BMJ Open

Ultrasound imaging to tailor the treatment of acute shoulder pain: a randomised controlled trial in general practice

Ramon P G Ottenheijm,[1] Jochen W L Cals,[1] Bjorn Winkens,[2] René E Weijers,[3] Rob A de Bie,[4] Geert-Jan Dinant[1]

CrossMark

For numbered affiliations see end of article.

Correspondence to
Dr Ramon PG Ottenheijm; ramon.ottenheijm@ maastrichtuniversity.nl

ABSTRACT

Objective: To determine the clinical effectiveness of ultrasound tailored treatment in patients with acute subacromial disorders.

Design: Pragmatic randomised controlled trial.

Setting: Dutch general practice.

Participants: Patients aged 18–65 years with acute (duration <3 months) unilateral shoulder pain and no previous treatment, in whom the general practitioner suspected a subacromial disorder was enrolled.

Interventions: All patients underwent ultrasound imaging of the affected shoulder. Patients who were still symptomatic after a qualification period of 2 weeks with standard treatment were randomised to treatment tailored to ultrasound diagnosis (disclosure of the ultrasound diagnosis) or usual care (non-disclosure of the ultrasound diagnosis).

Primary outcome measure: Patient-perceived recovery using the Global Perceived Effect questionnaire at 1 year.

Results: 129 patients were included. 18 patients recovered during the 2-week qualification period, resulting in 111 randomised patients; 56 were allocated to ultrasound tailored treatment and 55 to usual care. After 1 year, no statistically significant differences in recovery were found between the ultrasound tailored treatment group (72.5% (37/51)) and the usual care group (60% (30/50), OR 2.24 (95% CI 0.72 to 6.89; p=0.16)). Also, healthcare use was similar.

Conclusions: This study has shown no clinically significant difference in the primary outcome measure between the ultrasound tailored treatment and usual care groups. Furthermore, there was no overall difference in healthcare resources used between groups. Although no formal cost data are included, one can only assume that the ultrasound examinations are additional costs for the intervention group, which cannot be justified in routine practice based on this trial. Based on this study, no change in current pragmatic guidelines to incorporate early ultrasound imaging can be recommended.

Trial registration number: NTR2403; Results.

Strengths and limitations of this study

- This is the first primary care study investigating the effect of using ultrasound imaging in the management of patients with shoulder pain to target treatment to the specific underlying patho-anatomical disorders.
- This study was developed as a pragmatic trial to inform clinicians, guideline developers and policymakers to choose wisely between options for care.
- The trial does not give a conclusive answer to whether ultrasound tailored treatment improves outcome after 1 year, as our trial was under-enrolled. The limited size of this trial is a limitation that may have prevented the documentation of significant clinically important differences.

INTRODUCTION

General practitioners (GPs) are frequently consulted by patients with shoulder pain.[1–3] Prognosis is rather poor with 40% not being recovered after 1 year[4–6] and high recurrence rates.[7–9] These findings suggest that shoulder pain frequently progresses to a chronic disorder.

Shoulder pain is a symptom, not a diagnosis. Subacromial disorders are the most common cause of shoulder pain seen by GPs.[10 11] In general practice, accurate diagnosis of shoulder pain is difficult because findings from medical history and physical examination often poorly correlate with the underlying disorder.[12–14] Therefore, British and Dutch guidelines for shoulder pain advise GPs to start treatment based on patients' signs and symptoms rather than on the actual disorder.[10 15] GPs experience the current diagnostic process as complex, and the use of diagnostic ultrasound as helpful in establishing a more accurate diagnosis.[16–21]

Accuracy studies showed that diagnostic ultrasound is accurate for evaluating subacromial disorders.[22–24] Moreover, the full spectrum of subacromial disorders is observed in patients with shoulder pain presenting in general practice.[25–27] For each of these disorders, evidence-based treatments are available.[3 28 29] This implies that stratification of patients into diagnostic subgroups potentially allows for more tailored treatment than currently applied.[28] For instance, in general practice, patients with low back pain, a stratified management approach in which prognostic screening and treatment targeting were combined, improved patient outcome.[30]

In daily general practice, combining clinical information with ultrasound diagnosis is potentially helpful to tailor treatment to patients with shoulder pain. So far, no pragmatic trial has evaluated this test-treatment approach for shoulder pain in general practice. Pragmatic trials measure effectiveness in routine clinical practice and reflect variations between patients that occur in real clinical practice improving generalisability.[31] We conducted a pragmatic, randomised controlled trial, the Maastricht Ultrasound shoulder pain Trial (MUST), to study clinical effectiveness of ultrasound tailored treatment in patients with acute subacromial disorders.

METHODS
Design and participants
The study design and rationale of MUST were described elsewhere.[32] Patients were eligible if they had shoulder pain on abduction with painful arc; symptoms having lasted <3 months; no other episodes of shoulder pain in the previous 12 months and age between 18 and 65 years. Exclusion criteria were consultation or treatment for shoulder pain in the past 3 months; glenohumeral external rotation range of motion <45° as this is a reason to suspect a glenohumeral disorder like osteoarthritis or a frozen shoulder; history of fractures of the proximal humerus or acromion, dislocation and/or surgery of the affected shoulder; shoulder symptoms caused by rheumatic disease, suspected referred symptoms or extrinsic cause; history of depressive or anxiety disorders (negative prognostic factors) or pain catastrophising (irrational thought in believing pain is far worse than it actually is); inability to complete a questionnaire independently; unable to give informed consent (dementia or psychiatric disorders) and involved in disability or liability procedures. Initially, 21 GPs working in 11 practices in the *Westelijke Mijnstreek*, a region in the southern part of the Netherlands, recruited eligible patients. These GPs were asked to include sequential eligible patients within regular consultation hours. After 2 years, all 80 GPs in the aforementioned region were asked to recruit patients. All GPs received oral and written instructions.

All patients underwent ultrasound imaging of the affected shoulder by experienced musculoskeletal radiologists using a standardised protocol and criteria for pathology. Patients who were still symptomatic after a qualification period of 2 weeks with standard treatment were randomised to treatment tailored to ultrasound diagnosis (disclosure of the ultrasound diagnosis) or usual care (non-disclosure of the ultrasound diagnosis) (figure 1). The 2-week qualification period was used to perform diagnostic ultrasound and aimed to filter out patients with a favourable natural course. During this period, patients received treatment according to the shoulder pain guidelines of the Dutch College of General Practitioners: paracetamol or non-steroidal anti-inflammatory drugs (NSAIDs) in maximum dosage on a time contingent base, advice regarding activities of daily living, work, hobbies and sports.[10] Our reporting follows the Consolidated Standards of Reporting Trials (CONSORT) extension for pragmatic trials.[33] The study was approved by the Medical Ethics Committee of the Maastricht University Medical Centre, and the trial was registered at the Netherlands Trial Register (NTR2403). All patients signed an informed consent.

