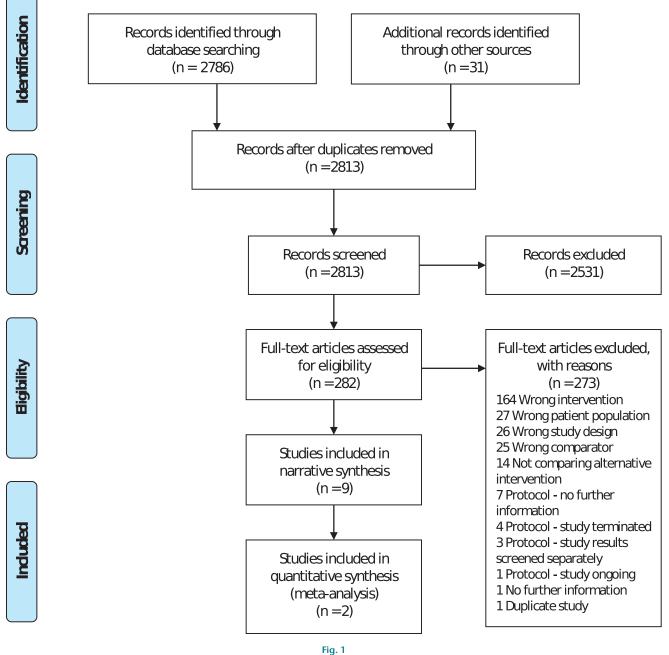le treatment choices are available with arthroscopic capsular release (ACR) being the most invasive followed by manipulation under anaesthesia (MUA), hydrodilatation, physiotherapy techniques with steroid injection (PTSI), and standard supportive care being the least invasive. Treatment is often offered from least to most invasive.[3]

Identification

| Records identified through database searching (n = 2786) | Additional records identified through other sources (n = 31) |

Screening

Records after duplicates removed (n = 2813)

Records screened (n = 2813) → Records excluded (n = 2531)

Eligibility

Full-text articles assessed for eligibility (n = 282) → Full-text articles excluded, with reasons (n = 273)
164 Wrong intervention
27 Wrong patient population
26 Wrong study design
25 Wrong comparator
14 Not comparing alternative intervention
7 Protocol - no further information
4 Protocol - study terminated
3 Protocol - study results screened separately
1 Protocol - study ongoing
1 No further information
1 Duplicate study

Included

Studies included in narrative synthesis (n = 9)

Studies included in quantitative synthesis (meta-analysis) (n = 2)

Fig. 1
Flow diagram.

There have been multiple systematic reviews on treatments for frozen shoulder specifically, or shoulder pain in general.[4-8] In 2012, a National Institute for Health Research Health Technology Assessment (NIHR HTA) programme-funded systematic review undertook a broad review of noninvasive and invasive treatments for frozen shoulder.[9] The authors concluded that there was limited clinical evidence on the effectiveness of different treatment options in the management of a primary frozen shoulder, including hydrodilatation, MUA, and ACR.[9] The review findings informed the design and funding of the multicentre randomized, parallel group, superiority trial called UK FROST.[10,11]

UK FROST compared the clinical and cost-effectiveness of PTSI, MUA, and ACR in secondary care for patients with primary frozen shoulder. Both MUA and ACR were followed with a programme of prespecified physiotherapy techniques.[12] As this was a large trial of three key treatment options for frozen shoulder it was timely to update the 2012 systematic review in the context of current randomized evidence for these specific interventions. The previous review included comparison of different

Table I. Summary of study characteristics.

Description	Country	Number randomized	Interventions	Number of sites	Number dropped out, n (%)	Mean age, yrs (SD)	Females, %	Included diabetic patients?	Diabetic patients, n (%)	Mean duration of symptoms, mths
UK FROST Unpublished data, since published)[11]	UK	503	1. MUA 2. ACR 3. Steroid and physio	35	57 (11.33)	54.3 (7.7)	63.42	Yes	150 (29.82)	10.9
De Carli[24]	Italy	46	1. ACR 2. Steroid and physio	1	2 (4.35)	55.6	54.35	Yes	6 (13.04)	N/A
Gallacher et al[29]	UK	50	1. ACR 2. Hydrodilatation	1	11 (22)	53.9 (9)	70.00	Yes	8 (16)	N/A
Jacobs[26]	UK	53	1. MUA 2. Hydrodilatation	1	10 (18.87)	56.75	66.04	No	0	4.4
Kivimaki[28]	Finland	125	1. MUA 2. Supportive care	3	46 (36.8)	53 (8.5)	68.00	Yes	18 (14.4)	7.2
Mukherjee et al[25]	India	60	1. ACR 2. Steroid and physio	1	4 (6)	50.4 (8.8)	58.93	Yes	16 (28.57)	6.3
Mun and Baek[31]	Korea	136	1. Steroid and physio 2. Hydrodilatation	1	15 (11.03)	53.01 (6.15)	62.81	Unknown	N/A	6.5
Quraishi[27]	UK	36	1. MUA 2. Hydrodilatation	1	3 (8.33)	54.87	58.33	Yes	6 (16.67)	8.9
Smitherman et al[30]	USA	26	1. ACR 2. Supportive care	1	9 (34.62)	51.75 (9.2)	N/A	Yes	N/A	N/A

ACR, arthroscopic capsular release; MUA, manipulation under anaesthesia; N/A, not available; SD, standard deviation.

physiotherapy techniques. This is outside the remit of the current review and has also been the subject of several other reviews.[3,4,13] In a survey of UK health professionals published in 2012, only 5% reported using hydrodilatation in the treatment of primary frozen shoulder.[14] Since then, there has been anecdotal evidence that this treatment option has increased in popularity and so was considered important to include in this review. Therefore, the aim of the review was to assess the effectiveness of MUA, ACR, hydrodilatation, and PTSI in the management of patients with a primary frozen shoulder.

Methods

The protocol for this review and the findings are reported in alignment with the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P) checklist and the PRISMA guidance, respectively.[15,16]

Protocol and registration. The review was undertaken based on a prospectively developed and registered protocol, International Prospective Register of Systematic Reviews (PROSPERO) registration number: CRD42019122999.

