Shoulder Pain in Primary Care Physiotherapy

Yasmaine Helga Jacques Marie Karel

Cover design: Ineke Goes
Layout and printed by: Optima Grafische Communicatie (www.ogc.nl)
ISBN/EAN: 978-94-6361-135-0

The Medical Ethics Committee of the Erasmus Medical Center in Rotterdam approved the study protocol (MEC-2011-414).
This study was financed by the SIA-RAAK grant and partly funded by a program grant of the Dutch Arthritis Foundation.

The printing of this thesis was financially supported by the Avans University of Applied Sciences, the European Society for Shoulder and Elbow Rehabilitation (EUSSER), the Erasmus Medical Center, the Scientific College Physical Therapy (WCF) of the Royal Dutch Society for Physical Therapy (KNGF) and the department of General Practice of the Erasmus Medical Centre.

Shoulder Pain in Primary Care Physiotherapy

Schouderpijn in de eerstelijns fysiotherapie praktijk

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam op gezag van de rector magnificus Prof. Dr. R.C.M.E. Engels en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 19 september 2018 om 11:30 uur door

Yasmaine Helga Jacques Marie Karel
geboren te Heerlen

Erasmus University Rotterdam

Promotiecommissie

Promotor:	Prof. Dr. B.W. Koes
Overige leden:	Prof. Dr. F.J.P.M. Huygen
	Prof. Dr. T.P.M. Vliet-Vlieland
	Prof. Dr. J.A.N. Verhaar
Copromotoren:	Dr. A.P. Verhagen
	Dr. G.G.M. Scholten-Peeters

CONTENT

Chapter 1

General Introduction

GENERAL INTRODUCTION

Shoulder pain

From all musculoskeletal disorders, shoulder pain is the third most common after low back- and neck pain in the general population [1]. Shoulder pain has a reported prevalence between 4.7 and 46.7% [2, 3, 4]. The difference in prevalence numbers might be attributed to the study settings or different definitions of shoulder pain. Shoulder pain can have a significant impact on patient health and can affect an individual's capacity to work and participate in social activities. The clinical course is unfavourable as it can persist for a long period of time whereas about 50% of the patients continue to have pain for over 6 months [5]. Musculoskeletal disorders are the second most costly health expenditure in the Netherlands [6]. Expenditures related to shoulder pain in primary care are estimated to be on average about 689 euros (for 6 months) on average per patient in 2003 [7].

Management in primary care

Most of patients with shoulder pain are managed in primary care [8]. According to the guideline of the Dutch College of General Practice (NHG) for general practitioners, the recommended treatment consists of providing information, lifestyle recommendations, prescriptions of (pain)medication and a possible referral to physiotherapy or a specialist in secondary care when conservative treatment fails [9]. A Dutch study showed that general practitioners refer about 38% of their shoulder patients of which 84% to physiotherapy and 16% to secondary care [8].

Diagnostic process in physiotherapy practice

Patients will visit their primary care physiotherapists, either through direct access or after referral by their general practitioner/medical specialist. The physiotherapist will gather information using history taking and start their clinical reasoning in order to determine the patient's problem. This clinical reasoning process is a continuous process of information gathering in order to generate an initial hypothesis. It is estimated that most patients (80-85%) with shoulder pain suffer from rotator cuff disease, otherwise called subacromial pain syndrome or subacromial impingement syndrome [10, 11, 12]. Research has shown that shoulder tests, regularly used in physiotherapy practice, do not lead to a valid patho-anatomical diagnosis and there is a lack of uniformity in these diagnosis in research and clinical practice [13, 14, 15]. Therefore, the term "non-specific shoulder pain" is often used, rather than a specific diagnostic label. Diagnosing patients with shoulder pain is complex. However, a clear working hypothesis/diagnosis as a starting point for physiotherapeutic management is important. With an accurate diagnosis, the patient has the best opportunity for a positive health outcome as the treatment can be better tailored [16].

Diagnostic imaging

Diagnostic imaging is commonly used for musculoskeletal disorders and is regarded as an important tool for the management of these conditions. For example, in the case of red flags in patients with low back pain or upper extremity disorders, diagnostic imaging can be used to identify specific pathology [17, 18]. Imaging usually only serves a purpose in the diagnosis of specific pathologies. Likewise, several studies conclude that routine imaging for patients with acute low back pain and knee pain is not indicated when looking at patient reported outcome measures, either due to asymptomatic findings or the absence of reassurance [19, 20, 21]. Diagnostic imaging in patients with shoulder pain is only recommended after ineffective treatment in primary care [9].

Recently, there has been an increase in the use of diagnostic ultrasound for musculoskeletal disorders in primary care [22, 23]. Diagnostic ultrasound is considered to be a safe, non-invasive and accessible method to visualize extra-articular lesions and could help the physiotherapist's in their diagnostic process [24, 25]. It could be a useful imaging method for patients with musculoskeletal disorders. Previous research showed that the interobserver reliability between experienced medical specialists (often radiologists) is good in patients with shoulder pain [26, 27, 28]. It might open subsequently the opportunity to tailor treatment [29]. Diagnostic ultrasound could even serve to monitor progress since 50% of newly symptomatic tears progress in size compared to 20% in asymptomatic tears [30, 31]. Whether the use of diagnostic ultrasound could lead to better treatment processes and improve recovery for patients with shoulder pain remains unknown. Contrary, the use of diagnostic imaging procedures could even lead to overdiagnosis and unnecessary referrals to secondary care when detecting asymptomatic findings, as pathology found does not always explain the complaints [32, 33]. For ultrasound operators, it is essential to realize the consequences of false positive or false negative results for patient expectations and health care costs. Only a small number of medical specialists report that they trust the ultrasound findings made by physiotherapists and general practitioners [34]. Consequently, the diagnostic ultrasound is commonly repeated in secondary care.

Prognosis

The natural course of shoulder pain is not favorable. Only between 25% and 50% of patients with shoulder pain report to be recovered after 6 months in primary care [35, 36]. Prognostic information is important because it may provide a greater knowledge of who is likely to recover, or who will or will not respond to physiotherapy. It ensures efficient use of resources since a subgroup of patients with chronic complaints could account for a large part of the total costs [37]. Furthermore, it can assist the clinical decision-making process. At the moment, we cannot reliably define subgroups based on traditional diagnostic labels and help the patient with their expectations on the course of their shoulder pain [15]. Previous studies showed that duration of complaints, lower disability scores, and

being younger are prognostic factors for recovery [38, 39]. What determines a prolonged course of complaints requires further investigation to determine whether improvements in diagnostic and prognostic processes may reduce recovery time.

Working alliance

An accurate patient history, physical examination and identification of prognostic factors seems to be important for establishing a targeted treatment plan. There should also be a mutual collaboration between the therapist and the patient that involves emotional bonding, and agreement on the tasks and goals of treatment [40]. Communication between the physiotherapist and the patient should be ongoing to monitor progress and address any issues that might aggravate physical or psychological symptoms. Shared decision making has become an important novel aspect in this the communication process [41, 42]. Shared-decision-making is a conjoint decision –making process in which the therapist and patient are actively involved in the treatment plan [43]. A good working alliance could strengthen the patient's participation in this shared decision-making process and compliance to treatment. Earlier studies have found a positive correlation between working alliance and treatment outcome [40, 44, 45]. Working alliance might therefore be an important prognostic factor for recovery in patients with shoulder pain.

Management in physiotherapy practice

Physiotherapy usually includes a range of different interventions like exercise therapy, stretching, advice, massage and/or electrotherapy aimed at controlling/relieving pain and improving function of the shoulder. The evidence statement for subacromial complaints of the Royal Dutch Association for Physiotherapists (KNGF) recommends exercise therapy with active movements of the glenohumeral and scapulothoracic joint when there is sufficient range of motion. Despite this evidence statement, physiotherapy treatment seems to be highly variable [46, 47]. Several studies have studied effects of different interventions for shoulder pain, however the heterogeneity of management protocols makes it difficult to follow guidelines [36, 48].

Objective of this thesis

In summary, current physiotherapy management in patients with shoulder pain is unknown. Additionally, little is known about the effect of diagnostic imaging procedures, especially diagnostic ultrasound, as a relatively new imaging procedure in primary care physiotherapy. The current evidence statement do not makes a recommendation on the use of diagnostic ultrasound. Due to the lack of reproducibility of traditional diagnostic labels, subgroups based on prognostic factors could help facilitate more appropriate treatment plans. Several prognostic factors have been described and it is believed that

diagnostic ultrasound and working alliance might also be potential prognostic factors for recovery.

Therefore, the main objectives of this thesis are (1) to describe current management in relation to diagnostic work-up (including the use of diagnostic ultrasound) and treatment strategies of physiotherapy care for patients with shoulder pain (2) to identify prognostic factors and develop a prognostic model (including the use of diagnostic ultrasound and working alliance) of recovery for patients with shoulder pain.

REFERENCES

1. Kooijman, M., Swinkels, I., Leemrijse, C., & Veenhof, C. (2009). *Jaarcijfers 2009 en trendcijfers 2003-2009 fysiotherapie*. Netherlands Institute for Health Services Research NIVEL. NIVEL.
2. Greving, K., Dorrestijn, O., Winters, J., Groenhof, F., van der Meer, K., Stevens, M., & Diercks, R. (2012). Incidence, prevalence, and consultation rates of shoulder complaints in general practice. *Scandinavian Journal or Rheumatology, 41*, 150-155.
3. Reilingh, M., Kuijpers, T., Tanja-Hafterkamp, A., & van der Windt, D. (2008). Course and prognosis of shoulder symptoms in general practice. *Rheumatology, 47*, 724-730.
4. Luime, J., Koes, B., Hendriksen, I., Burdorf, A., Verhagen, A., Miedema, H., & Verhaar, J. (2004). Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scandinavian Journal of Rheumatology, 33*(2), 73-81.
5. Kuijpers, T., van der Windt, D., Boeke, A., Twisk, J., Vergouwe, Y., Bouter, L., van der Heijden, G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. Pain, 120(3), 276-85
6. RIVM. (2011). *Health care costs in the Netherlands*. Opgeroepen op 2017, van www.kostenvanziekten.nl
7. Kuijpers, T., van Tulder, M., van der Heijden, G., Bouter, L., & van der Windt, D. (2006). Costs of shoulder pain in primary care consulters: a prospective cohort study in the Netherlands. *BMC Musculoskeletal Disorders, 7*(83).
8. Dorrestijn, O., Greving, K., van der Veen, W., van der Meer, K., Diercks, R., Winters, J., & Stevens, M. (2011). Patients with shoulder complaints in general practice: consumption of medical care. *Rheumatology, 50*(2), 389-95.
9. Winters, J., Van der Windt, D., Spinnewijn, W. D., Boeke, A., Feleus, A., & Geraets, J. (2008). NHG-Standaard Schouderklachten (Tweede herziening). . *Huisarts Wet, 51*(11), 555-565.
10. Östör, A., Richards, C., Prevost, A., Speed, C., & Hazleman, B. (2004). Diagnosis and relation to general health of shoulder disorders presenting in primary care. *Rheumatology, 44*(6), 800-805.
11. Browning, D., & Desai, M. (2004). Rotator cuff injuries and treatment. *Primary Care, 31*, 807-29.
12. Hanchard, N., Lenza, M., Handoll, H., & Takwoingi, Y. (2013). Physical tests fpr shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *Cochrane Database Syst Rev, 4*.
13. Hughes, P., Taylor, N., & Green, R. (2008). Most clinical tests cannot accurately diagnose rotator cuff pathology: a systemiatic review. *Aus J Physiother, 54*, 159-170.
14. Hegedus, E., Goode, A., Campbell, S., Morin, A., Tamaddoni, M. M., & Cook, C. (2008). Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med, 42*(2), 80-92.
15. Schellingerhout, J., Verhagen, A., Thomas, S., & Koes, B. (2008). Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach. *Man Ther, 13*(6), 478-83.
16. European Society of Radiology. (2015). Medical imaging in personalised medicine: a white paper of the research committee of the European Society of Radiology (ESR). *Insights Imaging, 6*(2), 141-155.
17. Bussières, A., Peterson, C., & Taylor, J. (2008). Diagnostic imaging guideline for musculoskeletal complaints in adults-an evidence-based approach-part 2: upper extremity disorders. *J Manipulative Physiol Ther., 31*(1), 2-32.