Interventions
Experienced musculoskeletal radiologists of the Zuyderland Medical Centre in Sittard-Geleen performed all ultrasound examinations using a standardised protocol.[32] The two study arms were treatment tailored to the ultrasound diagnosis or usual care (figure 1). The GP provided the allocated treatments.

Ultrasound tailored treatment
A key feature of this intervention was disclosure of the ultrasound diagnosis to the GP in order to tailor treatment. GPs treated patients according to the advised evidence-based, tailored treatment steps as presented in figure 1 and in detail published in the study protocol.[32] Advised treatment modalities depending on the ultrasound diagnosis were subacromial corticosteroid injections in case of bursitis or calcific tendonitis, referral to a physiotherapist in case of tendinopathy or partial-thickness tendon tear and an orthopaedic surgeon in case of full-thickness tendon tears.

In case there were no abnormal ultrasound findings, usual care according to the guideline for shoulder pain was provided. In cases where multiple ultrasound diagnoses were present, the most relevant diagnosis was selected on the basis of the clinical findings. Motivated and within the recommendations made in the guideline for shoulder pain, GPs were allowed to deviate from the advised treatment steps.

Usual care
In the control group, the ultrasound diagnosis was not disclosed; therefore, usual care according to the guidelines for shoulder pain was applied. It consisted of a pragmatic, stepwise approach; a wait-and-see policy with
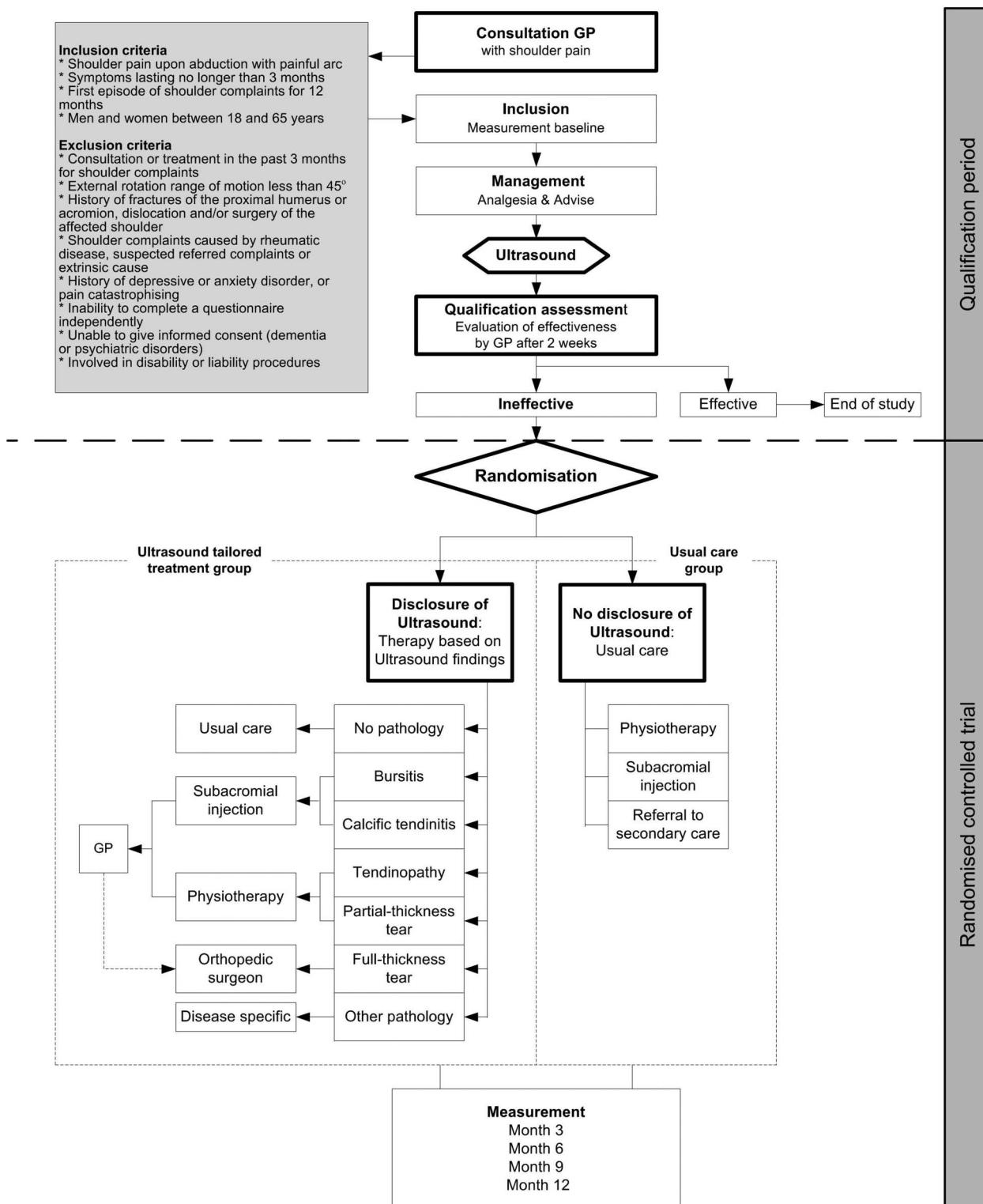
Figure 1 Flow chart of the study. GP, general practitioner.

advice and analgesia for another 2 weeks; corticosteroid injections and referral to a physiotherapist were advised as options in persisting cases, depending on the level of pain and functional limitations, respectively; referral to a hospital specialist was advised if conservative treatment failed.[10]

Randomisation and blinding
Based on a qualification assessment at 2 weeks, unrecovered patients were randomly assigned by central block randomisation (blocks of 4) to one of the study arms after stratification for age (cut-off ≥50 years), using an online application developed at the centre for data and

information management of Maastricht University. Recovery was measured by the Global Perceived Effect questionnaire (see outcomes).[34] Neither the patient nor the GP could be blinded for allocated treatment. However, ultrasound diagnoses were only disclosed to GPs of those patients in the ultrasound tailored treatment group. Radiologists were not allowed to communicate with the patient about the ultrasound imaging results.