Eligibility criteria. Randomized clinical trials (RCTs) including people aged 18 years or older and with over 90% of participants with idiopathic (primary) frozen shoulder (adhesive capsulitis), with or without diabetes, were included. Studies treating general shoulder conditions were only included if outcomes were reported separately for participants with a frozen shoulder.

Eligible interventions were: MUA with or without a steroid injection; ACR, with or without an MUA but not 'open' capsular release; any form of physiotherapy techniques with steroid injection (PTSI); and

hydrodilatation, with or without guidance by radiological imaging. MUA, ACR, and hydrodilatation were included with or without follow-up physiotherapy. Any of the included interventions, supportive care (e.g. leaflets, home exercises, watchful waiting, painkillers), or no treatment were allowed as a comparator.

The primary outcome was patient self-reported function and disability (e.g. Oxford Shoulder Score, Shoulder Disability Questionnaire). The primary endpoint was 12-month follow-up. Other outcomes were: quality of life (using standardized outcome measures), pain (e.g. visual analogue pain scores), time to recovery, return to work and recreation, complications, and adverse events (number and type). There was no restriction on the setting in which a study was undertaken. No language restrictions were applied (one abstract screened required translation from German, which was conducted by AK).

Information sources and search. The search strategy used in the 2012 review was adapted to search for any new RCTs of the interventions of interest (see Supplementary Material).[9] The searches for the original systematic review were undertaken in March 2010; therefore, the start date of January 2010 was used for the updated search, and overlapping was used to allow for any delays in articles being added to the bibliographic databases. MEDLINE, Embase, PEDro, Science Citation Index, and Clinicaltrials.gov were searched on 7 December 2018; Cochrane Central Register of Controlled Trials (CENTRAL) was searched on 5 December 2018; and WHO International Clinical Trials Registry was searched on 11 December 2018. The reference lists of eligible studies were reviewed for further

Table II. Baseline pain and functioning.

Study	Outcome	Outcome description	MUA	ACR	Steroid + physiotherapy	Hydrodilatation	Supportive care
UK FROST[11] Mean (SD)	Pain NRS	Range: 0 to 10; lower is better	6.8 (2.23)	7 (1.89)	6.9 (2.37)	N/A	N/A
	OSS	Range: 0 to 60; higher is better	20.5 (8.88)	19.1 (7.72)	20.3 (7.97)	N/A	N/A
De Carli[24] Mean (SD not available)	SST	Range: 0 to 100; higher is better	N/A	15.6	30.1	N/A	N/A
Gallacher et al[29] Mean (SD)	OSS	Range: 0 to 60; higher is better	N/A	17.3 (7.2)	N/A	16.2 (5.2)	N/A
Kivimaki[28] Mean (SD)	SDQ	Range: 0 to 28; lower is better	22.7 (4.9)	N/A	N/A	N/A	21.7 (4.6)
Mukherjee et al[25] Mean (SD)	Pain VAS	Range: 0 to 10; lower is better	N/A	7.1 (1.8)	7.1 (1.8)	N/A	N/A
Quraishi[27] Mean (range)	Pain VAS	Range: 0 to 10; lower is better	5.7 (3 to 8.5)	N/A	N/A	6.1 (4 to 10)	N/A
Smitherman[30] Mean (SD)	SPADI	Range: 0 to 100; lower is better	N/A	70 (11)	N/A	N/A	82 (12)

N/A, not applicable; NRS, numerical rating scale; OSS, Oxford Shoulder Score; SDQ, Shoulder Disability Questionnaire; SPADI, Shoulder Pain and Disability Index; SST, Simple Shoulder Test; VAS, visual analogue scale.

studies. The RCTs identified from the original review were reassessed for inclusion in this review.

Study selection. Literature search results were uploaded to Clarivate's EndNote referencing software (Clarivate Analytics, USA) and exported to Covidence,[17] an online systematic review programme to remove duplicates and facilitate collaboration.[17] Two researchers (SR and LK) independently screened all titles and abstracts, to identify potentially relevant studies. Full manuscripts of potentially relevant studies were independently assessed by the two researchers against the eligibility criteria. Disagreements over eligibility were resolved by recourse to a third researcher (SB).[17]

Data collection process. A data extraction form was developed in Microsoft Excel, piloted and adjusted using a small selection of studies. Data extraction was performed by one author (SR) and checked by a second researcher (LK). Any discrepancies were resolved through discussion or by recourse to a third researcher (SB). UK FROST results were obtained from the trial statistician which were unpublished data at the time of this review's analysis.

Data items. Information on study design, number randomized, loss to follow-up, country, setting, patient characteristics (including average age, sex, presence of diabetes, stage of condition), description of the interventions (and comparators), concomitant treatments, and outcome measures used were extracted from all the included studies. When the number of participants in an analysis was unclear, the number randomized minus the number of dropouts was used.
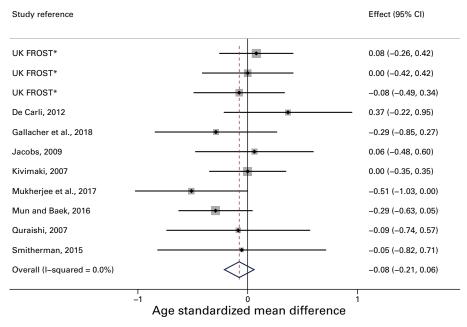
Risk of bias in individual studies. The Cochrane Risk of Bias Tool[18] was used to assess the risk of bias in included RCTs.[19] Quality assessment was undertaken by one researcher (SR) and checked by a second (LK); disagreements were adjudicated by a third researcher (SB). Baseline heterogeneity was explored by carrying out a random effects

meta-analysis of baseline age by treatment group.[20] This was to explore for any potential selection bias.

Summary measures. The primary outcome was patient self-reported shoulder function and disability. Outcomes were grouped as short- (three months or closest), medium- (six months or closest), and long-term (12 months or closest; primary) follow-up.