18. Bussières, A., Taylor, J., & Peterson, C. (2008). Diagnostic imaging practice guidelines for musculoskeletal complaints in adults-an evidence-based approach-part 3: spinal disorders. *J Manipulative Physiol Ther., 31*(1), 33-81.
19. van Tulder, M., Tuut, M., Pennick, V., Bombardier, C., & Assendelft, W. (2004). Quality of primary care guidelines for acute low back pain. *Spine, 29*(17), E357-62.
20. Oei, E., Nikken, J., Ginai, A., Krestin, G., Verhaar, J., van Vugt, A., & Hunink, M. (2009). Costs and effectiveness of a brief MRI examination of patients with acute knee injury. *Eur Radiol., 19*(2), 409-18.
21. Bryan, S., Weatherburn, G., Bungay, H., Hatrick, C., Salas, C., Parry, D., et.al. (2001). The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint. *Health Technol Assess., 5*(27), 1-95.
22. McKiernan, S., Chiarelli, P., & Warren-Forward, H. (2010). Diagnostic ultrasound use in physiotherapy, emergency medicine, and anaestesiology. *Radiography, 16*, 154-9.
23. Potter, C., Cairns, M., & Stokes, M. (2012). Use of ultrasound imaging by physiotherapists: a pilot study to survey use, skills and training. *Manual Therapy, 17*, 39-46.
24. Rutten, M., Jager, G., & Kiemeney, L. (2010). Ultrasound detection of rotator cuff Tears:Observer Agreement related to increasing experience. *AJR Am J Roentgenol, 195*(6), W440-W446.
25. de Jesus, J., Parker, L., Frangos, A., & Nazarian, L. (2009). Accuracy of MRI, MR arthrography, and ultrasound in the diagnosis of rotator cuff tears: a meta-analysis. *192*(6), 1701-1707.
26. O'Connor, P., Rankine, J., Gibbon, W., Richardson, A., Winter, F., & Miller, J. (2005). Interobserver variation in sonography of the painful shoulder. *J Clin Ultrasound, 33*(2), 53-6.
27. Ottenheijm, R., Jansen, M., Staal, J., van den Bruel, A., Weijers, R., de Bie, R., & Dinant, G. (2010). Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. *Arch Phys Rehabil, 91*(10), 1616-25.
28. Smith, T., Back, T., Toms, A., & Hing, C. (2011). Diagnostic accuracy of ultrasound for rotator cuff tears in adults: a systematic review and meta-analysis. *Clin Radiol, 66*, 1036-48.
29. Ottenheijm, R., van't Klooster, I., Starmans, L., Vanderdood, K., de Bie, R., Dinant, G., & Cals, J. (2014). Ultrasound-diagnosed disorders in shoulder patients in daily general practice: a retrospective observational study. *BMC Fam Prac, 10*(15), 115.
30. Yamaguchi, K., Tetro, M., Blam, O., Evanoff, B., Teefey, S., & Middleton, W. (2001). Natural history of asypmtomatic rotator cuff tears: A longitudinal analysis of asymptomatic tears detected sonography. *J Shoulder Elbow Surg*.
31. Mall, N., Kim, H., Keener, J., Steger-May, K., Teefey, S., Middleton, W., et.al. (2010). Symptomatic progression of Asymptomatic Rotator Cuff Tears: A Prospective Study of Clinical and Sonographic Variables. *Journalof Bone and Joint Surgery*.
32. Tempelhof, S., Rupp, S., & Seil, R. (1999). Age-related prevalence of rotator cuff tears in asymptomatic shoulders. *J Shoulder Elbow Surg, 8*, 296-9.
33. Schibany, N., Zehetgruber, H., Kainberger, F., Wurning, C., Ba-Ssalamah, A., Herneth, A., et.al. (2004). Rotator cuff tears in asymptomatic individuals: a clinical and ultrasonographic screening study. *European Journal of Rediology, 51*, 263-268.
34. Scholten-Peeters, G., Franken, N., Beumer, A., & Verhagen, A. (2014). The opinion and experiences of Dutch orthopedic surgeons and radiologists about diagnostic musculoskeletal ultrasound imaging in

primary care: A survey. *Manual Therapy, 19*(2), 109-113.

35. Feleus, A., Bierma-Zeinstra, S., Miedema, H., Verhagen, A., Nauta, A., Burdorf, A., et.al. (2007). Prognostic indicators for non-recovery of non-traumatic complaints at arm, neck and shoulder in general practice–6 months follow-up. *Rheumatology (Oxford)., 46*(1), 169-76.
36. van der Windt, D., Koes, B., Boeke, A., DeVille, W., Jong, d., BA, & Bouter, L. (1996). Shoulder disorders in general practice: prognostic indicators of outcome. *Br J Gen Pract, 46*, 519-523.
37. Virta, L., Joranger, P., Brox, J., & Eriksson, R. (2012). Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. *BMC Musculoskelet Disord., 13*, 17.
38. Kuijpers, T., van der Windt, D., Boeke, A., Twisk, J., Vergouwe, Y., Bouter, L., & van der Heijden, G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Oain, 120*(3), 276-85.
39. Chester, R., Shepstone, L., Daniell, H., Sweeting, D., Lewis, D., & Jerosch-Herold, C. (2013). Predicting resonse to physiotherapy treatment for musculoskeletal shoulder pain: a systematic review. *BMC Musculoskeletal Disorders, 14*.
40. Ferreira, P., Ferreira, M. M., Refshauge, K., Latimer, J., & Adams, R. (2013). The therapeutic alliance between clinicians and patients predicts outcome in chronic low back pain. *Physical Therapy, 93*, 470-8.
41. Fuertes JN, M. A. (2007). The physician-patient working alliance. *Patient Educ Couns, 66*(1), 29-36.
42. Charles, C., Whelan, T., & Gafni, A. (1999). What do we mean by partnership in making decisions about treatment? *BMJ, 319*(7212), 780-2.
43. Elwyn, G., Dehlendorf, C., Epstein, R., Marrin, K., White, J., & Frosch, D. (2014). Shared decision making and motivational interviewing: achieving patient-centered care across the spectrum of health care problems. *Ann. Fam. Med, 12*, 270-275.
44. Hall, A., Ferreria, P., Maher, C., Latimer, J., & Ferreira, M. (2010). The inlfuence of the therapist-patient relationship on treatment outcome in physical rehabilitation: A systematic review. *Physical Therapy, 90*(8), 1099-110.
45. Lakke, S., & Meerman, S. (2016). Does working alliance have an influence on pain and physical functioning in patients with chronic musculoskeletal pain; a systematic review. *Journal of Compassionate Health Care, 3*(1).
46. Littlewood, C., Lowe, S., & Moore, J. (2012). Rotator cuff disorders: survey of current UK physiotherapy practice. *Should Elb, 4*, 64-71.
47. Bernhardsson, S., Oberg, B., Johansson, K., Nilsen, P., & Larsson, M. (2015). Clinical practice in line with evidence? A survey among primary care physiotherapists in western Sweden. *Journal of Evaluation in Clinical Practice*.
48. Laslett, M., Steele, M., Hing, W., NcNair, P., & Cadogan, A. (2014). Shoulder pain patients in primary care - part 1: Clinical outcomes over 12 months following standardized diagnostic workup, corticosteroid injections, and community-based care. *J Rehabil Med, 46*, 898-907.

Chapter 2

Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: Design of a prospective cohort study

Published in BioMed Central Musculoskeletal Disorders

2013 (online open access)

Yasmaine Karel
Wendy Scholten-Peeters
Marloes Thoomes-de Graaf
Edwin Duijn
Ramon Ottenheijm
Maaike van den Borne
Bart Koes
Geert-Jan Dinant
Eric Tetteroo
Annechien Beumer
Joost van Broekhoven
Marcel Heijmans

ABSTRACT

Background

Shoulder pain is disabling and has a considerable socio-economic impact. Over 50% of patients presenting in primary care still have symptoms after 6 months; moreover, prognostic factors such as pain intensity, age, disability level and duration of complaints are associated with poor outcome. Most shoulder complaints in this group are categorized as non-specific. Musculoskeletal ultrasound might be a useful imaging method to detect subgroups of patients with subacromial disorders.

Aim

To present the design of a prospective cohort study evaluating the influence of known prognostic and possible prognostic factors, such as findings from musculoskeletal ultrasound outcome and working alliance, on the recovery of shoulder pain. Also, to assess the usual physiotherapy care for shoulder pain and examine the inter-rater reliability of musculoskeletal ultrasound between radiologists and physiotherapists for patients with shoulder pain.

Methods

A prospective cohort study including an inter-rater reliability study. Patients presenting in primary care physiotherapy practice with shoulder pain are enrolled. At baseline validated questionnaires are used to measure patient characteristics, disease-specific characteristics and social factors. Physical examination is performed according to the expertise of the physiotherapists. Follow-up measurements will be performed 6, 12 and 26 weeks after inclusion. Primary outcome measure is perceived recovery, measured on a 7-point Likert scale. Logistic regression analysis will be used to evaluate the association between prognostic factors and recovery.

Discussion

The ShoCoDiP (Shoulder Complaints and using Diagnostic ultrasound in Physiotherapy practice) cohort study will provide information on current management of patients with shoulder pain in primary care, provide data to develop a prediction model for shoulder pain in primary care and to evaluate whether musculoskeletal ultrasound can improve prognosis.

BACKGROUND

This paper describes the ShoCoDiP (Shoulder Complaints and using Diagnostic Ultrasound in Physiotherapy practice) cohort study. Publishing the design of a study provides insight into the objectives and procedures before publishing the results. It may also protect against (subconscious) selective outcome reporting. Shoulder disorders are the second most common musculoskeletal complaint in the general population with a point prevalence of 20.6% [1] and cause considerable functional disability, pain and healthcare costs [2]. The reported 12-month prevalence for shoulder disorders is 6.7 to 66.7% [3]. In the Netherlands, the annual incidence in general practice is 15-16/1000 person-years [4]. About 30-40% of the patients with shoulder pain consult a general practitioner (GP) due to these complaints [1]. Chronicity and recurrence of shoulder pain are common [5-7]. About 40% of the patients still experience pain after 12 months [6] and 40% re-consult their GP [2]. There is strong evidence that prognostic factors for shoulder pain such as age, high disability scores, duration of shoulder pain and pain intensity are associated with poor outcome [4,8,9]. Having a specific diagnosis like bursitis, rotator cuff tear and frozen shoulder is reported to be a predictor for increased recovery in patients with upper extremity disorders compared to patients with a non-specific diagnosis in general practice [8].