Outcomes

Follow-up was performed by postal questionnaires at baseline, 3, 6, 9 and 12 months. Primary outcome was patient-perceived recovery using the Global Perceived Effect questionnaire at 1 year. It consists of a one-item score concerning recovery following treatment, measured on a seven-point ordinal scale. Patients were considered recovered when they reported to be much improved or fully recovered.[34] Secondary outcomes included experienced shoulder pain using the Shoulder Pain Score,[35] performance of daily activities using the Shoulder Disability Questionnaire[36] and health-related quality of life using the Euroqol five-item quality of life questionnaire (EQ-5D).[37]

The SPS questionnaire consists of six pain symptom questions and a 10-point scale.[35] The SPS has been proved to be a useful instrument for following the course of the disorder over time and gives an indication when a patient feels cured. The score can range from 7 to 28 with a higher score indicating more pain. The SDQ contains 16 questions and is a useful discriminative instrument, especially in the primary care setting.[36] The SDQ score can range from 0 to 100 with a higher score indicating a more severe disability. The EQ-5D is one of the most used generic measures to quantify the health-related quality of life in participants with musculoskeletal disorders[38 39] and consists of two sections. The first section comprises five questions regarding five dimensions of health. Calculation of the index score was performed according to the British recommendations and ranges from −1 to 1, higher scores indicating better quality of life.[37] The second section is a visual analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).

At inclusion, patients of whom the GP doubted about pain catastrophising behaviour were given the Pain Catastrophizing Scale (PCS).[40] The PCS is a 13-item self-report scale to measure pain catastrophising and the score can range from 0 to 52 with a higher score indicating a higher level of catastrophising. We made use of a cut-off value of 20 points.[40] Participants were excluded by the research team if their score was >20.

To measure 1-year healthcare use, we obtained electronic patient records of all patients by contacting the GP and used the postal questionnaires. In case patients were referred to a physiotherapist, the number of visits was collected.

Sample size

We estimated a 20% improvement in recovery rate as the minimal clinical important difference, and therefore calculated our sample size on the ability to detect a difference in study arms of 20% (60% vs 80%) or more on recovery rate.[4] Using these data, we estimated that we needed 81 patients per study arm to show a significant difference at the 5% level (two-sided) with 80% power (based on the χ^2 test). Allowing for a 10% drop-out rate,[41 42] and the expectation that the qualification period would filter out 20% of the patients we needed to include 226 patients in total.

Statistical analysis

Data were analysed on an intention-to-treat basis. In addition, a per-protocol analysis of complete cases for the primary outcome at 12 months was performed to estimate the impact on the effect of loss to follow-up and protocol deviations. To handle single-item missing data, a sensitivity analysis based on best-case and worst case scenarios (highest and lowest scores) was performed. No multiple imputation was used if whole questionnaires were missing, since missing outcome data were dealt with using a likelihood-based approach, assuming missing at random.

For the primary outcome measure, analysis was performed using a logistic regression analysis with correction for the stratification variable age (cut-off at 50 years). In addition, a three-level logistic mixed model to correct and account for age (cut-off at 50 years) and variation at the level of the GP practice, patient and repeated observations with data from four time points (3, 6, 9 and 12 months) was used. Secondary outcomes were analysed using a linear mixed model correcting and accounted for age and variation at the level of GP practice, patient and repeated observations with data from four time points (0, 3, 6 and 12 months). Shoulder-related healthcare resource during 1-year follow-up was analysed using a logistic regression analysis with correction for the stratification variable age (cut-off at 50 years).

Putative prognostic factors at baseline, as well as recruitment strategy, were added as covariates to the mixed model. Owing to the expected small number of non-recovered patients, this correction was only applied to the numerical outcomes.

All analysis was performed in SPSS (V.21). A p value of ≤0.05 was considered statistically significant.

RESULTS

From November 2010 to December 2013, 129 patients were included in 26 GP practices. The flow of patients through the study is presented in figure 2. Eighteen patients (14%) recovered during the 2-week qualification period, resulting in 111 randomised patients; 56 were allocated to ultrasound tailored treatment and 55 to usual care. Ten patients were lost to follow-up at
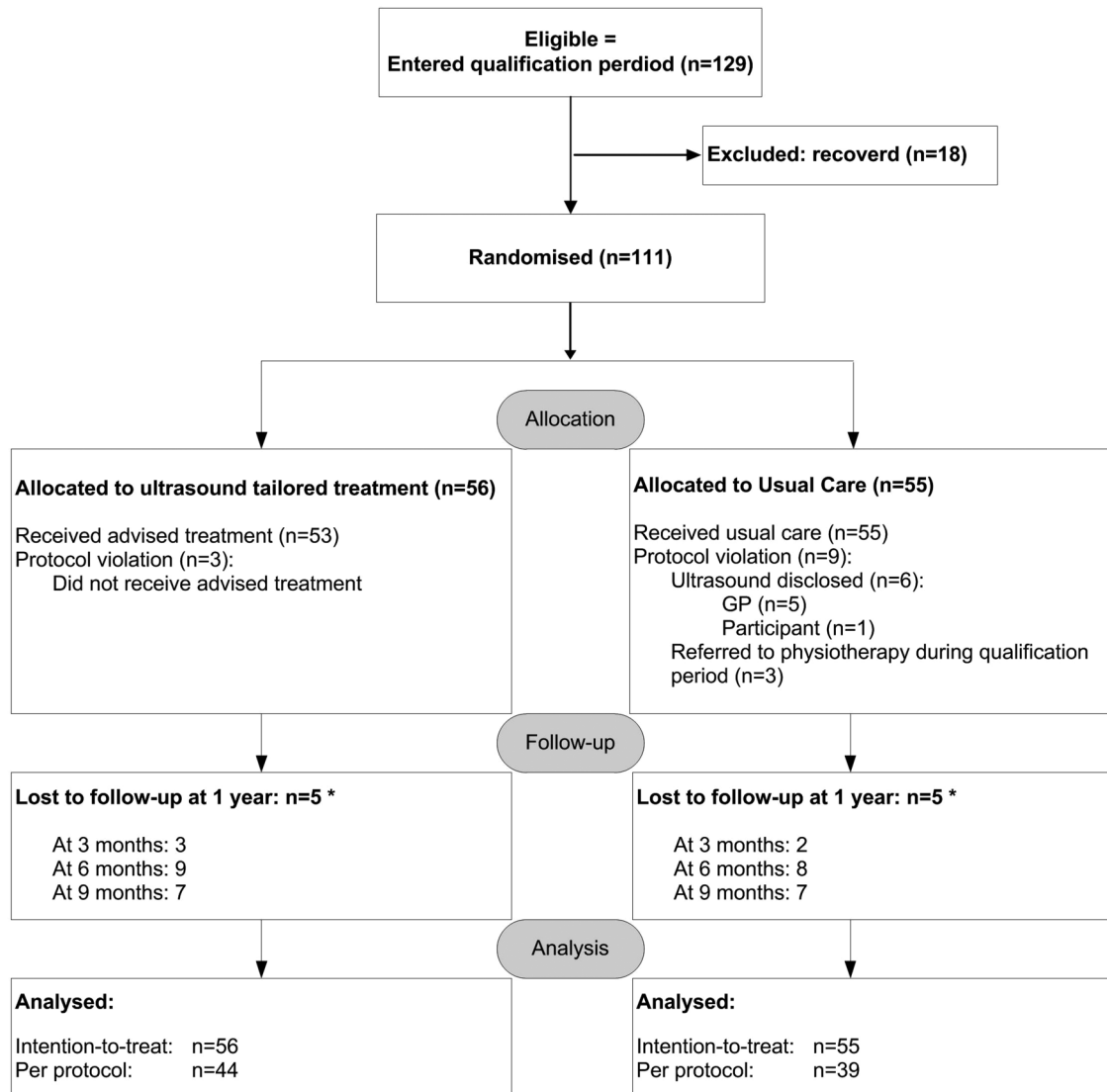
* Lost to follow-up was reported for each time point. Some participants missed 1 time point, but were followed up at a different time point