For continuous outcomes the post-intervention mean (standard deviation (SD) and number of participants) or the mean change from baseline for each group (SD of the change and number of participants) or the effect estimate and standard error (SE) were extracted. Both unadjusted and adjusted data were extracted, but the latter used where possible. Standard data imputation methods were used, where necessary, to calculate SDs or SEs.[19] When that was not possible the average SD across all trials or other existing literature were used. Where only median and ranges are reported, these were extracted, and, if necessary, used to calculate the mean and SD.[21]

Synthesis of results. A narrative and tabular summary of key study characteristics, results, and quality assessment were provided. An assessment was made about the clinical and methodological similarity of included RCTs to establish whether a quantitative synthesis was appropriate. Studies were grouped by intervention and comparator.

Where there were more than two studies reporting the same outcome, they were pooled using a random effects meta-analysis. A weighted mean difference (if the same scales were used) or standardized mean difference (SMD) (if different measurements scales were used) were calculated for continuous outcomes and risk ratio for dichotomous outcomes. Where estimates and SEs were extracted, a generic inverse variance approach was used. Statistical heterogeneity was assessed using the chi-squared test and quantified using the I-squared statistic. Trials with more than two treatment comparisons were treated as

Baseline age heterogeneity

Study reference	Effect (95% CI)
UK FROST*	0.08 (−0.26, 0.42)
UK FROST*	0.00 (−0.42, 0.42)
UK FROST*	−0.08 (−0.49, 0.34)
De Carli, 2012	0.37 (−0.22, 0.95)
Gallacher et al., 2018	−0.29 (−0.85, 0.27)
Jacobs, 2009	0.06 (−0.48, 0.60)
Kivimaki, 2007	0.00 (−0.35, 0.35)
Mukherjee et al., 2017	−0.51 (−1.03, 0.00)
Mun and Baek, 2016	−0.29 (−0.63, 0.05)
Quraishi, 2007	−0.09 (−0.74, 0.57)
Smitherman, 2015	−0.05 (−0.82, 0.71)
Overall (I−squared = 0.0%)	−0.08 (−0.21, 0.06)

Age standardized mean difference

NOTE: Weights are from random−effects model; DerSimonian−Laird estimator

Fig. 2

Baseline age heterogeneity *For UK FROST, total number of patients in each group were divided by three, to account for multiple comparisons. CI, confidence interval.
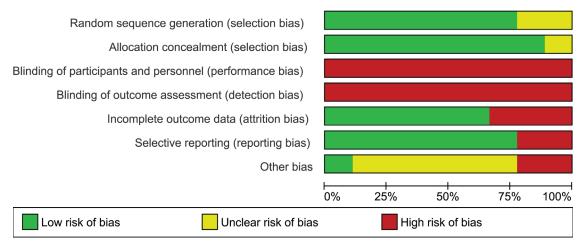
Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Other bias

0%   25%   50%   75%   100%

■ Low risk of bias   ■ Unclear risk of bias   ■ High risk of bias
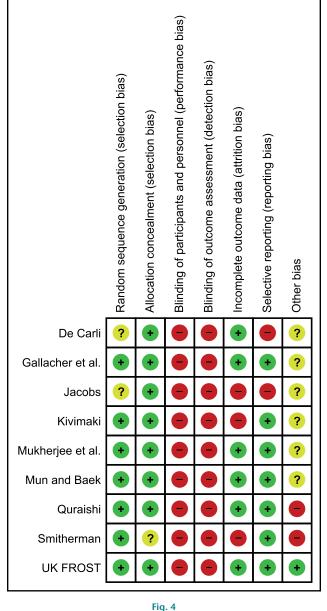
Fig. 3
Risk of bias summary (% of studies).

multiple individual trials by splitting the shared group into two or more groups. No formal statistical pooling of complications was undertaken, instead a descriptive summary was provided. All statistical analyses were carried out using RevMan 5 and STATA v15 (StataCorp, USA).

Risk of bias across studies. The effect of risk of bias on the effect estimate in a sensitivity analysis was not undertaken as planned by omitting studies that were judged to be of high risk of bias, as there were insufficient studies. To explore the potential for reporting bias, funnel plots would be generated.[22,23]

Additional analysis. Network meta-analysis and subgroup analysis on studies with and without diabetic patients were not undertaken as planned in the protocol due to the limited number of studies included.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
De Carli	?	+	−	−	+	−	?
Gallacher et al.	+	+	−	−	+	+	?
Jacobs	?	+	−	−	−	−	?
Kivimaki	+	+	−	−	−	+	?
Mukherjee et al.	+	+	−	−	+	+	?
Mun and Baek	+	+	−	−	+	+	?
Quraishi	+	+	−	−	+	+	−
Smitherman	+	?	−	−	−	+	−
UK FROST	+	+	−	−	+	+	+

Fig. 4

Risk of bias by study.

Results

Study selection. There were 2,786 studies identified through database screening and 31 studies identified through other sources, such as the previous review and citations from selected studies. Following deduplication, 2,813 studies were transferred to Covidence for screening. A further four studies were removed as duplicates. A total of 2,531 studies were excluded when screening the title and abstract. Overall, 282 full-text studies were assessed; 273 studies were excluded due to not meeting our eligibility criteria, or when insufficient information was available (clarification was sought from corresponding authors). The list of the excluded studies is available from the authors on request. Nine studies were eligible

for inclusion. Figure 1 summarizes the study selection process.

Study characteristics. Eight of the studies were published between 2007 and 2018 and unpublished data were supplied by the UK FROST trial team. The results of this trial have since been published in 2020.[11] Four of the included studies were published since the 2012 HTA review. There were eight two-arm trials and a single three-arm trial. UK FROST is the only study that compared MUA with PTSI or ACR with MUA. Three studies - UK FROST,[11] De Carli,[24] and Mukherjee et al[25] - compared ACR with PTSI. Two studies, Jacobs[26] and Quraishi,[27] compared MUA with hydrodilatation. Kivimaki[28] compared MUA with supportive care. Gallacher et al[29] is the only study that compared ACR with hydrodilatation. Smitherman[30] compared ACR with supportive care. Mun and Baek[31] compared hydrodilatation with PTSI.

The interventions varied considerably between studies, for example, ACR was performed with MUA in two studies,[11,24] and without MUA in three studies.[26,28,29] The eligibility criteria were consistent across studies. Final follow-up for most studies was 12 months but it ranged from 20 weeks to 24 months. Table I provides a summary of study characteristics. Additional details and eligibility criteria can be found in Supplementary Tables i and ii.