At first consultation GPs recommend a ´wait and see´ policy in about 40% of the patients, 39% receive oral NSAIDs and 16% are referred for physiotherapy [10]. Early treatment in general practice mainly consists of pain medication and advice [2]. The guideline for shoulder pain of the Dutch College of General Practitioners advises a referral for physiotherapy or a corticosteroid injection as a standard procedure in shoulder pain when these complaints are present for ≥ 2 weeks [2]. In the Netherlands, since 2006 patients can directly access physiotherapy care which means they do not need a referral to consult a physiotherapist (PT). Nevertheless, the Dutch institute for paramedical care reported that in 2009 49% of the patients who visited the PT were referred by their GP, 38% used self-referral, and the remaining 13% were referred by a medical specialist [11].

In primary care, the information gained during history taking and physical examination is used to make a diagnosis and decide on treatment options. Unfortunately, physical examination is not always a reliable or valid diagnostic tool [12-14]. As a result, most complaints are regarded as non-specific, because no specific pathology can be diagnosed. When additional diagnostic information is needed, GPs can refer patients to radiologists for further diagnostic imaging, such as musculoskeletal ultrasound (MSU).

Nowadays, in the Netherlands many PTs attend additional courses on MSU, which can be a reliable and relatively inexpensive tool for the diagnosis of patients with shoulder pain [15, 16]. A recent systematic review shows that MSU has a sensitivity of 95% and a specificity of 96% for full thickness rotator cuff tears, and a sensitivity of 72%

and specificity of 93% for partial thickness tears when performed by an experienced radiologist [17]. Therefore, MSU performed by an experienced examinator might help in accurately diagnosing rotator cuff tears. Knowing this, the question remains whether or not patients will respond better to treatment once this pathology is identified in primary care. An accurate diagnosis is essential to ensure that patients receive appropriate treatment and correct information about their prognosis. Apart from the proposed treatment, the prognosis can be influenced by the patient's experience in the perceived health care or acquired treatment goals. This involves a therapeutic encounter between the patient and PT, hereafter referred to as 'working alliance'. A recent systematic review indicated that working alliance has a consistent positive correlation with treatment outcome in a physical rehabilitation setting [18]. The present study will evaluate whether working alliance and pathology detected on MSU are possible prognostic factors in primary care patients with shoulder pain.

MSU used by PTs probably could help to identify subgroup of patients who might better respond to physiotherapy treatment. We assume that a more specific diagnosis will lead to more specific treatment choices and better patient prognosis. Classifying these shoulder disorders seems to be a diagnostic challenge and therefore a shift from diagnostic research to prognostic research might help in the first steps of consultation [19].

The primary aim of the ShoCoDiP study is to evaluate physiotherapy care and prognostic factors in patients with shoulder pain and investigate whether MSU and the working alliance are related to patient recovery. Secondary aims are to assess the inter-rater reliability of MSU between PTs and radiologists, and whether patient characteristics of those who receive MSU differ from those who do not receive MSU.

METHODS

Design

A prospective cohort study including patients with shoulder pain presenting in primary care physiotherapy (Figure 1). Furthermore, a nested case cohort design will be used to evaluate whether patient characteristics differ between patients who do and do not receive MSU (Figure 1). The control group will be randomly selected from the total cohort. These patients are matched to patients who received MSU, based on the PT's decision, by age and sex. Patients who received MSU via the PT are also scanned by a radiologist to evaluate the inter-rater reliability. The Medical Ethics Committee of the Erasmus Medical Center in Rotterdam approved the study protocol (MEC-2011-414).

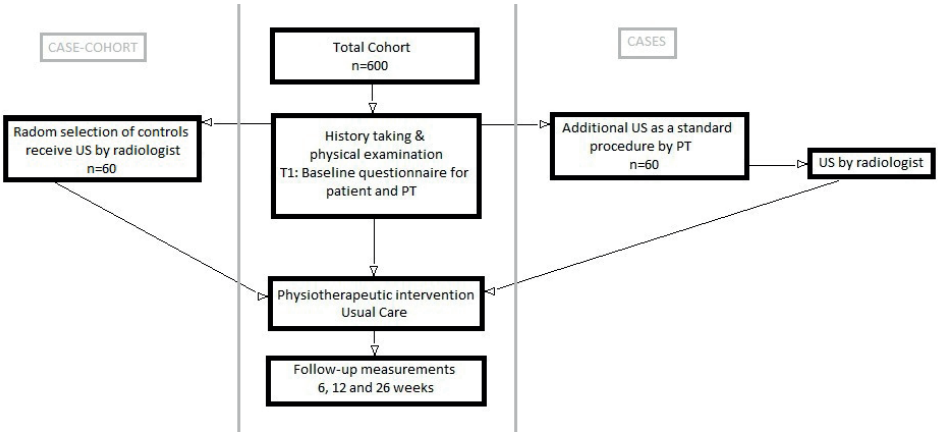


Figure 1. Flow chart of the study protocol.

Recruitment of PT, radiologists and patients

Physiotherapists (PTs)

PTs in the southwest region of the Netherlands will be asked to participate in the study. An introductory meeting was organized to explain study procedures and data collection. Selection criteria for PTs using MSU are: 1)PTs having ≥ 1 year of experience after their MSU course, 2) PTs performed ≥ 100 MSU examinations of the shoulder, 3) the transducer should have a minimum frequency of 7.5 MHz, and 4) having appropriate software (beamforming technology). These PTs were trained to use the MSU protocol by Jacobson [20] during a special consensus meeting.

Radiologists

Radiologists in the southwest region of the Netherlands are invited by telephone and email. Only radiologists who are specialized in musculoskeletal radiology and perform MSU in their hospitals are invited to participate. A total of 9 radiologists from 4 hospitals participate in the study. One of the researchers visits to inform them about the study procedures and the MSU protocol as described by Jacobson [20].

Patients

From November 2011 to November 2012 PTs will recruit consecutive patients in primary care. Patients eligible for the study suffer from shoulder pain, are aged ≥ 18 years and have adequate understanding of the Dutch language. Patients are excluded if they have serious pathologies (infection, cancer or fracture), previous surgery of the shoulder in the last 12 months, or received diagnostic imaging techniques such as MSU, MRI or X-ray of the shoulder in the 3 months prior to start of the study. All patients provided written informed consent.

Data collection

Data will be collected using online Limeservice software and safely stored by both the investigators and the software holders. Patients will receive a digital questionnaire at baseline and at 6, 12 and 26 weeks after inclusion. PTs receive questionnaires at baseline and at 3, 6 and 12 weeks follow-up. Whenever a PT performs an MSU, within 1 week the patient will undergo a second MSU by a blinded radiologist. To reduce the chance of missing data, an email reminder will be sent at 2 and 5 days to the patient or PT. Newsletters will be sent every month to the participating PTs to encourage adherance to the study. Moreover, all PTs will be contacted by telephone every 3 months to ensure adherence to the study protocol, and stimulate them to recruit eligible patients.

Baseline assessment

Patient characteristics, prognostic factors and disease-specific information will be collected at baseline. These include demographic variables and complaint-specific variables. PTs will report data on physical examination and their interpretation after history taking and physical examination. Possible hypotheses are described in Table 1.

Table 1. Hypotheses
Hypotheses are build and edited by the clinical opinions of 5 PT's.

0	Possible sub-acromial impingement
1	Possible internal (posterior) impingement
2	Possible instability of the glenohumeral joint
3	Possible SLAP leasion
4	Possible biceps tendinopathy
5	Possible frozen shoulder/capsulitis
6	Possible disorder of cervic0-thoracic spinal column and adhering costea
7	Possible myofascial triggerpoint in neck and shoulder region
8	Possible disorder of the acromioclavicular/sternoclavicular joint
9	Possible hypertonia in neck/shoulder region
10	Possible strain or sprain in neck/shoulder region
11	Not possible to specify a clear hypothesis
12	Other non-specified

Prognostic factors

Possible prognostic factors on recovery for patients with shoulder pain are extracted from the literature [4, 21-24] and consist of pain intensity, duration of complaints, age, gender, disability, highest level of education, job description (physically heavy work, static repetitive work or work with awkward postures; yes/no), sick leave due to shoulder complaint (yes/no), and complaints worsen during work (yes/no). Also, exploratory MSU outcome

and the Dutch version of the working alliance inventory (WAV-12) will be assessed as possible prognostic factors as they might be related to patient recovery.

Physiotherapy management

Descriptive factors like the frequency of diagnostic hypotheses, the treatment period, costs, and treatment goals and related interventions in physiotherapy practice will be assessed in the PT questionnaire.

Sample size

Based on the literature about 40% of the patients with shoulder pain will recover within 6 months [7]. We will estimate to include 15 prognostic variables in our prognostic model. Based on the 1 in 10 rule of 10 events per variable, a total of 150 events are needed in the smallest outcome (recovered or not). Therefore, the total study population should include about 300 subjects. Adjusting for about 20% missing values, the total population will comprise a minimum of 400 subjects.

Outcome measures

Primary outcome

Our primary outcome is recovery measured with the Global Perceived Effect (GPE) scale [25] (Table 2). The GPE uses a 7-point Likert scale scoring whether the patient's condition has improved or deteriorated since the start of their physiotherapy treatment. This scale ranges from 'worse than ever' to 'fully recovered'. Patients are considered to be recovered when they score 'strongly improved' or 'completely recovered' [25].

Secondary outcome

Functional disability will be measured with the Dutch version of the Shoulder Disability Questionnaire (SDQ-NL). The SDQ has 16 items which are answered with either 'yes', 'no', or 'not applicable'. The score ranges from 0 to 100, with a high score indicating more functional disability. This questionnaire has good construct validity [23], and appears to be a useful discriminative instrument in primary care [26]. The Shoulder Pain Disability Index (SPADI) is measured in conjunction with the SDQ-NL to validate the SPADI questionnaire in Dutch. The SPADI has 8 questions designed to measure the degree of difficulty someone has with various activities of daily living that require the use of upper extremities. Internal consistency is good (Cronbach's alpha: 0.90). Test-retest reliability of the SPADI and the intraclass correlation for the disability subscale ranges from 0.57-0.84 [27].

Pain severity will be assessed with the Shoulder Pain Score (SPS); this instrument has 6 questions about pain symptoms experienced in the last 24 hours scored on a 4-point scale, and an 11-point Numeric Rating Scale. Internal consistency for the SPS is good (Cronbach's alpha: 0.82) [28].

Table 2. Baseline to follow-up measures.

	Baseline	T1: 3 weeks	T2: 6 weeks	T3: 12 weeks	T4: 6 months
In- en exclusion criteria	X				
Demographic data	X				
GPE		X	X	X	X
SPS	X	X	X	X	X
SDQ-NL	X	X	X	X	X
SPADI	X	X	X	X	X
EQ5D	X	X	X	X	X
WAV-12			X		
Medical Consumption	X		X	X	X
	Physiotherapist				
	Baseline	T1: 3-weeks	T2: 6 weeks	T3: 12 weeks	T4: 6 months
Interpretation from physical examination and patient history	X				
Change in treatment plan	X	X	X	X	
Treatment goals	X	X	X	X	
Number of treatments	X	X	X	X	

Legends: GPE: General Perceived Effect, SPS: Shoulder Pain Score, SDQ-NL: Dutch Shoulder Disability Questionnaire, SPADI: Shoulder Pain and Disability Index, EQ5D: Euroquol five-item quality of life questionnaire, WAV-12: Dutch Working Alliance Scale (Short Form).