Figure 2 Flow of patients through the study. GP, general practitioner.

1 year (five in each randomised group). Baseline characteristics were similar in both randomised groups (table 1). Two patients were prompted by their GP to complete the Pain Catastrophizing Scale, and both were classified as not pain catastrophising.

Patient-perceived recovery

Although more patients in the ultrasound tailored treated group perceived to be recovered after 1 year according to the patients' Global Perceived Effect assessment (72.5% (37/51) vs 60% (30/50), OR 1.86 (95% CI 0.79 to 4.36)), this difference was not significant (p=0.15). The intention-to-treat analysis of this primary outcome measure is presented in table 2. Logistic mixed model analysis adjusting for age (cut-off at 50 years) and variation at the level of the GP practice, patient and

repeated observations with data from four time points (3, 6, 9 and 12 months) showed similar results (OR 2.24 (95% CI 0.72 to 6.89), p=0.16) for patient-perceived recovery after 1 year.

Secondary outcomes

After 1 year, the mean differences in the Shoulder Pain Score (1.7 points, 95% CI −3.9 to 0.5) and Shoulder Disability Questionnaire (6.9 points, 95% CI −19.9 to 6.1) were in favour of the ultrasound tailored treated group, but these differences were not significant (p=0.15 and 0.29, respectively). Also health-related quality of life, measured with the EQ-5D and EQ-5D Visual Analogue Scale, was not significantly different between groups (table 3).

Table 1 Baseline characteristics

Variables	Ultrasound tailored treatment (n=56)	Usual care (n=55)	Recovered before randomisation (n=18)
Demographic variables			
Age, mean (SD)	49.0 (9.9)	49.4 (10.9)	48.6 (12.4)
Female, n (%)	26 (46.4)	20 (36.4)	10 (55.6)
Specific disease variables			
Duration of pain in weeks (SD)	6.2 (3.8)	5.5 (3.5)	2.1 (1.8)
Acute onset, n (%)	24 (42.9)	28 (50.9)	13 (72.2)
Concomitant neck symptoms, n (%)	20 (35.7)	14 (25.5)	6 (33.3)
Dominant shoulder affected, n (%)	33 (58.9)	37 (67.3)	9 (50.0)
Ultrasound findings, n (%)			
Calcific tendonitis	26 (46.4)	28 (50.9)	11 (61.1)
Tendinopathy	16 (28.6)	16 (29.1)	5 (27.8)
Bursitis	14 (25.0)	10 (18.2)	2 (11.1)
Partial-thickness tears	11 (19.6)	13 (23.6)	1 (5.6)
Full-thickness tears	0 (0.0)	3 (5.5)	1 (5.6)
Impingement	12 (21.4)	7 (12.7)	2 (11.1)
Number of disorders, n (%)			
No disorder	12 (21.4)	8 (14.5)	4 (22.2)
1 disorder	20 (35.7)	26 (47.3)	7 (38.9)
≥2 disorders	24 (42.9)	21 (38.2)	7 (38.9)

Analyses adjusting for the putative prognostic factors did not alter the results. No significant differences were found in the proportion of patients referred to physiotherapy (59% of patients receiving ultrasound tailored treatment vs 64% of patients receiving usual care), their mean number of physiotherapy sessions (12.9 vs 12.4), corticosteroid injections or referrals to secondary care (39% vs 31% and 20% vs 13%). Shoulder-related healthcare use after 1 year between the two groups is summarised in table 4. No adverse events or side effects were reported.

Sensitivity analysis

Per-protocol analyses of the primary outcome measure and the sensitivity analyses with best-case and worst-case scenarios to handle single-item missing data of the secondary outcome measures produced similar effects as in the intention-to-treat analysis (table 5).

DISCUSSION
Summary

This study has shown no clinically significant difference in the primary outcome measure between the ultrasound tailored treatment and usual care groups. Furthermore, there was no overall difference in healthcare resources used between the groups. However, our trial does not give a conclusive answer to whether ultrasound tailored treatment improves outcome after 1 year, as our trial was under-enrolled. The limited size of this trial is a limitation that may have prevented the documentation of significant clinically important differences. Although no formal cost data are included, one can only assume that the ultrasound examinations are

Table 2 Primary outcome measure, global perceived effect

Recovered	Ultrasound tailored treatment %	Usual care %	Between-group difference						
			Diff. (%)	OR*	95% CI	p Value	OR†	95% CI	p Value
3 months	41.5 (22/53)	32.1 (17/53)	9.4	1.52	0.67 to 3.38	0.30	2.18	0.75 to 6.37	0.15
6 months	46.8 (22/47)	44.7 (21/47)	2.1	1.10	0.47 to 2.48	0.82	1.15	0.39 to 3.50	0.80
9 months	53.1 (26/49)	60.4 (29/48)	−7.3	0.75	0.33 to 1.73	0.50	0.59	0.20 to 1.80	0.35
12 months	72.5 (37/51)	60.0 (30/50)	12.5	1.86	0.79 to 4.36	0.15	2.24	0.72 to 6.89	0.16

*Based on logistic regression analysis correcting for age (stratification variable, cut-off 50 years).
†Based on logistic mixed model analysis correcting for age (stratification variable, cut-off 50 years) and variation at the level of the GP practice, patient and repeated observations. The level-3 variance (GP practice) was equal to 0. The interaction between group and time was not significant.
Diff., difference.