Only patient-reported shoulder function scores and pain scores were extracted for baseline as these were the most commonly reported outcomes. Table II provides an overview of pain and shoulder function measures reported in the individual studies and baseline scores for each intervention group. The outcome measures reported at baseline were insufficient, therefore comparison between studies could not be made. Two included studies show evidence of baseline imbalance between the two groups, which could affect the precision of the unadjusted differences in outcome.[24,30]

At baseline, the mean age of patients in all RCTs was the sixth decade of their life (fifties). The percentage of females ranged from 54% to 70%. Seven of the nine RCTs were known to include diabetics. Duration of symptoms ranged from four to 11 months. Baseline heterogeneity in age was assessed for the different treatment groups (Figure 2). Only one study resulted in an upper confidence interval that was not positive (i.e. zero).[25] The overall I^2 measure of heterogeneity was 0%. This suggests that that there is no reason to suspect selection bias in the included studies based on age at baseline.[20] Other baseline variables were not included for this assessment, as they were not fully reported.

Risk of bias within studies. Most of the studies reported how the allocation sequence was generated and concealed. Due to the nature of the interventions, participants were not blinded to their treatment group. Since the outcomes of interest for this review were all patient-reported, all studies were marked as high risk of bias for

Shoulder scores

Treatment comparisons and studies	Effect (95% CI)

ACR vs Hydrodilatation
Gallacher et al., 2018 (Medium) — 0.77 (0.12, 1.42)

ACR vs MUA
UK FROST (Short) — −0.35 (−0.57, −0.13)
UK FROST (Medium) — −0.13 (−0.35, 0.09)
UK FROST (Long) — 0.21 (−0.01, 0.43)

ACR vs Standard care
Smitherman, 2015 (Short) — 0.20 (−0.76, 1.16)
Smitherman, 2015 (Long) — 0.13 (−0.83, 1.09)

ACR vs Steroid + Physio
UK FROST (Short) — −0.50 (−0.75, −0.25)
De Carli, 2012 (Short) — 0.61 (0.00, 1.22)
UK FROST (Medium) — 0.11 (−0.14, 0.36)
De Carli, 2012 (Medium) — 0.50 (−0.11, 1.11)
De Carli, 2012 (Long) — 0.25 (−0.34, 0.84)
UK FROST (Long) — 0.33 (0.08, 0.58)

MUA vs Standard care
Kivimaki, 2007 (Short) — −0.04 (−0.43, 0.35)
Kivimaki, 2007 (Medium) — 0.20 (−0.23, 0.63)
Kivimaki, 2007 (Long) — 0.00 (−0.45, 0.45)

MUA vs Steroid + Physio
UK FROST (Short) — −0.15 (−0.40, 0.10)
UK FROST (Medium) — 0.24 (−0.01, 0.49)
UK FROST (Long) — 0.11 (−0.14, 0.36)

−1   0   1
Favours less–invasive intervention   Favours more–invasive intervention

Standardized mean difference

Fig. 5

Shoulder function scores. *Shoulder Disability Questionnaire and Shoulder Pain Disability Index Scores were reversed so that positive result implies better outcome. This was done to be consistent with other shoulder scores (e.g. Oxford Shoulder Score and Simple Shoulder Test). ACR, arthroscopic capsular release; CI, confidence interval; MUA, manipulation under anaesthesia.

'blinding of participants and personnel' and 'blinding of outcome assessment', following Cochrane guidance.[18]

Studies with high attrition rate (i.e. over 30% in any single arm) were marked as high risk of bias for 'incomplete outcome data'.[26,28,30] Two studies were marked as high risk of bias for selective reporting, as one of them reported an outcome measure for only one of their time-points,[24] and another did not reporting on their final follow-up.[26]

Three studies did not provide clear reasons for non-consent and dropouts;[24,28,29] one study only followed up patients for 20 weeks;[25] and one was from a single institution.[31] These studies were marked as 'unclear' for other risk of bias. One study had all MUA treatments performed by a single surgeon and hydrodilatation by another doctor, potentially introducing treatment bias.[27] Another reported treatment effectiveness based on a very low sample size.[30] These studies were considered as 'high' for other risk of bias. Figures 3 and 4 illustrate risk of bias in the included studies.

Due to the limited number of eligible studies with multiple treatment comparisons, publication bias was hard to establish. A funnel plot was used to visualize the standardized mean difference of shoulder function in the long term (Supplementary Figure b).

Primary outcomes. The primary outcome of shoulder function score was reported for five studies comparing six combinations of interventions.

Gallacher et al[29] compared ACR with hydrodilatation. At their final follow-up at six months, the Oxford Shoulder Score (OSS) was significantly better in the ACR group than in the hydrodilatation group (unadjusted mean difference: 5.3 (95% CI 1.16 to 9.44)).

Smitherman et al[30] compared ACR with supportive care. At their final follow-up at 12 months, participants randomized to ACR had slightly better SPADI scores than the supportive care group but this difference was not statistically significant (unadjusted mean difference: -5.00 (95% CI -29.16 to 19.16)).

Three studies compared ACR to PTSI,[10,24,25] and two reported function and disability at 12 months,[10,24] with the other study reporting pain only.[25] De Carli[24] reported the primary outcome (Simple Shoulder Test) at 12 months (unadjusted mean difference: 0.59 (95% CI -0.77 to 1.95)) and UK FROST reported the primary outcome (OSS) at 12 months (adjusted mean difference: 3.06 (95% CI 0.71 to 5.41)).[11] De Carli is a single centre study, while UK FROST is multicentre.
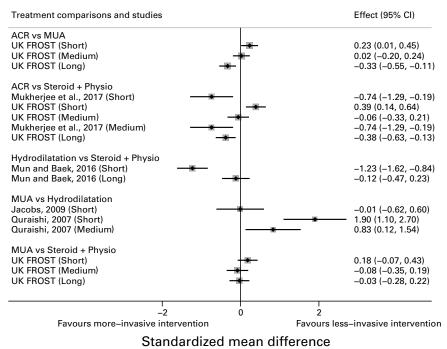
Kivimaki[28] compared MUA with supportive care at their final follow-up at 12 months, and did not find any

Table III. Summary of results.