Health-related quality of life will be measured using the Euroquol five-item quality of life questionnaire (EQ-5D) [29]. This questionnaire covers 5 dimensions of health, and a visual analogue scale ranging from 0-100. The five dimensions of health are: mobility, self-care, usual activities, complaints/discomfort and anxiety/depression. The patient can score three levels of severity in each dimension (1=no problem, 2=moderate problem, 3=severe problem). Scoring will be calculated according to the European guideline recommendations [30].

Working alliance will be measured with a Dutch version of the Working Alliance Inventory (WAV-12). The WAV-12 will be assessed after 6 weeks. This questionnaire has three subscales designed to assess three primary components of the working alliance: 1) how closely client and therapist agree on and are mutually engaged in the goals of treatment (Cronbach's alpha: 0.85), 2) how closely client and therapist agree on how to reach the treatment goals (Cronbach's alpha: 0.83), and 3) the degree of mutual trust, acceptance, and confidence between client and therapist. Patients score on a 5-point scale ranging from rarely to always [31, 32].

MSU will be standardized in terms of 11 primary outcome categories: 1) tendinopathy, 2) calcification, 3) full or 4) partial thickness tear, 5) Biceps tendon tear, 6) subacromial-subdeltoid bursitis, 7) subacromial impingement, 8) osteoarthritis of the

acriomio-clavicular joint, 9) cortical discontinuity of superior aspect of the acromion, 10) no specific pathology, or 11) other. In case a diagnosis in category 1-2 was made, it could be specified in the following diagnostic subgroups; supraspinatus, infraspinatus, teres minor or subscapularis and biceps tendon. For category 3-4 it could be specified in; supraspinatus, infraspinatus and teres minor or subscapularis tendon. This resulted in a total of 11 diagnostic categories (Table 3) [17].

Table 3. MSU outcome classification.

Pathology	Anatomical site
1. Tendinopathy	m. supraspinatus
	m.subscapularis
	m. infraspinatus
	m. teres minor
2. Calcification	m. supraspinatus
	m.subscapularis
	m. infraspinatus
	m. teres minor
3. Full-thickness tear	m. supraspinatus
	m.subscapularis
	m. infraspinatus
	m. teres minor
4. Partial-thickness tear	m. supraspinatus
	m.subscapularis
	m. infraspinatus
	m. teres minor
5. Biceps tendon tear	
6. Bursitis acromialis (>2mm low frequency)	
7. Subacromial impingement (with active abduction)	
8. Artritis or arthrosis	acriomio-clavicular joint
9. Cortical discontinuity	superior aspect of the acromion
10. No specific pathology	
11. Other non-specified	

Statistical analysis

Descriptive statistics, including frequencies for categorical variables and means with standard deviations (SD) for continuous variables, will be used to describe the characteristics of the patients, PTs and radiologists. We intend to develop a prognostic model using logistic regression analysis with recovery (GPE) after 6 months as the primary outcome. Missing values will be handled using multiple imputation techniques. All candidate

predictors will be included in our prognostic model. All assumptions (homogeneity of variance, independence-normality of residuals, linearity and multicollinearity) for building a regression model will be checked before model building. Internal validation of the model will be assessed by a bootstrap procedure (200 repetitions) to assess the accuracy of the regression analysis. The inter-rater reliability will be evaluated with a KAPPA statistic. Statistical analysis will be performed using SPSS 20.0. A p-value of >0.05will be considered statistically significant.

DISCUSSION

The proposed study will describe the current management of shoulder pain in primary care and will help to determine which factors can predict patient recovery in PT practice. This study is designed to include key methodological features in order to minimize bias. These features include sampling of a representative cohort from physiotherapy setting with a high rate of follow-up.

Based on the sample of patients that will be recruited from physiotherapy practices, we aim to produce a pragmatic prediction model for PTs in primary care.

Possible prognostic factors and confounders are selected based on previous research [4]. The selected population of PTs in primary care enables us to include possible additional predictors such as characteristics from the PT and ultrasonographer. All medical consumption besides physiotherapy will be registered during follow-up questionnaires. Completeness of data collection will be stimulated by means of email reminders.

Although we will select a heterogeneous group of patients with shoulder complaints, we stress two important exclusion criteria. The first is that patients who had surgery of the shoulder in the previous 12 months are excluded, since these patients seem to differ in pathology and prognosis. Excluding these patients will ensure a more valid prediction model. Secondly, we postulate that PTs base their diagnosis and interventions on imaging techniques that were performed in the past; moreover, in case of the inter-rater reliability study, this could threaten blinding because most patients know the results of diagnostic imaging. Therefore, this study also excludes patients who had imaging of the shoulder in the 3 months prior to the start of physiotherapy treatment. PTs will be instructed to act as usual and are not instructed to adhere to a specific intervention protocol. This study aims to report on usual care in physiotherapy practice and provide insight into the diagnostic and therapeutic management of patients. Because patients are selected in primary care physiotherapy, we assume that they will represent the usual population consulting the PT with shoulder pain.

Patients in the control group will be randomly matched (by age and sex) to patients that receive an MSU by their PT. To avoid disease progression bias, their second MSU

will be performed within 1 week; we do not expect that partial or full-thickness ruptures or calcifications will heal within 1 week. However, we cannot be certain that patient recovery is related to changes in patho-anatomical findings on MSU. Furthermore, the literature describes a high prevalence of rotator cuff tears in asymptomatic populations [33, 34]. Therefore, we cannot ensure that these pathologies found on MSU images cause symptoms or constraints in daily activities for patients.

Radiologists and PTs will be blinded to each other's findings. Moreover, they will be blinded to clinical information that was not intended to form part of the MSU assessment. Radiologists are instructed to keep the patient blinded from MSU outcome. Blinding will be evaluated in the follow-up questionnaire of the patient.

From previous research it is known that MSU is operator dependent [35]. PTs and radiologists are instructed to use a standardized scanning protocol [20], to ensure comparability in MSU procedures. Current management with MSU does not standardize pathology criteria. To assess the effect of current management of MSU in primary care we chose not to define criteria for pathology in this study. Nevertheless, we standardized possible outcome definitions for both the radiologist and PT in order to be able to categorize data.

We assume that inter-rater reliability between PT and radiologist might be influenced by the quality of ultrasound equipment and experience. Therefore, only equipment with transducer frequencies of at least 7.5 MHz will be used in physiotherapy practice and PTs should have at least 1 year of experience with ≥ 50 examinations of the shoulder.

Until now, reliability studies generally evaluated the inter-rater reliability between radiologists. However, PTs increasingly use MSU in daily practice and the reliability between different professions has not yet been evaluated.

It is hoped that this prospective cohort study will help improve the current management and prognosis of patients with shoulder pain.

REFERENCES

1. Picavet, H.S., Schouten, J.S. (2003). Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3) study. *Pain*, 102, 167–78.
2. Winters, J.C., van der Windt, D.A.W.M., Spinnewijn, W.E.M., De Jongh, A.C., Van der Heijden, G.J.M.G., Buis, P.A.J., Boeke, A.J.P., Feleus, A. & Geraets, J.J.X.R. (2008). Dutch College of General Practitioners: practice guideline for shoulder complaints, *Huisarts wet*, 222-231.
3. Luime, J.J., Koes, B.W., Hendriksen, I.J., Burdorf, A., Verhagen, A.P., Miedema, H.S. & Verhaar, J.A.N. (2004). Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scandinavian Journal of Rheumatology*, 33, 73-81.
4. Croft, P., Pope, D., Silman, A. (1996). The clinical course of shoulder pain: prospective cohort study in primary care. *British Medical Journal*, 313, 601-2.
5. Van der Windt, D.A.W.M., Koes, B.W.K., Boeke, A.J., Devillé, W., de Jong, B.A. & Bouter, L.M. (1996). Shoulder disorders in general practice: prognostic indicators of outcome. *British Journal of General Practice*, 46(410), 519-23.
6. Winters, J.C., Sobel, J.S., Groenier, K.H., Arendzen, J.H. & Meyboom-de Jong, B. (1999). The long-term course of shoulder complaints: a prospective study in general practice. *Rheumatology*, 38, 160-163.
7. Kuijpers, T., Van der Windt, D.A.W.M., Van der Heijden, G.J.M.G. & Bouter, L.M. (2004). Systematic review of prognostic cohort studies in shoulder disorders. *Pain*, 109, 420-431.
8. Feleus, A., Bierma-Zeinstra, S.M.A., Miedema, H.S., Verhagen, A.P., Nauta, A.P., Burdorf, A., Verhaar, J.A.N. & Koes, B.W. (2006). Prognostic indicators for non-recovery of non-traumatic complaints at arm, neck and shoulder in general practice - 6 months follow-up. *Rheumatology*, 46(1), 169-176.
9. Mallen, C.D., Peat, G., Thomas, E., Dunn, K.M., Croft, P.R. (2007). Prognostic factors for musculoskeletal pain in primary care: a systematic review. *British Journal of General Practice*, 57, 655-661.
10. Dorrestijn, O., Greving, K., van der Veen, W.J., van der Meer, K., Diercks, R.L., Winters, J.C & Stevens, M. (2011). Patients with shoulder complaints in general practice: consumption of medical care. *Rheumatology*, 50(2), 389-395.
11. Swinkels, I.C.S., Kooijman, M.K. & Leemrijse, C.J. (2011). *Four Years of self-referral to physiotherapy in the Netherlands*. World Physical Therapy Congress. 2011., LIPZ: Amsterdam.
12. Hegedus, E.J., Goode, A., Campbell, A., Morin, A., Tamaddoni, M., Moorman, C.T. & Cook. C. (2008). Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *British Journal Sports Medicine*, 42, 80–92.
13. Hughes, P.C., Taylor, N.F. & Green, R.A. (2008). Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. *Australian Journal of Physiotherapy*, 54(3), 159-170.
14. Beaudreuil, N., R., Thomas, T., Peyre, M., Liotard, J.P., Boileau, P., Marc, T., Dromard, C., Steyer, E., Bardin, T., Orcel, P. & Walch, G. (2009). Contribution of clinical tests to the diagnosis of rotator cuff disease: A systematic literature review. *Joint Bone Spine*, 76, 15-19.
15. Diehr, S., Ison, D., Jamieson, B. & Oh. R. (2006). Clinical inquiries: What is the best way to diagnose a suspected rotator cuff tear? *Journal of Family Practice*, 55, 621-624.
16. Dinnes, J., Loveman, E., McIntyre, L. & Waugh, N. (2003). The effectiveness of diagnostic tests for the assessment of

shoulder pain due to soft tissue disorders: a systematic review. *Health Technology Assessment, 7*(1-166).