Table 3 Secondary outcome measures

| Outcome variable | Ultrasound tailored treatment | | Usual care | | Between-group difference | | | |
	N	Mean (SD)	N	Mean (SD)	Mean difference*	95% CI Lower	Upper	p Value
SPS								
Baseline	56	20.6 (4.1)	55	19.5 (4.3)				
3 months	53	14.3 (5.1)	53	15.2 (5.5)	−2.1	−4.2	0.1	0.06
6 months	47	13.1 (4.7)	47	13.5 (5.4)	−1.7	−3.8	0.5	0.13
12 months	51	11.1 (5.3)	50	11.5 (4.8)	−1.7	−3.9	0.5	0.12
SDQ								
Baseline	56	69.6 (17.3)	55	68.3 (17.2)				
3 months	53	47.2 (32.4)	53	49.4 (29.8)	−4.3	−14.9	6.3	0.42
6 months	47	40.7 (29.7)	47	36.3 (27.1)	3.2	−8.1	14.4	0.58
12 months	51	24.3 (30.2)	50	31.0 (29.7)	−6.9	−19.9	6.1	0.29
EQ-5D								
Baseline	56	0.68 (0.20)	55	0.73 (0.22)				
3 months	53	0.76 (0.21)	53	0.79 (0.22)	0.01	−0.07	0.09	0.74
6 months	47	0.83 (0.18)	47	0.84 (0.18)	0.05	−0.04	0.13	0.26
12 months	51	0.81 (0.27)	50	0.87 (0.16)	0.002	−0.10	0.10	0.97
EQ-5D VAS								
Baseline	56	69.4 (15.2)	55	69.9 (14.2)				
3 months	53	72.6 (17.6)	53	73.1 (13.7)	0.13	−5.16	4.90	0.96
6 months	47	78.4 (12.0)	47	76.5 (13.7)	2.82	−2.64	8.29	0.31
12 months	51	78.3 (15.8)	50	77.4 (14.5)	1.62	−4.76	8.00	0.62

*Linear mixed model analysis corrected for baseline and age (stratification variable, cut-off 50 years) and variation at the level of the GP practice, patient and repeated observations. The level-3 variance (GP practice) was equal to 0 for SPS and EQ-5D. The interaction between group and time was not significant.
EQ-5D VAS, Visual Analogue Scale 0–100 (100=best health status); EQ-5D, Euroqol five-item quality of life questionnaire tariff −1 to 1 (1=highest health-related quality of life); SDQ, Shoulder Disability Questionnaire 0–100 (100=most severe disability); SPS, Shoulder Pain Score 7–28 (28=most pain).

Table 4 Use of healthcare resources during 1-year follow-up

	Ultrasound tailored treatment n=54*	Usual care n=55	OR†	95% CI	p Value
GP re-consultation, n (%)	24 (43)	30 (55)	1.05	0.49 to 2.25	0.90
No. of re-consultations (mean, SD)	1.7 (0.9)	2.1 (1.3)			0.26
Diagnostic imaging ordered by GP, n (%)‡	56 (100)	13 (24)	0.58	0.17 to 2.04	0.40
– Plain radiography, n (%)	1 (2)	6 (11)	1.02	0.21 to 4.92	0.98
– Ultrasound imaging, n (%)‡	56 (100)	12 (22)	0.72	0.20 to 2.60	0.62
Physiotherapist referral, n (%)	33 (59)	35 (64)	0.83	0.38 to 1.81	0.64
No. of physiotherapy sessions (mean, SD)§	12.9 (9.0)	12.4 (10.2)			0.64
Medication used, n (%)	31 (55)	36 (65)	1.59	0.73 to 3.47	0.25
– GP prescription, n (%)	12 (21)	21 (38)	0.85	0.37 to 1.95	0.70
– OTC, n (%)	25 (45)	27 (49)	1.17	0.55 to 2.49	0.69
No. of GP prescriptions (mean, SD)	1.6 (1.2)	1.4 (0.7)			1.00
Corticosteroid injections by GP, n (%)	22 (39)	17 (31)	1.27	0.57 to 2.82	0.56
No. of injections (mean, SD)	1.4 (0.67)	1.5 (0.51)			0.50
Secondary care referral, n (%)	11 (20)	7 (13)	2.23	0.73 to 6.82	0.16
Surgery, n (%)	1 (2)	3 (5)	0.74	0.10 to 5.54	0.77

*One patient gave no consent to obtain her patient record, and one GP did not deliver the patient record.
†Logistic regression analysis corrected for age (stratification variable, cut-off 50 years).
‡Including intervention ultrasound.
§From five patients in each group, the number is lacking.
GP, general practitioner; OTC, over-the-counter medication.

Table 5 Secondary outcome measures: linear mixed model analysis adjusting for age or all putative prognostic factors with sensitivity analysis