Study	Scale*	Short term (≤ 3 mths)	Medium term (> 3 and ≤ 6 mths)	Long term (> 6 and ≤ 12 mths)
ACR vs Hydrodilatation				
OSS (higher is better)				
Gallacher et al,[29] 2018 (ACR: n = 25, Hydro: n = 25)	MD		5.3 (1.16 to 9.44)	
	SMD		0.77 (0.12 to 1.42)	
ACR vs MUA				
OSS (higher is better)				
UK FROST[11] (ACR: n = 203, MUA: n = 201)	MD	-3.36 (-5.27 to -1.45)	-1.17 (-3.02 to 0.67)	2.01 (0.1 to 3.91)
	SMD	-0.35 (-0.56 to -0.14)	-0.13 (-0.34 to 0.08)	0.21 (0.00 to 0.42)
NRS - Pain (lower is better)				
UK FROST[11] (ACR: n = 203, MUA: n = 201)	MD	0.59 (0.1 to 1.07)	0.05 (-0.43 to 0.52)	-0.73 (-1.2 to -0.25)
	SMD	0.24 (0.03 to 0.44)	0.00 (-0.21 to 0.21)	-0.32 (-0.53 to -0.11)
ACR vs Supportive care				
SPADI (lower score is better)				
Smitherman et al[30] (ACR: n = 13, Supportive care: n = 13)	MD	-5 (-29.16 to 19.16)		-2(-15.39 to 11.39)
	SMD	-0.20 (-1.17 to 0.77)		-0.13 (-1.10 to 0.83)
ACR vs PTSI				
OSS (higher is better)				
UK FROST[11] (ACR: n = 203, PTSI: n = 99)	MD	-4.72 (-7.06 to -2.39)	0.98 (-1.31 to 3.26)	3.06 (0.71 to 5.41)
	SMD	-0.50 (-0.76 to -0.24)	0.11 (-0.15 to 0.38)	0.33 (0.07 to 0.59)
SST (higher is better)				
De Carli[24] (ACR: n = 25, PTSI: n = 21)	MD	1.44 (0.08 to 2.8)	1.18 (-0.18 to 2.54)	0.59 (-0.77 to 1.95)
	SMD	0.61 (0.01 to 1.22)	0.50 (-0.10 to 1.11)	0.25 (-0.34 to 0.85)
NRS - Pain (lower is better)				
UK FROST[11] (ACR: n = 203, PTSI: n = 99)	MD	1.02 (0.42 to 1.61)	-0.14 (-0.74 to 0.45)	-0.81 (-1.39 to -0.23)
	SMD	0.38 (0.13 to 0.64)	-0.09 (-0.36 to 0.18)	-0.38 (-0.64 to -0.12)
VAS – Pain (lower is better)				
Mukherjee et al[25] (ACR: n = 30 PTSI: n = 30)	MD	-1.2 (-2.04 to -0.36)	-1.2 (-2.04 to -0.36)	
	SMD	-0.74 (-1.28 to -0.20)	-0.74 (-1.28 to -0.20)	
Hydrodilatation vs PTSI				
VAS – Pain (lower is better)				
Mun and Baek[31] (Hydro: n = 67, PTSI: n = 69)	MD	-0.9 (-1.16 to -0.64)		-0.1 (-0.39 to 0.19)
	SMD	-1.23 (-1.62 to -0.84)		-0.12 (-0.48 to 0.23)
MUA vs Hydrodilatation				
VAS – Pain (lower is better)				
Jacobs[26] (MUA: n = 28, Hydro: n = 25)	MD	-0.02 (-1.15 to 1.11)		
	SMD	-0.01 (-0.61 to 0.59)		
Quraishi[27] (MUA: n = 17, Hydro: n = 19)	MD	2.3 (1.51 to 3.09)	1 (0.21 to 1.79)	
	SMD	1.90 (1.08 to 2.73)	0.83 (0.12 to 1.53)	
MUA vs Supportive care				
SDQ (lower score is better)				
Kivimaki[28] (MUA: n = 65, Supportive care: n = 60)	MD	0.3 (-2.69 to 2.75)	-1.7 (-5.3 to 1.9)	0 (-3.2 to 3.2)
	SMD	0.04 (-0.35 to 0.43)	-0.2 (-0.63 to 0.23)	0 (-0.44 to 0.44)
MUA vs PTSI				
OSS (higher is better)				
UK FROST[11] (MUA: n = 201, PTSI: n = 99)	MD	-1.36 (-3.7 to 0.98)	2.15 (-0.12 to 4.42)	1.05 (-1.28 to 3.39)
	SMD	-0.15 (-0.4 to 0.10)	0.24 (-0.02 to 0.51)	0.12 (-0.14 to 0.37)
NRS - Pain (lower is better)				
UK FROST[11] (MUA: n = 201, PTSI: n = 99)	MD	0.43 (-0.17 to 1.03)	-0.19 (-0.78 to 0.4)	-0.08 (-0.66 to 0.5)
	SMD	0.17 (-0.09 to 0.42)	-0.09 (-0.35 to 0.18)	-0.04 (-0.30 to 0.21)

SST score was reported as a percentage so was converted to original scale. SD was not provided so was imputed by taking the average SD reported by Yoon et al.[32]
VAS Pain SD was not reported by Quraishi.[27] This value was imputed by taking the average SD reported from other VAS scores reported.
*Due to the variability in measures used in the included studies, the standardized mean difference in addition to mean difference are reported to allow comparison between studies.
ACR, arthroscopic capsular release; MD, mean difference; MUA, manipulation under anaesthesia; NRS, numerical rating scale; OSS, Oxford Shoulder Score; PTSI, physiotherapy techniques with steroid injection; SD, standard deviation; SDQ, Shoulder Disability Questionnaire; SMD, standardized mean difference; SPADI, Shoulder Pain and Disability Index; SST, Simple Shoulder Test; VAS, visual analogue scale.