17. Ottenheijm, R.P., Jansen, M.J., Staal, J.B., van den Bruel, A., Weijers, R.E., de Bie, R.A. & Dinant, G.J. (2010). Accuracy of diagnostic ultrasound in patients with suspected subacromial disorders: a systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation*, 91(10), 1616-1625.
18. Hall, A.M., Ferreira, P.H., Maher, C.G., Latimer, J. & Ferreira, M.L. (2010). The influence of the Therapist-Patient Relationship on Treatment Outcome in Physical Rehabilitation: A Systematic Review. *Journal of the American Physical Therapy Association*, 90, 1099-1110.
19. Dinant, G.J., Buntinx, F. & Butler, C.C. (2007). The necessary shift from diagnostic research to prognostic research. *BMC Family Practice*, 8(53).
20. Jacobson, J.A. (2011). Shoulder US: Anatomy, Technique, and Scanning Pitfalls. *Radiology*, 260(1).
21. Ginn, K.A.C., M.L. (2003). Conservative treatment for shoulder pain: Prognostic indicators of outcome. *Archives of Physical Medicine and Rehabilitation*, 85(8), 1231-1235.
22. Kennedy, C.A., Manno, M., Hogg-Johnson, S., Haines, T., Hurley, L., McKenzie, D. & Beaton, D.E. (2006). Prognosis in Soft Tissue Disorders of the Shoulder: Predicting Both Change in Disability and Level of Disability After Treatment. *Journal of the American Physical Therapy Association*, 86(7), 1018-1032.
23. Holtermann, A., Hansen, J.V., Burr, H. & Sogaard, K. (2010). Prognostic factors for long-term sickness absence among employees with neck-shoulder and low-back pain. Scandinavian Journal of Work, *Environment & Health*, 36(1), 34-41.
24. Zengh, X., Simpson, J.A., van der Windt, D.A. & Elliott, A.M. (2005). Data from a study of effectiveness suggested potential prognostic factors related to the patterns of shoulder pain. *Journal of Clinical Epidemiology*, 58(8), 823-830.
25. Kamper, S.J., Ostelo, R.W., Knol, D.L., Maher, C.G., de Vet, H.C. & Hancock, M.J. (2010). Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *Journal of Clinical Epidemiology*, 63(7), 760-766.
26. de Winter, A.F., van der Heijden, G.J., Scholten, R.J., van der Windt, D.A. & Bouter, L.M. (2007). The Shoulder Disability Questionnaire differentiated well between high and low disability levels in patients in primary care, in a cross-sectional study. *Journal of Clinical Epidemiology*, 60(11), 1156-1163.
27. Bot, S.D.M., Terwee, C.B., van der Windt, D.A.W.M., Bouter, L.M., Dekker, J. & de Vet, H.J.C. (2004). Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Annals of the Rheumatic Diseases*, 63, 335-341.
28. Winters, J.C., Sobel, J.S., Groenier, K.H., Arendzen, J.H. & Meyboom-De Jong, B. (1996). A shoulder pain score: a comprehensive questionnaire for assessing pain in patients with shoulder complaints. *Scandinavian Journal of Rehabilitative Medicine*, 28(3), 163-167.
29. Picavet, H.S., Hoeymans, N. (2004). Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DCM3 study. *Annals of the Rheumatic Diseases*, 63(6), 723-729.
30. Rabin, R., Oemar, M., Oppe, M., Janssen, B. & Herdman, M. (2011). *EQ-5D User Guide; Basic information on how to use the EQ-5D-5L instrument*. The EuroQol Group.
31. Stinckens, N., Ulburghs, A. & Claes, L. (2009). De werkalliantie als sleutelelement in het therapie gebeuren; Meting met behulp van de WAV-12, de Nederlandstalige

verkorte versie van Working Alliance Inventory. *Tijdschrift Klinische Psychologie*, 39, 44-60.

32. Hatcher, R.L.G., J.A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, 16, 12-15.
33. Girish, G., Lobo, L.G., Jacobson, J.A., Moraq, Y., Miller, B. & Jamadar, D.A. (2011). Ultrasound of the shoulder: asymptomatic findings in men. *American Journal of Roentgenology*, 197(4), 13-19.
34. Milgrom, C., Schaffler, M., Gilbert, S. & van Holsbeeck, M. (1995). Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender. *Journal of Bone Joint Surgery*, *77*(2), 296-298.
35. Ohrndorf, S., Naumann. L., Grundey, J., Scheel, T., Scheel, A.K., Werner, C. & Backhaus, M. (2010). Is musculoskeletal ultrasonography an operator-dependent method or a fast and reliably teachable diagnostic tool? Interreader agreements of three ultrasonographers with different training levels. *International Journal of Rheumatology,* 164518

Chapter 3

Physiotherapy for patients with shoulder pain in primary care: a descriptive study of diagnostic- and therapeutic management

Published in Physiotherapy

2017, 103(4)

Yasmaine Karel
Wendy Scholten-Peeters
Marloes Thoomes-de Graaf
Edwin Duijn
Joost van Broekhoven
Bart Koes
Arianne Verhagen

INTRODUCTION

Shoulder complaints are the third most common musculoskeletal complaint [1, 2]. The annual incidence of shoulder pain in the Netherlands is about 34 patients per 1000 [3]. About 10% of the patients presenting in physiotherapy practice have shoulder complaints [4]. In a Dutch study, 76% of these patients were referred by their general practitioner, 12% by a medical specialist and 12% accessed the physiotherapist without a referral [4]. About 50% of patients with shoulder pain in primary care have symptoms for more than six months [5, 6].

Frequently mentioned causes of shoulder pain in primary care are rotator cuff disease (subacromial impingement syndrome), glenohumeral disorders, acromioclavicular joint disease or referred neck pain [7]. Rotator cuff diseases are the most common cause of shoulder pain. The incidence is estimated to be 85% of the total population with shoulder pain in primary care, although more than one clinical diagnosis is made in 77% of the patients [8].

Most clinical tests are not valid in making a confident statement for pathology in patients with shoulder complaints [9, 10]. In the Netherlands physiotherapists increasingly use diagnostic ultrasound to assist their clinical decision-making, but the impact and specific aims of this diagnostic tool remain unknown [11, 12].

The most widely used interventions for patients with shoulder complaints are exercises, mobilization and/or massage [4, 13]. Current conclusions from (systematic) reviews describe moderate evidence for the effect of exercise therapy, manipulative therapy and NSAIDs (non-steroidal anti-inflammatory drugs) [14-16]. Physiotherapeutic interventions (exercise therapy and joint mobilizations) show a favorable outcome for patients with shoulder complaints [13, 15, 17, 18]. Several studies have shown good outcomes of non-operative management for patients with subacromial impingement syndrome [15, 19-21]. Despite physiotherapy treatment, in many patients (40%) the disability and physical impairments persist for over a year after the first symptom experience [6].

In The Netherlands, there is a Clinical Guideline for General Practitioners (GP) for the management of patients with shoulder pain and an evidence statement released by the Dutch Physiotherapist Society for patients suspected of having subacromial pain [22, 23]. Both the guideline and the evidence statement classify patients with non-specific shoulder pain into three subgroups: 1) pain during abduction (complaints arising from the subacromial space), 2) passive restricted range of motion (complaints arising from the glenohumeral joint) and 3) painful abduction and restricted passive range of motion (instability, complaints from the acromioclavicular joint or the neck).

To date, knowledge about the diagnostic strategies and therapeutic intervention(s) applied in primary care is limited [13, 24]. There is a lack of information on characteristics of physical examination and treatment in physiotherapy practice [4, 6].

Given the lack of clinical information for patients with various shoulder complaints in primary care, we aim to gain insight into current physiotherapy management, diagnostic- and treatment strategies. Gaining insight into current physiotherapy management may help guide further research and health care decisions.

METHODS.

Study design

This study was a prospective cohort study with a follow-up of 26 weeks in physiotherapy practice of patients with non-specific shoulder complaints. Primary aims of the ShoCo-DiP study were to evaluate physiotherapy care and prognostic factors in patients with shoulder pain. Secondary aims were to assess the inter-rater reliability of diagnostic ultrasound (US) between physiotherapists (PTs) and radiologists, and to assess whether patient characteristics of those who receive US differ from those who do not receive US. Details of the study design are published elsewhere [25]. The Medical Ethics Committee of the Erasmus Medical Center approved the study protocol (MEC-2011-414). In the current manuscript, the focus is on the description of PT care (diagnostic and therapeutic management) in the first 12 weeks of management and reported recovery after 12 and 26 weeks.

Physiotherapists & Patients

In total 125 physiotherapists from the South West region of the Netherlands participated in the study and they recruited patients, either referred by their GP or through direct access, from November 2011 until November 2012. Patients with shoulder pain were eligible when they were 18 years or over and adequately understood the Dutch language. Exclusion criteria were: patients with serious pathologies (infection, cancer or fracture), shoulder surgery in the past 12 months or diagnostic imaging techniques (musculoskeletal ultrasound, magnetic resonance imaging or radiography) performed on the shoulder in the past 3 months.

Data collection

Data was collected from both the PTs and the patients using digital questionnaires. Patient- and clinical characteristics were measured, and patients received follow-up questionnaires after 6 and 12 weeks concerning recovery. Characteristics (age, sex, work experience and/or specialization) of the PTs were reported before the start of the study. Physiotherapists reported their daily management at 3, 6 and 12 weeks in terms of clinical hypotheses after patient history (max. 3) and physical examination, initial clinical diagnosis, the use of diagnostic ultrasound (US), pathologic findings on US, changes in

clinical diagnosis after US and initial therapeutic management of the patient. Whenever a treatment plan changed during follow-up, the PTs reported the reasons for change and treatment goal(s).

Outcomes

Diagnostic process

We predefined a set of possible clinical diagnoses based on literature and consensus: subacromial impingement, internal impingement, glenohumeral instability, SLAP lesion, biceps tendinopathy, frozen shoulder, acromio-clavicular or sterno-clavicular joint pathology, sprain or strain, triggerpoints in the muscles of the shoulder and neck, muscular hypertension/hypotension, cervical-thoracic pathology or no clear clinical diagnosis.

Diagnostic US

The following pathological findings were listed: tendinopathy, calcification, full thickness/partial thickness tears, biceps tendon rupture, biceps halo, bursitis, subacromial impingement, glenohumeral discontinuity, acromion discontinuity, labrum tear/SLAP, capsular thickening, and rotator cuff atrophy.

Treatment process

Physiotherapists estimated patient's the prognosis at baseline (full recovery, clinical relevant reduction of complaints, stabilizing complaints or not estimable) and also reported their treatment of choice. Possible interventions were categorized into: information/advice, exercise therapy, massage, manual joint mobilization/manipulation, extracorporeal shockwave therapy (EST), transcutaneous electrical nerve stimulation, trigger point therapy, taping/bracing or posture correction. Each follow-up moment PTs could report whether 1) treatment was ended (additional information about number of treatments and reasons), 2) if any changes in planned treatment interventions were made and 3) if patients remained under treatment without any changes in treatment since baseline.

Recovery

Recovery status of the patient was measured with the Global Perceived Effect scale (GPE). The GPE uses a 7-point Likert scale indicating whether the patient's condition had improved or deteriorated since the start of their treatment. The outcome was dichotomised into "recovered" and "not recovered", with " recovered" defined as "completely recovered" or "much improved" [26-28]. The GPE is validated for patients with musculoskeletal complaints [29].

Statistical Analysis

Descriptive analyses were conducted using SPSS 22.0 statistical software. Descriptive statistics included patient's clinical and symptom characteristics, physiotherapists' characteristics, information from history taking, physical examination, utility of diagnostic ultrasound, treatment plan, average treatment period, possible changes of treatment plan since initiation at baseline, recovery or referrals to other (para)medical care.