| Outcome variable | 3 months | | 6 months | | 12 months | | Overall |
	Mean difference (95% CI)	p Value	Mean difference (95% CI)	p Value	Mean difference (95% CI)	p Value	p Value
SPS*							
ITT—all	−2.02 (−4.17 to 0.13)	0.065	−1.57 (−3.73 to 0.58)	0.155	−1.59 (−3.76 to 0.58)	0.149	0.316
PP—age 50	−1.73 (−3.97 to 0.52)	0.130	−1.30 (−3.52 to 0.93)	0.251	−1.38 (−3.70 to 0.93)	0.239	0.485
PP—all	−1.71 (−3.96 to 0.54)	0.133	−1.27 (−3.51 to 0.96)	0.259	−1.36 (−3.68 to 0.95)	0.245	0.494
Best case							
ITT—age 50	−2.05 (−4.15 to 0.06)	0.057	−1.71 (−3.84 to 0.43)	0.116	−1.70 (−3.85 to 0.45)	0.119	0.275
ITT—all	−2.02 (−4.13 to 0.10)	0.061	−1.64 (−3.79 to 0.51)	0.133	−1.62 (−3.79 to 0.54)	0.139	0.298
PP—age 50	−1.74 (−3.96 to 0.48)	0.123	−1.35 (−3.55 to 0.86)	0.228	−1.39 (−3.69 to 0.91)	0.233	0.468
PP—all	−1.74 (−3.96 to 0.48)	0.123	−1.35 (−3.55 to 0.86)	0.228	−1.39 (−3.69 to 0.91)	0.233	0.468
Worst case							
ITT—age 50	−1.88 (−3.99 to 0.23)	0.080	−1.89 (−4.05 to 0.27)	0.086	−1.75 (−3.89 to 0.40)	0.109	0.285
ITT—all	−1.85 (−3.96 to 0.27)	0.086	−1.82 (−3.99 to 0.35)	0.099	−1.67 (−3.83 to 0.49)	0.127	0.315
PP—age 50	−1.74 (−3.96 to 0.48)	0.123	−1.57 (−3.81 to 0.67)	0.166	−1.44 (−3.74 to 0.86)	0.217	0.436
PP—all	−1.74 (−3.96 to 0.48)	0.123	−1.57 (−3.81 to 0.67)	0.166	−1.44 (−3.75 to 0.86)	0.216	0.436
SDQ†							
ITT—all	−4.41 (−15.00 to 6.16)	0.412	3.03 (−8.24 to 14.31)	0.597	−6.86 (−19.81 to 6.08)	0.296	0.320
PP—age 50	−3.15 (−14.82 to 8.51)	0.595	3.12 (−8.78 to 15.03)	0.606	−6.57 (−20.31 to 7.16)	0.345	0.425
PP—all	−3.36 (−15.07 to 8.34)	0.572	2.94 (−8.96 to 14.84)	0.627	−6.59 (−20.16 to 6.98)	0.338	0.434
Best case							
ITT—age 50	−3.38 (−13.82 to 7.05)	0.524	2.94 (−8.35 to 14.24)	0.609	−5.24 (−18.20 to 7.72)	0.425	0.495
ITT—all	−3.24 (−13.72 to 7.24)	0.543	3.05 (−8.25 to 14.35)	0.596	−5.06 (−17.91 to 7.78)	0.437	0.507
PP—age 50	−2.56 (−14.15 to 9.02)	0.663	2.22 (−9.81 to 14.26)	0.716	−5.64 (−19.54 to 8.25)	0.422	0.618
PP—all	−2.56 (−14.23 to 9.09)	0.665	2.22 (−9.81 to 14.26)	0.716	−5.62 (19.26 to 8.01)	0.416	0.620
Worst case							
ITT—age 50	−3.75 (−14.25 to 6.74)	0.482	2.57 (−8.81 to 13.96)	0.657	−5.03 (−18.18 to 8.12)	0.450	0.531
ITT—all	−3.61 (−14.15 to 6.93)	0.501	2.69 (−8.71 to 14.08)	0.643	−4.85 (−17.88 to 8.18)	0.463	0.545
PP—age 50	−2.90 (−14.59 to 8.78)	0.625	1.90 (−10.27 to 14.07)	0.759	−5.43 (−19.55 to 8.69)	0.448	0.660
PP—all	−2.90 (−14.66 to 8.85)	0.627	1.90 (−10.26 to 14.06)	0.759	−5.41 (−19.26 to 8.43)	0.440	0.662
EQ-5D*							
ITT—all	0.01 (−0.07 to 0.09)	0.766	0.05 (−0.04 to 0.13)	0.284	−0.004 (−0.10 to 0.09)	0.931	0.615
PP—age 50	0.02 (−0.07 to 0.10)	0.715	0.07 (−0.03 to 0.16)	0.155	0.002 (−0.11 to 0.11)	0.976	0.410
PP—all	0.02 (−0.07 to 0.10)	0.715	0.07 (−0.03 to 0.16)	0.155	0.001 (−0.11 to 0.11)	0.992	0.407
EQ-5D VAS†							
ITT—all	−0.23 (−5.26 to 4.80)	0.929	2.68 (−2.79 to 8.14)	0.336	1.39 (−5.00 to 7.78)	0.667	0.714
PP—age 50	−1.30 (−7.07 to 4.47)	0.657	2.76 (−3.16 to 8.68)	0.359	2.09 (−4.53 to 8.71)	0.533	0.529
PP—all	−1.30 (−6.94 to 4.34)	0.650	2.76 (−3.14 to 8.66)	0.357	1.95 (−5.03 to 8.92)	0.580	0.536

*The level-3 variance (GP practice) was equal to 0. Final model was based on lowest Akaike's Information Criterion, which resulted in an unstructured covariance structure for repeated measures (level 1) and no random effects at levels 2 (patients) and 3 (GP practice).

†Best model (lowest Akaike's Information Criterion) was model with random intercept and random slope (time) on patient level and random intercept at GP practice level.

Age, stratification variable (cut-off 50 years); All, all putative prognostic factors; EQ-5D, Euroqol five-item quality-of-life questionnaire tariff; ITT, intention-to-treat analysis; PP, per-protocol analysis; SDQ, Shoulder Disability Questionnaire; SPS, Shoulder Pain Score; VAS, Visual Analogue scale.

additional costs for the intervention group, which cannot be justified in routine practice based on this trial. Based on this study, no change in current pragmatic guidelines to incorporate early ultrasound imaging can be recommended.

Strengths and limitations

To the best of our knowledge, this is the first primary care study investigating the effect of using ultrasound imaging in the management of patients with shoulder pain to target treatment to the specific underlying patho-anatomical disorders. This study has several strengths. First, this study was developed as a pragmatic trial to inform clinicians, guideline developers and policymakers to choose wisely between options for care.[33] Pragmatic trials measure effectiveness in routine clinical practice and reflect variations between patients that occur in real clinical practice improving generalisability.[31] For instance, our trial reflects variations between GPs and physiotherapists who applied the treatments, and between radiologists who performed the ultrasounds. Second, we filtered out patients with a favourable natural course (14%) during the 2-week qualification period to prevent non-responders in the course of the trial. Third, since blinding for ultrasound diagnosis in the usual care group had to be taken into account to prevent information bias, radiologists were not allowed to discuss their findings with patients. We evaluated this by asking patients one question in the questionnaire at 3 months; blinding of patients for the ultrasound diagnosis was violated only once. On the other hand, in five patients in the usual care group, the ultrasound diagnosis was disclosed to the GP by mistake. We incorporated these protocol violations in the per-protocol analysis. Finally, our study seems representative as the recovery rate of 60% in the usual care group corresponds to recovery rate used in the sample size calculation, which was based on previous studies.[4 5]

Our study knows some limitations related to under-enrollment. First, under-enrollment hampers the interpretation of the results. We found a non-statistically significant difference in favour of ultrasound tailored treatment. However, this lack of evidence does not necessarily mean there is no effect.[43] The original sample size was calculated on a between-group difference in recovery rate of 20% (80% vs 60% for ultrasound tailored treatment and usual care, respectively) after 1 year. However, we observed a non-significantly difference of 12.5% (72.5% vs 60%) in recovery rate. Based on a binominal distribution, we have calculated that there is a chance of 10.6% to observe a difference in effect of 20%, if we had reached the target number of patients. Recalculation of the between-group difference for a reduced sample size of 111 randomised patients indicated that a recovery rate of ~23% was necessary to reach the level of significance. Very limited literature is available on the minimal clinically important difference for the Global Perceived Effect questionnaire. We

choose this outcome measure as it provides the patient's perspective on the impact of disorder and treatment.[44] One can argue that a difference of >10% is clinically important, but population and context determine what is clinically important.[45] The observed difference of 12.5% would have been statistically significant if we had included 448 patients, far more than calculated and not feasible in this design. Second, although we carefully planned the recruitment aspects and GPs beforehand responded positively to the feasibility, overtime it became clear that the target number of patients would be difficult to achieve within 2 years of recruitment. To facilitate greater patient inclusion, we relocated the inclusion procedure from the GP practice to the Department of Radiology at the Zuyderland Medical Centre. GPs were asked to refer eligible patients to the radiology department for inclusion and ultrasound imaging of the shoulder. This provided two advantages: all GPs working in the region of the Zuyderland Medical Centre were capable of referring patients for inclusion, and time spent by GPs on patient recruitment decreased to a minimum. This relocation was combined with the extension of the recruitment period with an additional year. This adjusted inclusion procedure yielded 42 additional patients. As this recruitment strategy might have induced a form of selection bias, it was added as a covariate to the mixed model analysis.