Pain scores

Treatment comparisons and studies	Effect (95% CI)

ACR vs MUA
UK FROST (Short) — 0.23 (0.01, 0.45)
UK FROST (Medium) — 0.02 (−0.20, 0.24)
UK FROST (Long) — −0.33 (−0.55, −0.11)

ACR vs Steroid + Physio
Mukherjee et al., 2017 (Short) — −0.74 (−1.29, −0.19)
UK FROST (Short) — 0.39 (0.14, 0.64)
UK FROST (Medium) — −0.06 (−0.33, 0.21)
Mukherjee et al., 2017 (Medium) — −0.74 (−1.29, −0.19)
UK FROST (Long) — −0.38 (−0.63, −0.13)

Hydrodilatation vs Steroid + Physio
Mun and Baek, 2016 (Short) — −1.23 (−1.62, −0.84)
Mun and Baek, 2016 (Long) — −0.12 (−0.47, 0.23)

MUA vs Hydrodilatation
Jacobs, 2009 (Short) — −0.01 (−0.62, 0.60)
Quraishi, 2007 (Short) — 1.90 (1.10, 2.70)
Quraishi, 2007 (Medium) — 0.83 (0.12, 1.54)

MUA vs Steroid + Physio
UK FROST (Short) — 0.18 (−0.07, 0.43)
UK FROST (Medium) — −0.08 (−0.35, 0.19)
UK FROST (Long) — −0.03 (−0.28, 0.22)

−2   0   2
Favours more–invasive intervention   Favours less–invasive intervention

Standardized mean difference

Fig. 6

Pain scores across different studies. ACR, arthroscopic capsular release; CI, confidence interval; MUA, manipulation under anaesthesia.

statistically significant difference between treatment groups (adjusted mean difference: 0.0 (95% CI -3.2 to 3.2)). This study was conducted from three regional hospitals.

UK FROST found that at 12 months, participants randomized to ACR had significantly better OSS scores than MUA (2.01 (95% CI 0.10 to 3.91)) after adjusting for baseline OSS, age, sex, and diabetes. There was no statistically significant difference between MUA and PTSI (1.05 (95% CI -1.28 to 3.39)). None of these differences met the threshold of clinical significance agreed for the trial i.e. a four-point difference on the OSS between MUA and ACR (i.e. a SMD of 0.33), and a five-point difference between PTSI and either surgery (i.e. a SMD of 0.42).

Due to the limited number of included studies and heterogeneity (I² ranging from 27% to 100%), for the other follow-up points, only a pooled estimate of the primary outcome of interest is reported (Supplementary Figure a).

The differences in short- and medium-term pain and shoulder function varied considerably between studies, whereas long-term outcomes tended to favour more invasive treatment, such as ACR generally. In Figure 5 a positive SMD indicates a more favourable outcome for the first treatment in each comparison e.g. favours ACR in ACR versus PTSI. In Figure 6 a

positive SMD indicates a more favourable outcome for the second treatment in each comparison e.g. favours PTSI in ACR versus PTSI.

Table III contains both mean and standard mean differences for short-, medium-, and long-term follow-up periods for both shoulder function and pain score.

Secondary and additional outcomes: patient-reported pain scores. The secondary outcome of patient-reported pain score was reported for five studies comparing five combinations of interventions.

Mukherjee et al[25] compared ACR with PTSI. At their final follow-up at 20 weeks, participants in ACR group had statistically significant lower (less) pain than the PTSI group (unadjusted mean difference: -1.20 (95% CI -2.04 to -0.36)).

Mun and Baek[31] compared hydrodilatation with PTSI. At their final follow-up at 12 months, there was no statistically significant difference between groups on VAS pain scores (adjusted mean difference: -0.10 (95% CI -0.39 to 0.19)). This was a single-centre study.

Jacobs[26] and Quraishi[27] compared MUA with hydrodilatation. Jacobs,[26] in a single-centre study, reported VAS mean difference at 16 weeks and there was no statistically significant difference between groups (adjusted mean difference: -0.02 (95% CI -1.15 to 1.11)). Quraishi[27] at six months' follow-up found that VAS pain in the MUA group was significantly higher (worse) than in the

hydrodilatation group (unadjusted mean difference: 1.00 (95% CI 0.21 to 1.79)).

UK FROST at 12 months compared pain scores for MUA versus PTSI (adjusted mean difference: -0.08 (95% CI -0.66 to 0.50)); ACR versus PTSI (adjusted mean difference: -0.81 (95% CI -1.39 to -0.23)) and ACR versus MUA (adjusted mean difference: -0.73 (95% CI -1.20 to -0.25)).

Complications reported. Of the nine RCTs included in the systematic review, five reported no complications in any of the treatment groups.[24,26,29–31] One RCT did not report whether there were any complications at all, therefore it is unclear whether complications were assessed.[27] Kivimaki[28] reported that there were no major complications in the MUA group but that small injuries of the joint were possible, as verified on arthroscopy. No complications were reported for the supportive care group. Mukherjee et al[25] found one case of articular cartilage scuffing of glenoid and one case of the humeral head in the ACR group. UK FROST reported more complications in the ACR group compared to MUA and PTSI.[11]

Other outcomes of interests such as quality of life (using standardized outcome measures), time to recovery, and return to work and recreation were not reported by studies apart from UK FROST and consequently were not reported here.

Discussion

Nine eligible RCTs comparing PTSI, MUA, ACR, or hydrodilatation were identified. Five RCTs,[24,25,29–31] including UK FROST, were eligible since the 2012 review which had found an evidence gap in the effectiveness of these interventions.

UK FROST, to the authors' knowledge, is the largest multicentre RCT comparing three of the treatments of interest, while most of the other comparisons between treatments are informed by single site studies with limited sample sizes. UK FROST provided unbiased evidence except for blinding, which can be argued to be neither feasible with the interventions being evaluated nor necessarily desirable in a pragmatic trial design that reflects the real world delivery of care.[33] In comparison, other RCTs were susceptible to additional bias concerning, for example, incomplete outcome data or selective reporting of outcomes. In the UK FROST trial, although ACR was statistically significant superior to PTSI, the treatment effect was small and not clinically important.[11] Therefore, none of the treatment options compared were clinically superior. In UK FROST, MUA was the most cost-effective option to the NHS.[11,34]

Four RCTs that evaluated hydrodilatation were included,[26,27,29,31] with a duration of follow-up ranging from six months to two years. Evidence was inconclusive with limited sample sizes (ranging from 20 to 60 patients in the hydrodilatation arm) and two of the four RCTs were at risk of more than three types of bias.[26,30] Although a recent network meta-analysis of non-surgical treatment strategies for frozen shoulder conducted by Zhang et al[4] found hydrodilatation to be one of the highest ranked treatments for short-term pain relief, the diverse group of interventions included and lack of longer-term follow-up (i.e. 12 months) makes it difficult to make robust conclusions for clinical practice and policy. Another network meta-analysis conducted by Challoumas et al[35] supports the Zhang et al[4] conclusion that hydrodilatation is better for short term pain relief, while mid-term results favoured PTSI. A ten-year review of 2,432 hydrodilatations conducted by Nicholson[36] found the repeat intervention rate to be 7.6% and only 1.7% requiring a subsequent ACR, suggesting that hydrodilatation may prevent the need for more costly and invasive ACR. An evidence gap, however, remains for high-quality evidence for the effectiveness of hydrodilatation compared with other commonly used treatments for primary frozen shoulder.