Descriptive statistics were presented in mean scores for continuous data with a normal distribution. In all other cases, median scores and the interquartile range (IQR) were used. Hypotheses after patient history were categorized according to the guidelines (complaints arising from pathology/dysfunction in: 1) the subacromial space (subacromial impingement, internal impingement & sprain/strain), 2) glenohumeral joint (glenohumeral joint instability, frozen shoulder, biceps tendinopathy & SLAP), 3) acromioclavicular (AC)/sternoclavicular (SC) joint, 4) cervico-thoracic spine and 5) other and presented in a scaled rectangle diagram [30]. The number of missings were reported with all data.

RESULTS

Physiotherapists (n=125) were mostly men with a mean age of 39. Of all physiotherapists 50% (51/102) were specialized in manual therapy, and 37% (38/102) were trained to use diagnostic ultrasound. The response rate of the physiotherapists was 94% (366/389) at baseline and 93% (362/389) after 12 weeks.

A total of 389 patients with a mean age of 50 years (standard deviation of 13) were included (see Table 1 for baseline characteristics). After 26 weeks 70% (272/389) of patients had returned one or more follow-up questionnaires. No significant differences in baseline characteristics were found between the responders and non-responders.

Clinical diagnosis.

History taking: After history taking 48% (174/365) of patients had a suspected subacromial impingement as primary hypothesis, 14% (51/365) was rated with shoulder pain due to a cervical or thoracic dysfunction, 8% (29/365) was rated with a frozen shoulder, 5% (17/365) with glenohumeral joint instability and 4% (13/365) with AC/SC joint pathology (Table 2). As PTs could give a maximum of three hypotheses, the overlap between clinical hypotheses is presented in figure 1. In 92 patients the PT suspected either a subacromial impingement or pathology in the glenohumeral joint, and for 52 patients the PT suspected a subacromial impingement or pathology in the cervical-thoracic spine after history taking.

Physical examination: Frequently used specific test for a suspected subacromial impingement were Neer's Sign (177/241, 73%), Hawkins-Kennedy Test (193/241, 80%),

Table 1. Patient characteristics

	Total (n= 389)
Gender, men (%)	170 (43)
Age, mean (SD)	50 (13)
Duration in weeks, med (IQR)	12 (6-26)
History of shoulder pain (yes, %)	158 (40)
Onset (%)	
Sudden onset	118 (33)
Slow onset	246 (67)
Cause (%)	
Traumatic	79 (21)
Work related	132 (36)
Unclear	128 (35)
Other	29 (8)
Dominant side affected (Yes, %)	224 (57)
Shoulder surgery in the past (yes, %)	16 (4)
Corticosteroid injection (yes, %)	32 (8)
Medication (yes, %)	183 (47)
Comorbidity (yes, %)	236 (60)
Level of education:	
high school diploma or less	239 (65)
higher degree	127 (35)
Work	261 (67)
NRS, med (IQR)	6.0 (4-7)
SPS, med (IQR)*	18 (15-21)
SDQ, med (IQR)	62.5 (44-81)
EQ5D Tariff, med (IQR)	0.83 (0.77-0.87)

SD Standard Deviation, Med Median, IQR Interquartile Range, NRS Numeric Rating Scale, SDQ Shoulder Disability Questionnaire, EQ5D EuroQol-5 Dimensions, SPS Shoulder Pain Score
*The shoulder pain score consists of 6 pain symptoms questions together with the NRS

Empty/Full Can (204/241, 85%) and Painful Arc (154/241, 64%). For glenohumeral joint instability, the tests most frequently used were the O'Brien (25/54, 46%), the Relocation Test (38/54, 70%), the Apprehension Test (39/54, 72%), the Biceps Load 1&2 (12/54, 22%) and a Sulcus Sign (14/54, 26%). In the case of suspected AC joint pathology, the acromioclavicular joint play test (73/88, 83%) was most frequently used [31].

In 22% (73/333) of the patients, the physiotherapists changed the primary hypothesis after physical examination, but no specific patterns in these changes were found. After physical examination 39% (122/316) were diagnosed with subacromial impingement, 17% (54/316) with shoulder complaints due to a cervical of thoracic origin, 9% (29/316)

with a frozen shoulder, 7% (24/316) with glenohumeral joint instability, 7% (21/316) with a sprain or strain and 5% (17/316) with AC/SC joint pathology (Table 2).

Table 2. Clinical diagnosis (%) after patient history, physical examination and/or diagnostic ultrasound

	Clinical hypothesis after patient history (n=365)	Clinical diagnosis after physical examination and/or US (n= 316)
Subacromial impingement	174 (48)	122 (39)
Internal impingement	24 (7)	18 (6)
GH joint instability	17 (5)	24 (7)
SLAP lesion	1 (0.3)	2 (1)
Biceps tendinopathy	12 (3)	8 (3)
Frozen shoulder	29 (8)	29 (9)
Cervical/thoracic origin	51 (14)	54 (17)
AC/SC origin	13 (4)	17 (5)
Sprain/strain	17 (5)	21 (7)
Triggerpoints	-	2 (0.5)
Muscular hypertension	3 (1)	1 (0)
No clear clinical diagnoses		2 (0.5)
Other	20 (5)	16 (5)

GH Glenohumeral, AC/SC Acromio-clavicular/sterno-clavicular, SLAP Superior labrum anterior posterior, US Diagnostic Ultrasound

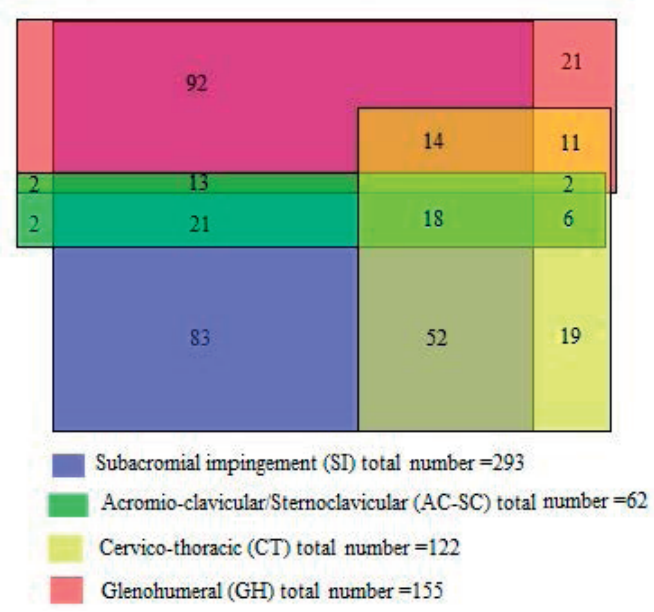


Figure 1. Scaled rectangle diagram showing the overlap for selected clinical hypothesis (max 3 per patient) by physiotherapists after patient history. Colors show the base color for each clinical hypothesis.

Diagnostic ultrasound (US).

A diagnostic US was performed in 31% (n=122) of all patients. In 92% (109/122) of these patients the US was performed before, or instead of, physical examination; in 38% (41/109) of these patients, the PT chose not to perform a physical examination anymore. In 34% (42/122) of all patients the reason to use a diagnostic US was that the PT expected this would lead to a more specific clinical diagnosis, and in 13% (16/122) that it would help the PT in selecting the most appropriate intervention. In 12% of the patients (15/122) the PTs used the US findings to confirm their initial diagnosis and in another 11% (14/122) to better inform the patient about their complaints. The results of the US were: a tendinopathy of the rotator cuff in 57% of the patients (70/122), a calcification of the rotator cuff in 38% (46/122), a partial thickness tear of the rotator cuff in 20% (24/122) and a full thickness tear of the rotator cuff in 5% (6/122) (table 3). Pathological findings were most frequently detected in the supraspinatus tendon.

The PTs assumed that the use of diagnostic US resulted in a better prediction of the integrity of tendon tissue in 51% of patients (62/122), a more specific exercise therapy in 42% (n=51), a better advice and better assessment of prognosis for 48% (59/122), more specific home exercises for 35% (43/122), behavior change in 33% (40/122) and an indication for EST in 16% (19/122) of the patients. Only in 11% (14/122) the results of US led to a hands-off policy and in 7% (8/122) the PTs stated that the use of diagnostic US had no consequence for the treatment plan.

In 16% (19/122) of the patients who had a diagnostic US the consequence of diagnostic US resulted in a referral to the general practitioner. Only 8% (21/267) of patients without a diagnostic US were referred (back) to their GP. These patients were mostly suspected with calcific tendinitis of the supraspinatus in 42% (8/19) and tendinopathy of the supraspinatus in 42% (8/19). Overall, the clinical diagnosis changed in 29% (35/122) of the patients after diagnostic US. In 31% (11/35) of these cases, the diagnoses changed from various diagnoses to a sprain (trauma) or strain.

Treatment plan

Baseline

At baseline, physiotherapists estimated full recovery in 50% (161/323) of the patients and a clinically relevant reduction of complaints in another 47% (152/323) within the estimated treatment period. Physiotherapists estimated full recovery for 80% (43/54) of patients with shoulder pain due to a suspected cervical or thoracic dysfunction. Estimated recovery was lower in all other diagnostic categories. The longest treatment period (>26 weeks) was estimated for patients with a suspected frozen shoulder.

The PTs chose a variety of interventions but most commonly gave advice (331/365, 91%) and exercise therapy (296/365, 81%) (Table 4). The aims for exercise therapy were to improve muscle functions of the rotator cuff and improve stability function of the

Table 3. Findings on US

	US (n=122)
No structural pathology	9 (7)
Not interpretable	2 (2)
Tendinopathy:	70 (57)
Biceps	15 (12)
Supraspinatus	45 (37)
Infraspinatus	4 (3)
Subscapularis	6 (5)
Calcification:	46 (38)
Biceps	2 (2)
Supraspinatus	34 (29)
Infraspinatus	4 (3)
Subscapularis	6 (5)
Full thickness tear:	6 (5)
Supraspinatus	5 (4)
Infraspinatus	-
Subscapularis	1 (1)
Partial thickness tear:	24 (20)
Supraspinatus	20 (16)
Infraspinatus	-
Subscapularis	4 (3)
Biceps tendon rupture	2 (2)
Biceps halo	7 (6)
Bursitis	13 (11)
Subacromial impingement	20 (16)
Arthritis/Arthrosis of AC joint	12 (10)
Glenohumeral discontinuity	4 (3)
Acromion discontinuity	2 (2)
Labrum tear/SLAP	2 (2)
Capsular thickening	1 (1)
Rotatorcuff athrophy	3 (2)
Other	3 (2)

scapulo-thoracic joint. A smaller portion of the PTs chose transcutaneous electric nerve stimulation (TENS) (5/365, 1%), massage (27/365, 7%) and tape/bracing techniques (54/365, 15%). For patients with a suspected subacromial impingement syndrome, 92% (112/122) of the patients received advice and exercise therapy. For patients with a suspected cervical or thoracic dysfunction, the preferred treatment strategy was advice (50/53, 93%) and manual mobilization/manipulation of the spine (49/53,

Table 4. Planned PT interventions

PT interventions	Total n=365	Patients with SI n=122	Cervical / thoracic origin n=53	Frozen shoulder n=29	GH instability n=24
Information/advice (%)	331 (91)	112 (92)	50 (93)	29 (100)	20 (83)
Exercise therapy (%)	296 (81)	112 (92)	36 (67)	18 (62)	23 (96)
Massage (%)	27 (7)	5 (4)	11 (20)	4 (14)	-
Manipulation/ mobilization (%)	208 (57)	57 (47)	49 (91)	23 (79)	6 (25)
Shockwave (%)	39 (11)	29 (24)	1 (2)	-	-
Transcutaneous Electric stimulation therapy (%)	5 (1)	-	-	2 (7)	-
Trigger point therapy (%)	32 (9)	11 (9)	5 (9)	3 (10)	2 (8)
Taping/bracing (%)	54 (15)	29 (24)	-	-	11 (46)
Posture correction (%)	7 (2)	2 (2)	3 (5)	-	-
Other (%)	25 (7)	4 (3)	3 (5)	3 (10)	1 (4)

PT Physiotherapy, GH Glenohumeral, SI subacromial impingement

91%). Patients with frozen shoulder also received mostly advice (29/29, 100%) and manual joint mobilization of the shoulder or cervical spine (23/29, 79%).