Another limitation is that we are unaware of the number of eligible patients as GPs rarely registered numbers and reasons for non-participation. However, this reflects the consultation process in general practice, where recruitment is rarely straightforward, but might influence the external validity of our findings.[46]

Comparison with existing literature

Previous studies, in general practice, focused on the effectiveness of treatments in patients with subacromial impingement syndrome, for example, the effectiveness of corticosteroid injections, exercise or manual therapy.[47–49] Instead of combining a diagnostic imaging test to inform GPs about underlying patho-anatomical disorders to tailor treatment to the observed underlying disorder, all these studies used the generic term subacromial impingement syndrome. We chose to label and define the shoulder disorders based on findings from physical examination followed by ultrasound imaging instead of solely findings from physical examination, as this often poorly correlates with the underlying disorder.[12–14] This lack in uniformity in the way shoulder disorders are labelled and defined hampers comparisons.[50–52] Therefore, content and effectiveness of this study cannot be compared to other randomised studies.

Implications for practice and research

Our findings do not support the ordering of ultrasound imaging at the initial visit of patients in whom the GP suspects a subacromial disorder. The current shoulder pain guidelines state that ultrasound imaging should be

considered in cases with an abnormal course or therapy resistance, an advice that still holds.[10] [15] As our study was under-enrolled, larger studies are needed to provide conclusive evidence.

GPs generally tend to ignore recommendations on additional imaging in guidelines as ultrasound imaging is frequently applied in patients with acute shoulder pain.[21] [26] [53] If GPs still consider ultrasound imaging, it might be used preferentially in patients aged 40 years or older, as over 90% of shoulder pain patients in this age group have a subacromial disorder, compared to a much lower prevalence in younger age groups.[25]

In patients with persistent pain, ultrasound imaging may yield a more specific diagnosis, provide a rationale for further treatment and inform patients about the prognosis of their disorder. For daily practice, it is important that ultrasound findings are considered in the clinical context, as asymptomatic findings may be detected. A situation that can pose a real challenge.

The question whether ultrasound tailored treatment is cost-effective has not been answered. Although not calculated, healthcare costs in this study probably do not differ between groups, as mean healthcare use is similar in both study groups. However, it is suggested that episodes of shoulder pain generate additional costs for productivity losses due to sick leave.[54] In an economic evaluation, these costs should be incorporated.

We showed that most patients were diagnosed with calcific tendonitis; therefore, attention can be focused predominantly on this disorder. There is growing evidence that extracorporeal shock wave therapy and barbotage are effective treatment options for calcific tendonitis, and that corticosteroid injections can be harmful,[55–58] which emphasises the importance of evaluating these treatment options in general practice.

Author affiliations
[1]Department of Family Medicine, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, The Netherlands
[2]Department of Methodology and Statistics, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, The Netherlands
[3]Department of Radiology, Maastricht University Medical Centre, Maastricht, The Netherlands
[4]Department of Epidemiology, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, The Netherlands

REFERENCES

1. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003;102:167–78.
2. Linsell L, Dawson J, Zondervan K, *et al*. Prevalence and incidence of adults consulting for shoulder conditions in UK primary care; patterns of diagnosis and referral. *Rheumatology (Oxford)* 2006;45:215–21.
3. Murphy RJ, Carr AJ. Shoulder pain. *BMJ Clin Evid* 2010;2010:1107.
4. van der Windt DA, Koes BW, Boeke AJ, *et al*. Shoulder disorders in general practice: prognostic indicators of outcome. *Br J Gen Pract* 1996;46:519–23.
5. Winters JC, Sobel JS, Groenier KH, *et al*. The long-term course of shoulder complaints: a prospective study in general practice. *Rheumatology (Oxford)* 1999;38:160–3.
6. Kuijpers T, van der Windt DA, Boeke AJ, *et al*. Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain* 2006;120:276–85.
7. Andersson HI. The course of non-malignant chronic pain: a 12-year follow-up of a cohort from the general population. *Eur J Pain* 2004;8:47–53.
8. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. *J Rheumatol* 1998;25:1612–15.
9. Luime JJ, Kuiper JI, Koes BW, *et al*. Work-related risk factors for the incidence and recurrence of shoulder and neck complaints among nursing-home and elderly-care workers. *Scand J Work Environ Health* 2004;30:279–86.
10. Winters JC, van der Windt DAWM, Spinnewijn WEM, *et al*. Shoulder pain guideline of the Dutch College of General Practitioners (in Dutch). *Huisarts Wet* 2008;51:555–65.
11. Bigliani LU, Levine WN. Subacromial impingement syndrome. *J Bone Joint Surg Am* 1997;79:1854–68.
12. Dinnes J, Loveman E, McIntyre L, *et al*. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. *Health Technol Assess* 2003;7:1–166.
13. Hegedus EJ, Goode A, Campbell S, *et al*. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med* 2008;42:80–92.
14. Hermans J, Luime JJ, Meuffels DE, *et al*. Does this patient with shoulder pain have rotator cuff disease? The rational clinical examination systematic review. *JAMA* 2013;310:837–47.
15. NHS/NICE Clinical Knowledge Summaries. Shoulder pain guidelines. http://www.cks.nhs.uk/shoulder_pain (accessed Apr 2016).
16. Britt H, Miller GC, Knox S. Shoulder syndrome. In imaging orders by general practitioners in Australia 1999–2000. AIHW Cat. No. GEP 7. General Practice Series No. 7. http://www.aihw.gov.au/publications/index.cfm/title/6949 (accessed Apr 2016).
17. Broadhurst NA, Gialamas A, McElroy HJ, *et al*. How do Australian GPs manage shoulder dysfunction? *Aust Fam Physician* 2004;33:861–4.
18. Buchbinder R, Staples MP, Shanahan EM, *et al*. General practitioner management of shoulder pain in comparison with rheumatologist expectation of care and best evidence: an Australian national survey. *PLoS ONE* 2013;8:e61243.
19. Loebenberg MI, Rosen JE, Ishak C, *et al*. A survey of decision-making processes in the treatment of common shoulder ailments among primary care physicians. *Bull Hosp Jt Dis* 2006;63:137–44.