The protocol for this review was pre-registered in PROSPERO and followed PRISMA guidance for preparing a protocol and reporting of a systematic review and meta-analyses.[15,16] Therefore, a comprehensive and rigorous review was undertaken. We specifically focused on putting UK FROST in the context of other randomized evidence, which is why we concentrated on physiotherapy techniques with a steroid injection as recommended by evidence-based clinical guidelines when designing the trial.[37] Including hydrodilatation identified the important finding of the lack of high-quality evidence to support its increasing popularity.

Given the number of treatment options being investigated, it would have been ideal to perform a network meta-analysis, however this was not possible due to the limited number of eligible RCTs, nor was the sub-analysis on studies with and without diabetic patients feasible. The follow-up timepoints were variable, as were the interventions, and RCTs were often of limited sample sizes and may have been underpowered. The high risk of bias present in some RCTs also limited the quality of evidence. Only patient-reported outcomes measures such as shoulder function and pain scores, alongside complications, were evaluated in this review as these more specifically assess the impact of the condition on participants' lives and their experience of the illness than clinical measurements such as shoulder range of motion.[38]

Overall, the findings from a recent multicentre RCT provided the strongest evidence that when compared with each other neither PTSI, MUA, nor ACR are clinically superior. Evidence from smaller RCTs did not change this conclusion. The effectiveness of hydrodilatation from

four RCTs was inconclusive and remains an evidence gap.

Take home message

- There were few high-quality and adequately powered randomized studies that evaluated key treatment choices in the management of patients with primary frozen shoulder.
- UK FROST has addressed this and provided the strongest evidence that when compared with each other, neither physiotherapy techniques with steroid injection, manipulation under anaesthesia, nor arthroscopic capsular release are clinically superior. Evidence from smaller randomized controlled trials (RCTs) did not change this conclusion.
- There is an evidence gap suggesting the need for a rigorously designed and adequately powered RCT to evaluate the comparative effectiveness of hydrodilatation.

Twitter

Follow S. S. Rex @Sal_MedStats
Follow L. Kottam @SouthTees, @STRIVEacademic, @LKottam
Follow C. McDaid @HealthSciYork @Catriona_McDaid
Follow S. Brealey @YorkTrialsUnit
Follow C. E. Hewitt @YorkTrialsUnit
Follow A. Keding @YorkTrialsUnit
Follow K. Wright @crd_york
Follow A. Rangan @YorkTrialsUnit, @SouthTees

Supplementary material

All the search strategies followed for this systematic review, and additional tables and figures quoted in the main document.