Follow-up

At 6 weeks, 41% (118/285) of patients reported to be recovered, 57% (152/269) at 12 weeks and 60% (164/272) at 26 weeks.

In total, 12% (44/362) of the patients ended treatment at 3 weeks, 29% (109/373) at 6 weeks and 59% (214/363) at 12 weeks. Of 69% (148/214) of all patients that ended treatment within 12 weeks the physiotherapist decided to stop treatment because treatment goals had been achieved; 13% (27/214) of the patients had stopped the treatment themselves, 10% (21/214) had been referred to their general practitioner, and 5% (11/214) had been referred to another health care professional. The referral rate in the first 3 weeks was higher compared to later follow-up moments.

Figure 2 shows the course of recovery for each follow up moment per diagnostic category. The subgroup of patients with frozen shoulder worsened during follow-up. At 6 weeks most patients with a subacromial impingement syndrome (SIS) had recovered. For patients with SIS, 41% (36/88) reported being recovered at 6 weeks. Patients with SIS, who reported no recovery, 73% (38/52) were still under treatment after 6 weeks and 50% (17/34) after 12 weeks.

During the treatment period, the PTs changed the treatment plan in 16% (58/365) of the patients and 3% (11/365) the PT changed the treatment plan twice. Reasons for changing the treatment plan were because of the absence of progression (in 38% (22/58) of the patients), a change in the course of the disease (in 26% (15/58)) or unforeseen dysfunctions (in 17% (10/58)).

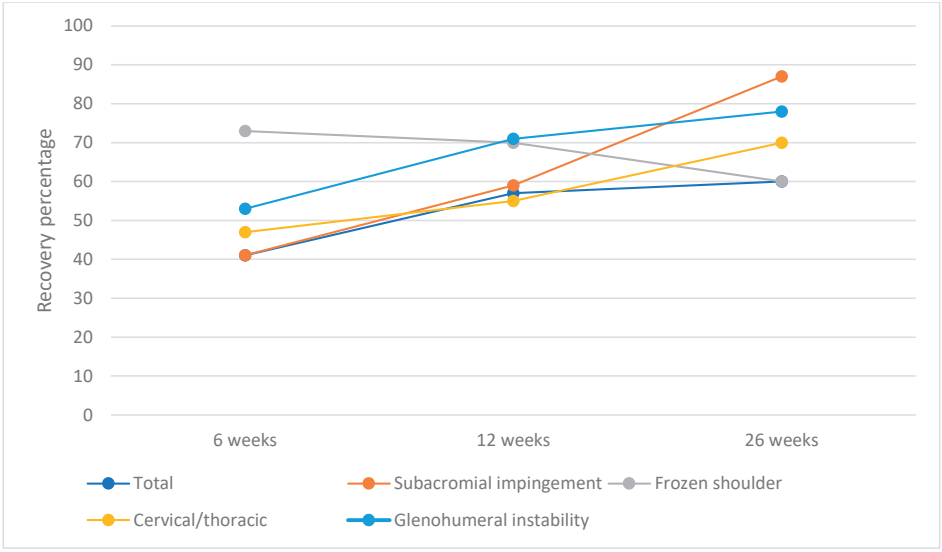


Figure 2. Percentage of recovery (GPE) per diagnostic category for all follow-up moments.

Overall, the median number of treatment sessions was 7 (IQR 6). When treatment stopped between 0 and 3 weeks, the median number of sessions was 4 (IQR 3). Between 3 and 6 weeks the median was 6 (IQR 4). Between 6 and 12 weeks the median was 7 (IQR 5).

DISCUSSION

Main findings

Physiotherapists (PTs) suspected subacromial impingement and complaints due to a cervical/thoracic origin to be the cause of their shoulder pain in most patients. In 31% of patients, diagnostic US was used. Pathologies most frequently found using US were in the supraspinatus tendon. Almost all patients received information and advice by their PT. Patient with suspected subacromial impingement, besides advice, received exercise therapy most frequently and patients with cervical thoracic originated shoulder pain received mostly manipulation/mobilization. About 33% of the patients stopped treatment within 12 weeks with a median number of treatment sessions of 7. After 26 weeks 60% of patients reported being recovered.

Comparison with existing literature

Patient demographics were similar to populations described by other studies [4, 32, 33]. The mean duration of complaints before seeking help for their shoulder complaint was

relatively long but comparable with another observational study from the Netherlands [4]. We found a median time of 12 weeks before seeking help, which might be a valuable time-period in which PTs can advise patients to reduce the chance of chronicity.

No previous cohort study collected descriptive data concerning the diagnostic process of the PT and the related interventions used. The most frequent clinical hypothesis found in this study was subacromial impingement. Also, other primary care studies described subacromial impingement to be the most common cause of shoulder complaints [34-37]. The scaled rectangle diagram (figure 1) showed the greatest overlap/concomitance between the formulated hypothesis subacromial impingement and cervical/thoracic originated shoulder pain. Subacromial impingement is probably caused by multiple factors and sometimes suggested to be a secondary complaint [38, 39]. Literature suggests that targeting adjunctive regions (cervical and thoracic spine) of the shoulder has beneficial effects and thus might be related to subacromial impingement possibly causing PTs to believe that the cervical region is related to subacromial impingement syndrome [39-41]. Similar to our study, a survey amongst physiotherapists in the United Kingdom, concluded that advice and exercise were administered the most in patients with rotator cuff disease [42].

Recovery rate at 12 weeks was 44% and 60% at 26 weeks, which is similar to the previously reported recovery rates in the literature [5, 43, 44]. However, not all recovered patients received an equal amount of physiotherapy treatment.

Strengths and limitations

This is the first study that describes the daily management of physiotherapy in patients with shoulder pain. Furthermore, this study is the first that evaluates different treatment strategies based on clinical characteristics from patient history and/or physical examination.

The response rate for participants was 70% after 12 weeks of follow-up. Dropout rates in observational studies remain challenging and in order to prevent dropouts proper actions were described in the study protocol [25]. All participants were sent personal links to their e-mail address at the time of follow-up with 2 reminders and for patients without computers or Internet the questionnaire was sent on paper. The PTs were sent the questionnaires to their e-mail, and the response rate was 93%. Both PTs and patients were telephoned twice during the study period to keep dropout and loss to follow-up at a low level.

This study found a median complaint duration of 12 weeks at baseline. Complaint duration could be this long since there is a possibility that patients were seen by their GP, who could apply a wait and see policy, before referring to physiotherapy treatment. It was unknown whether patients used direct access or if they were referred by their GP.

The list of potential diagnoses used was developed based on the rationale of clinical experts. No protocols or standardizations on diagnostic categories, tests, or treatments were used due to the nature of the study, as we wanted to describe daily management.

Based on the literature we know that specific tests are not valid in making a confident statement for pathology, and therefore the possibility of error in the clinical diagnosis made by the PTs is a problem. However, even if strict criteria for subgroups are stated in advance, the interobserver agreement in classification of current used subgroups, and the clinical tests leading to their diagnosis is only fair to moderate (Kappa 0.2–0.6) [45]. This implies that the usefulness of the currently used subgroups is still hampered by the lack of reproducibility of the diagnostic criteria. On the contrary PTs have to deal with diagnostic uncertainties; Figure 1 shows the overlap of diagnostic categories indicating that physiotherapists do not work with one hypothesis. Another study in GP practice came to the same conclusion [46]. These uncertainties might contribute to the increased use of diagnostic ultrasound. Physiotherapists will use US for a variety of other reasons depending on their level of expertise, specialization or the complaint of the patient (ie biofeedback for low back pain). It is imperative that physiotherapists are allowed to utilize US to optimize the effectiveness of their interventions, but it should be determined how this tool can best benefit the patient.

A limitation of this study may be the generalization of the results. Dutch PTs were asked to participate in a cohort study for shoulder complaints collecting data about physiotherapy management. Secondary aims about diagnostic US were also mentioned. It was possible that PTs who were specialized or more interested in the use of US would be more likely to participate leading to a higher frequency of US scans resulting in a biased sample of physiotherapists in this study. Sampling bias was however taken into account by recruiting physiotherapists in different ways (emails to the network of PT from the applied university, emails to PT supervisors of students of the applied university, physically addressing PTs to participate in symposia and emails to the shoulder networks in the region).

The utilization of US might have been influenced by one of the original study purposes (namely: the inter-rater reliability of US between radiologists and physiotherapists) explaining the large number of ultrasonographers. The total number of PTs with an ultrasound machine was 44 (35%), which might be higher than average. A second US scan by the radiologist was only requested when PTs had reported performing an US scan, representing usual care. Physiotherapists were never asked to conduct an US scan. For one third of the cases, the physiotherapist changed their clinical diagnosis after diagnostic US and believed the complaints were due to a sprain (trauma) or strain like rotator cuff tears. The findings on US could then lead to an increase in the number of referrals to the GP. However, we did not collect data on further interventions if patients were referred to the general practitioner or orthopedic surgeon. There is a controversy regarding the management of rotator cuff tears. Whether these patients would be better off being referred to the GP or orthopedic surgeon is not clear. It might be argued that small tears should be repaired to relieve symptoms and prevent tear progression, but little

evidence exists to support this view. Another study has found good results for patients with partial or full thickness tears receiving physiotherapy [51]. Furthermore, pathologies seen on US might not be the cause of symptoms experienced by the patient [47, 48]. Findings on diagnostic US should be interpreted with caution as studies have found a high number of pathologies in asymptomatic shoulders. Reliability of diagnostic US between radiologists and physiotherapists is substantial for full thickness rotator cuff tears [49].

All diagnostic subgroups, except patients with a suspected frozen shoulder, showed an improvement after 26 weeks. We are not sure whether these positive results could be attributed to the diagnostic subgroup or whether recovery was a reflection of the therapeutic intervention or the natural course of shoulder pain. However as might be expected from literature the complaint of patients with a suspected frozen shoulder got worse over time in the pain and stiffness stage [50]. A large group of patients were still not recovered after 12 weeks of treatment which might be attributed to the heterogeneous sample, the adherence of patients or the natural disease process.

Specific interventions were chosen in patients with subacromial impingement, cervical or thoracic dysfunction or frozen shoulder. Variability might exist on the exact interpretation of physiotherapeutic interventions. In the case of exercise therapy for subacromial impingement syndrome, specific exercises were not standardized.