20. Masters S, O'Doherty L, Mitchell GK, *et al*. Acute shoulder pain in primary care—an observational study. *Aust Fam Physician* 2007;36:473–6.

21. Broadhurst N, Baghurst T, MacLaren S. Ultrasound imaging for shoulder pain in general practice. *Aust Fam Physician* 2004;33:668–9.

22. Ottenheijm RPG, Jansen MJ, Staal JB, *et al*. Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. *Arch Phys Med Rehabil* 2010;91:1616–25.

23. Smith TO, Back T, Toms AP, *et al*. Diagnostic accuracy of ultrasound for rotator cuff tears in adults: a systematic review and meta-analysis. *Clin Radiol* 2011;66:1036–48.

24. Lenza M, Buchbinder R, Takwoingi Y, *et al*. Magnetic resonance imaging, magnetic resonance arthrography and ultrasonography for assessing rotator cuff tears in people with shoulder pain for whom surgery is being considered. *Cochrane Database Syst Rev* 2013; CD009020.

25. Ottenheijm RP, Cals JW, Weijers R, *et al*. Ultrasound imaging for tailored treatment of patients with acute shoulder pain. *Ann Fam Med* 2015;13:53–5.

26. Ottenheijm RP, van't Klooster IG, Starmans LM, *et al*. Ultrasound-diagnosed disorders in shoulder patients in daily general practice: a retrospective observational study. *BMC Fam Pract* 2014;15:115.

27. Cadogan A, Laslett M, Hing WA, *et al*. A prospective study of shoulder pain in primary care: prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskelet Disord* 2011;12:119.

28. Baring T, Emery RJH, Reilly P. Management of rotator cuff disease: specific treatment for specific disorders. *Best Prac Res Clin Rheumatol* 2007;21:279–94.

29. Royal Dutch Society for Physical Therapy. Diagnosis and treatment of subacromial disorders (in Dutch). http://www.kngfrichtlijnen.nl (accessed Apr 2016).

30. Hill JC, Whitehurst DG, Lewis M, *et al*. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet* 2011;378:1560–71.

31. Roland M, Torgerson DJ. What are pragmatic trials? *BMJ* 1998;316:285.

32. Ottenheijm RP, Joore MA, Walenkamp GH, *et al*. The Maastricht Ultrasound Shoulder pain Trial (MUST): ultrasound imaging as a diagnostic triage tool to improve management of patients with non-chronic shoulder pain in primary care. *BMC Musculoskeletal Disord* 2011;12:154.

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. Kamper SJ, Ostelo RW, Knol DL, *et al*. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol* 2010;63:760–6.

35. Winters JC, Sobel JS, Groenier KH, *et al*. A Shoulder Pain Score: a comprehensive questionnaire for assessing pain in patients with shoulder complaints. *Scand J Rehabil Med* 1996;28:163–7.

36. de Winter AF, van der Heijden GJ, Scholten RJ, *et al*. The shoulder disability questionnaire differentiated well between high and low disability levels in patients in primary care, in a cross-sectional study. *J Clin Epidemiol* 2007;60:1156–63.

37. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.

38. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.

39. Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Ann Rheum Dis* 2004;63:723–9.

40. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psycholog Assess* 1995;7:524–32.

41. Bergman GJ, Winters JC, Groenier KH, *et al*. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. *Ann Intern Med* 2004;141:432–9.

42. Geraets JJ, Goossens ME, de Bruijn CP, *et al*. Cost-effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints. *Int J Technol Assess Health Care* 2006;22:76–83.

43. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995;311:485.

44. Revicki D, Hays RD, Cella D, *et al*. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.

45. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002;14:109–14.

46. Jenkinson CE, Winder RE, Sugg HV, *et al*. Why do GPs exclude patients from participating in research? An exploration of adherence to and divergence from trial criteria. *Fam Pract* 2014;31:364–70.

47. Crawshaw DP, Helliwell PS, Hensor EM, *et al*. Exercise therapy after corticosteroid injection for moderate to severe shoulder pain: large pragmatic randomised trial. *BMJ* 2010;340:c3037.

48. Penning LI, de Bie RA, Walenkamp GH. The effectiveness of injections of hyaluronic acid or corticosteroid in patients with subacromial impingement: a three-arm randomised controlled trial. *J Bone Joint Surg Br* 2012;94:1246–52.

49. Kromer TO, Tautenhahn UG, de Bie RA, *et al*. Effects of physiotherapy in patients with shoulder impingement syndrome: a systematic review of the literature. *J Rehabil Med* 2009;41:870–80.

50. Green S, Buchbinder R, Glazier R, *et al*. Systematic review of randomised controlled trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *BMJ* 1998;316:354–60.

51. Schellingerhout JM, Verhagen AP, Thomas S, *et al*. Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach. *Man Ther* 2008;13:478–83.

52. de Witte PB, de Groot JH, van Zwet EW, *et al*. Communication breakdown: clinicians disagree on subacromial impingement. *Med Biol Eng Comput* 2014;52:221–31.

53. Ottenheijm RP, Hesselmans NJ, Kemper A, *et al*. GPs' perspectives on the diagnostic work-up in patients with shoulder pain: a qualitative study. *J Eval Clin Pract* 2014;20:239–45.

54. Kuijpers T, van Tulder MW, van der Heijden GJMG, *et al*. Costs of shoulder pain in primary care consulters: a prospective cohort study in The Netherlands. *BMC Musculoskelet Disord* 2006;7:83.

55. Bannuru RR, Flavin NE, Vaysbrot E, *et al*. High-energy extracorporeal shock-wave therapy for treating chronic calcific tendinitis of the shoulder: a systematic review. *Ann Intern Med* 2014;160:542–9.

56. de Witte PB, Selten JW, Navas A, *et al*. Calcific tendinitis of the rotator cuff: a randomized controlled trial of ultrasound-guided needling and lavage versus subacromial corticosteroids. *Am J Sports Med* 2013;41:1665–73.

57. Dean BJ, Franklin SL, Murphy RJ, *et al*. Glucocorticoids induce specific ion-channel-mediated toxicity in human rotator cuff tendon: a mechanism underpinning the ultimately deleterious effect of steroid injection in tendinopathy? *Br J Sports Med* 2014;48:1620–6.

58. Verstraelen FU, In den Kleef NJ, Jansen L, *et al*. High-energy versus low-energy extracorporeal shock wave therapy for calcifying tendinitis of the shoulder: which is superior? A meta-analysis. *Clin Orthop Rel Res* 2014;472:2816–25.