References

1. **van der Windt DA**. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis*. 1995;54(12):959–964.
2. **Walker-Bone K**. Prevalence and impact of musculoskeletal disorders of the upper limb in the general population. *Arthritis Rheum*. 2004;51(4):642–651.
3. **Rangan A**, **Hanchard N**, **McDaid C**. What is the most effective treatment for frozen shoulder. *BMJ*. 2016;354:i4162.
4. **Zhang J**, **Zhong S**, **Tan T**, **et al**. Comparative Efficacy and Patient-Specific Moderating Factors of Nonsurgical Treatment Strategies for Frozen Shoulder: An Updated Systematic Review and Network Meta-analysis. *Am J Sports Med*. 2021;49(6):1669–1679.
5. **Buchbinder R**. Arthrographic distension for adhesive capsulitis (frozen shoulder). *Cochrane Database Syst Rev*. 2008;1:Cd007005.
6. **Blanchard V**, **Barr S**, **Cerisola FL**. The effectiveness of corticosteroid injections compared with physiotherapeutic interventions for adhesive capsulitis: a systematic review. *Physiotherapy*. 2010;96(2):95–107.
7. **Lin MT**. Comparative efficacy of intra-articular steroid injection and distension in patients with frozen shoulder: A systematic review and network meta-analysis. *Arch Phys Med Rehabil*. 2018;99(7):1383–1394.
8. **Kraal T**. Manipulation under anaesthesia for frozen shoulders: outdated technique or well-established quick fix. *EFORT Open Rev*. 2019;4(3):98–109.
9. **Maund E**. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. *Health Technol Assess*. 2012;16(11):1–264.
10. **Brealey S**. United Kingdom Frozen Shoulder Trial (UK FROST), multi-centre, randomised, 12 month, parallel group, superiority study to compare the clinical and cost-effectiveness of Early Structured Physiotherapy versus manipulation under anaesthesia versus arthroscopic capsular release for patients referred to secondary care with a primary frozen shoulder: study protocol for a randomised controlled trial. *Trials*. 2017;18(1):614.
11. **Rangan A**, **Brealey SD**, **Keding A**, **et al**. Management of adults with primary frozen shoulder in secondary care (UK FROST): A multicentre, pragmatic, three-arm, superiority randomised clinical trial. *Lancet*. 2020;396(10256):977–989.
12. **Hanchard NCA**. Physiotherapy for primary frozen shoulder in secondary care: Developing and implementing stand-alone and post operative protocols for UK FROST and inferences for wider practice. *Physiotherapy*. 2020;107:150–160.
13. **Koh KH**. Corticosteroid injection for adhesive capsulitis in primary care: a systematic review of randomised clinical trials. *Singapore Med J*. 2016;57(12):646–657.
14. **Dennis L**. Managing Idiopathic Frozen Shoulder: A Survey of Health Professionals' Current Practice and Research Priorities. *Shoulder & Elbow*. 2010;2(4):294–300.
15. **Moher D**. Preferred reporting Items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
16. **Shamseer L**, **Moher D**, **Clarke M**, **et al**. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;350:g7647.
17. **No authors listed**. Covidence: Systematic review management. https://www.covidence.org (date last accessed 23 July 2021).
18. **Higgins JPT**. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
19. **Higgins JPT**, **Green S**. Cochrane handbook for systematic reviews of interventions version 5.1.0. 2011. https://www.training.cochrane.org/handbook (date last accessed 18 August 2021).
20. **Hicks A**, **Fairhurst C**, **Torgerson DJ**. A simple technique investigating baseline heterogeneity helped to eliminate potential bias in meta-analyses. *J Clin Epidemiol*. 2018;95:55–62.
21. **Hozo SP**, **Djulbegovic B**, **Hozo I**. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13.
22. **Sterne JAC**. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
23. **The Cochrane Collaboration**. Review Manager (RevMan). 2020. https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman (date last accessed 18 August 2021).
24. **De Carli A**, **Vadalà A**, **Perugia D**, **et al**. Shoulder adhesive capsulitis: Manipulation and arthroscopic arthrolysis or intra-articular steroid injections? *Int Orthop*. 2012;36(1):101–106.
25. **Mukherjee RN**, **Pandey RM**, **Nag HL**, **Mittal R**. frozen shoulder - a prospective randomized clinical trial. *World J Orthop*. 2017;8(5):394–399.
26. **Jacobs LG**. Manipulation or intra-articular steroids in the management of adhesive capsulitis of the shoulder? A prospective randomized trial. *J Shoulder Elbow Surg*. 2009;18(3):348–353.
27. **Quraishi NA**. Thawing the frozen shoulder. A randomised trial comparing manipulation under anaesthesia with hydrodilatation. *J Bone Joint Surg Br*. 2007;89-B(9):1197–1200.
28. **Kivimaki J**. Manipulation under anesthesia with home exercises versus home exercises alone in the treatment of frozen shoulder: a randomized, controlled trial with 125 patients. *J Shoulder Elbow Surg*. 2007;16(6):722–726.
29. **Gallacher S**, **Beazley JC**, **Evans J**, **et al**. A randomized controlled trial of arthroscopic capsular release versus hydrodilatation in the treatment of primary frozen shoulder. *J Shoulder Elbow Surg*. 2018;27(8):1401–1406.
30. **Smitherman JA**, **Struk AM**, **Cricchio M**, **et al**. Arthroscopy and manipulation versus home therapy program in treatment of adhesive capsulitis of the shoulder: A prospective randomized study. *J Surg Orthop Adv*. 2015 Spring;24(1):69–74.
31. **Mun SW**, **Baek CH**. Clinical efficacy of hydrodistention with joint manipulation under interscalene block compared with intra-articular corticosteroid injection for frozen shoulder: A prospective randomized controlled study. *J Shoulder Elbow Surg*. 2016;25(12):1937–1943.
32. **Yoon JP**, **Chung SW**, **Kim J-E**, **et al**. Intra-articular injection, subacromial injection, and hydrodilatation for primary frozen shoulder: A randomized clinical trial. *J Shoulder Elbow Surg*. 2016;25(3):376–383.
33. **Zwarenstein M**, **Treweek S**, **Gagnier JJ**, **et al**. Improving the reporting of pragmatic trials: An extension of the CONSORT statement. *BMJ*. 2008;337:a2390.
34. **Brealey S**, **Northgraves M**, **Kottam L**, **et al**. Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: The UK FROST three-arm RCT. *Health Technol Assess*. 2020;24(71):1–162.
35. **Challoumas D**, **Biddle M**, **McLean M**, **Millar NL**. Comparison of treatments for frozen Shoulder: A systematic review and meta-analysis. *JAMA Netw Open*. 2020;3(12):e2029581.
36. **Nicholson JA**, **Slader B**, **Martindale A**, **Mckie S**, **Robinson CM**. Distension arthrogram in the treatment of adhesive capsulitis has a low rate of repeat intervention. *Bone Joint J*. 2020;102-B(5):606–610.
37. **Hanchard NC**, **Goodchild L**, **Thompson J**, **O'Brien T**, **Davison D**, **Richardson C**, **et al**. Evidence-based clinical guidelines for the diagnosis, assessment and physiotherapy management of contracted (frozen) shoulder: Quick reference summary. *Physiotherapy*. 2012;98(2):117–120.

38. **Makrinioti H**, **Bush A**, **Griffiths C**. What are patient-reported outcomes and why they are important: improving studies of preschool wheeze. *Arch Dis Child Educ Pract Ed.* 2020;105(3):185–188.

Author information:
- S. S. Rex, BA, PG Dip, MSc, Research Fellow Statistician
- C. McDaid, PhD, MSc, BSSc, Reader
- S. Brealey, BSc, PhD, Trial Manager
- C. E. Hewitt, BSc, MSc, PhD, Professor of Medical Statistics, Deputy Director York Trials Unit
- A. Keding, BSc, MSc, Statistician
  York Trials Unit, Department of Health Sciences, University of York, York, UK.
- L. Kottam, BSc, MSc, PhD, Trials Manager, Department of Trauma & Orthopaedic Surgery, The James Cook University Hospital, Middlesbrough, UK.
- J. Dias, MB BS, FRCS, MD, Consultant Orthopaedic Surgeon, Leicester General Hospital, University Hospitals of Leicester NHS Trust, Leicester, UK.
- S. E. Lamb, D Phil, MSc, MA, PGCE, Graduate Diploma, Associate Dean for Research and & Mireille Gillings Professor of Health Innovation, College of Medicine and Health, University of Exeter, Exeter, UK.
- K. Wright, BA, MA, Information Service Manager, Centre for Reviews and Dissemination, University of York, York, UK.
- A. Rangan, FRCS(Tr&Orth), Mary Kinross Trust and Royal College of Surgeons Chair, Professor of Orthopaedic Surgery, Consultant Orthopaedic Surgeon, York Trials Unit, Department of Health Sciences, University of York, York, UK; Department of Trauma & Orthopaedic Surgery, The James Cook University Hospital, Middlesbrough, UK; Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.