During data collection, physiotherapists could select the physical examination tests based on the hypothesis, or multiple hypotheses, after patient history. Most physiotherapists formulated multiple hypotheses and therefore analyzing the tests used for each clinical diagnosis was impossible. Furthermore, we hypothesized that the decision for the interventions was primary based on the clinical diagnosis. However, patient preferences or other factors could have influenced these decisions. This study assumed that physiotherapists mostly use a patho-anatomical model to generate an early hypothesis. However new strategies, like the symptom modification procedure, use symptom provoking procedures to select whether treatment should focus on the glenohumeral joint, the scapula or the cervical/thoracic spine [51]. This procedure was proposed because clinicians recognized the complexity of making a definitive diagnosis and it might be that physiotherapists in the Netherlands already use this model in practice. However research for this new method of assessment is still unavailable.

Implications for practice

Subacromial impingement and complaints due to a cervical/thoracic origin were in most patients suspected to be the cause of their shoulder pain. Shoulder and neck pain often coincide together, but it's not clear whether PTs can distinguish between the two. The evidence statement for subacromial syndrome recommends exercise therapy (if there is sufficient mobility) and manual mobilizations (when absolutely necessary) [22]. This is consistent with observations in clinical practice for patients with subacromial impingement

who for the most part received exercise therapy. A small proportion chose interventions (TENS, massage and tape/bracing) not recommended by the evidence statement.

The evidence statement furthermore states that patients should be referred to the general practitioner or orthopedic surgeon if pain and activity levels did not improve [22]. Although we observed that 73% of patients with subacromial impingement, who had no or insufficient improvement, still received treatment after 6-12 weeks. This means that most patients were not treated according to recommendations from the evidence statement, which states referral to the GP when no improvement is seen after 6-12 weeks of physiotherapy. However, to date, there is no good evidence that referral to the GP, possible surgery, medications or injections are better than conservative management [52].

Implications for research

Our results show that PTs frequently use diagnostic US as a replacement for physical examination. The latest review of diagnostic tests in shoulder complaints described moderate accuracy for some shoulder tests but not yet validated by multiple studies [10, 53]. Whether US could assist diagnostic accuracy for the physiotherapist in primary care should be investigated by studying the combined effect of physical tests and US in large clinical trials.

This study describes physiotherapy care for patients with shoulder complaints. However, the exact reasons for the clinical decisions, like the number of treatments or the presumed prognosis, should be investigated further.

CONCLUSIONS

We observed that most patients were suspected of having subacromial impingement, or cervical thoracic originated shoulder pain. Exercise therapy and manual mobilizations were most frequently utilized and consistent with interventions recommended for patients with subacromial impingement syndrome. Diagnostic ultrasound was utilized in one-third of the patients and PTs expected that this would lead to a more specific clinical diagnosis, but the effect on patient recovery remains unknown. Modest differences for the choice of interventions were observed and consensus is required.

REFERENCES

1. Adamson, J., Ebrahim, S., Dieppe, P., Hunt, K. (2006). Prevalence and risk factors for joint pain among men and women in the West of Scotland Twenty-07 study. *Annals of Rheumatic Diseases,* 65, 520-4.
2. Macfarlane, G.J., Beasley, M., Smith, B.H., Jones, G.T., Macfarlane, T.V. (2015). Can large surveys conducted on highly selected populations provide valid information on the epidemiology of common health conditions? An analysis of UK Biobank data on musculoskeletal pain. *British Journal of Pain,* 9(4), 203-212.
3. Prins, M., Hek, K., Verberne, L., Nielen, M., Opperhuizen, G., Verheij, R. (2015). *Zorg door de huisarts; jaarcijfers 2014 en trendcijfers 2010-2014."* NIVEL Zorgregistraties, Utrecht.
4. Kooijman, M., Swinkels, I., van Dijk, C., de Bakker, D., Veenhof, C. (2013). Patients with shoulder syndromes in general and physiotherapy practice: an observational study. *BioMedCentral Musculoskeletal Disorders,* 14(128).
5. Winters, J., Sobel, J.S., Groenier, K.H., Arendzen, J.H. Meybook-de Jong, B. (1999). The long-term course of shoulder complaints: a prospective study in general practice. *Rheumatology,* 38, 160-3.
6. Kuijpers, T., van der windt, D.A.W.M., van der Heijden, G.J.G.M., Bouter, L.M. (2004). Systematic review of cohort studies on shoulder disorders. *Pain,* 109(3), 420-31.
7. Holmes, R.E., Barfield, W.R., Woolf, S.K. (2015). Clinical evaluation of nonarthritic shoulder pain: Diagnosis and treatment. *The Physician and sportsmedicine,* 43(3), 262-8.
8. Östör, A.J.K., Richards, C.A., Prevost, A.T., Speed, C.A., Hazleman, B.L. (2004). Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology,* 44(6), 800-805.
9. Hermans, J., Luime, J., Meuffels, D.E., Reijman, M., Simel, D.L., Bierma-Zeinstra, S.M.A. (2013). Does this patient with shoulder pain have rotator cuff disease? The rational clinical examination systematic review. *JAMA,* 310(8).
10. Hegedus, E.J., Goode, A.P., Cook, C.E., Michener, L., Myer, C.A., Myer, D.M., Wright, A. (2012). Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine,* 46(14), 964-78.
11. Hanchard, N.C., Lenza,M., Handoll, H.H., Takwoingi, Y. (2013). Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *The Cochrane Database of systematic reviews,* 30(4).
12. Wright, A.A., Wassinger, C.A., Frank, M., Michener, L.A., Hegedus, E.J. (2013). Diagnostic accuracy of scapular physical examination tests for shoulder disorders: a systematic review. *British Journal of Sports Medicine,* 4(13), 886-92.
13. Green, S., Buchbinder, R., Hetrick, S.E. (2003). Physiotherapy interventions for shoulder pain (Review). *The Cochrane Library*, (2).
14. Schellingerhout, J.M., Thomas, S., Verhagen, A.P. (2007). Aspecific shoulder complaints: literature review to assess the efficacy of current interventions." *Nederlands Tijdschrift voor Geneeskunde,* 151(52), 2892-7.
15. Michener, L.A., Walsworth, M.K., Burnet, E.N. (2004). Effectiveness of rehabilitation for patients with Subacromial impingement syndrome: a systematic review. *Journal of Hand Therapy,* 17(2), 152-164.
16. Kuhn, J.E. (2009). Exercise in the treatment of rotator cuff impingement: A systematic review and a synthesized evidence-based

rehabilitation protocol. *Journal of Shoulder and Elbow Surgery,* 18, 138-160.

17. Kelly, S.M., Wrightson, P.A., Meads, A.C. (2010). Clinical outcomes of exercise in the management of subacromial impingement syndrome: a systematic review. *Clinical Rehabilitation,* 24, 99-109.
18. Hanratty, C.E., et al. (2012). The effectiveness of physiotherapy exercises in subacromial impingement syndrome: a systematic review and meta-analysis. *Seminars in Arthritis and Rheumatism* 42(3), 297-316.
19. Bang, M.D., Deyle, G.D. (2000). Comparison of supervised exercise with and without manual physical therapy for patients with shoulder impingement syndrome. *J Orthop Sports Phys Ther,* 30(3), 126-137.
20. Dickens, V.A., Williams, J.L., Bhamra, M.S. (2005). Role of physiotherapy in the treatment of subacromial impingement syndrome: a prospective study. *Physiotherapy,* 91(3), 159-164.
21. Conroy, D.E., Hayes, K.W. (1998). The effect of joint mobilization as a component of comprehensive treatment for primary shoulder impingement syndrome. *J Orthop Sports Phys Ther.,* 28(1), 3-14.
22. Jansen, M.J., et al. (2011). KNGF Evidence Statement Subacromiale klachten. *Nederlands Tijdschrift voor Fysiotherapie,* 121, no. 1.
23. Winters, D. (2008). NHG-standaard schouderklachten. *Huisarts Wet,* 51, 555-65.
24. Schellingerhout, J.M., Verhagen, A.P., Thomas, S., Koes, B.W. (2008). Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach. *Manual Therapy,* 16,(6), 478-83.
25. Karel, Y.H.J.M., Scholten-Peeters, W.G.M., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R.P.G., van den Borne, M.J.P., Koes, B.W., Verhagen, A.P. (2013). Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. *BioMedCentral Musculoskeletal Disorders,* 14(62).
26. Beurskens, A.J., de Vet, H.C., Koke, A.J. (1996). Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain,* 65, 71-76.
27. Pool, J.J., Ostelo, R.W., Koke, A.J., Bouter, L.M., de Vet, H.C. (2006). Comparison of the effectiveness of a behavioural graded activity program and manual therapy in patients with sub-acute neck pain: design of a randomized clinical trial. *Manual Therapy,* 11, 297-305.
28. Vonk, F., Verhagen, A.P., Geilen, M., Vos, C.J., Koes, B.W. (2004). Effectiveness of behavioural graded activity compared with physiotherapy treatment in chronic neck pain: design of a randomised clinical trial. *BMC Musculoskeletal Disorder, 5*(34).
29. Kamper, S., Ostelo, R., Knol, D., Maher, C., de Vet, H., Hancock, M. (2010). Global perceived effect scales provided reliable assessments of health transition on people with musculoskeletal disorders, but ratings are strongly influences by current status. *Journal of Clinical Epidemiology,* 63(7), 760-6.
30. Marshall, R.J. (2005). Scaled rectangle diagram can be used to visualize clinical and epidemiological data. *Journal of Clinical Epidemiology,* 58, 974-981.
31. Funk, L. Shoulder examination tests. Shouderdoc. https://www.shoulderdoc.co.uk/section/497. Accessed June 27th, 2016
32. Van der Windt, D.A, Koes, B.W., Deville, W., Boeke, A.J.P., Bouter, B.A., de Jong, L.M. (1998). Effectiveness of corticoid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomiser trial. *British Medical Journal,* 317, 1292-6.
33. Hay, E.M, Thomas, E., Paterson, S.M., Dziedzic, K., Croft, P.R. A pragmatic randomized controlled trial of local corticosteroid injection and physiotherapy for the treatment of new episodes of unilateral

shoulder pain in primary care. *Annal of Rheumatic Diseases,* 62, 94-9.

34. Stevenson, J.H., Trojian, T. (2002). Evaluation of shoulder pain. *Journal of Family Practice,* 51, 605-11.
35. Arcuni, S.E. (2000). Rotator cuff pathology and subacromial impingement. *Nurse Practitioner,* 25(58), 65-6.
36. de Winter, A.F, Jans, M.P., Scholten, R.J.P.M., Devillé, W., van Schaardenburg, D., Bouter, L.M. (1999). Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. *Annals of Rheumatic Disease,* 58, 272-277.
37. Mitchell, C., Adebajo, A., Hay, E., Carr, A. (2005). Shoulder pain: diagnosis and management in primary care. *BMJ,* 331(7525), 1124-1128.
38. Rossi, F. (1998). Shoulder impingement syndromes. *European Journal of Radiology* 27,S42–S48.
39. Walser, R.F., Meserve, B.B., Boucher, T.R. (2009). The effectiveness of thoracic spine manipulation for the management of musculoskeletal conditions: a systematic review and meta-analysis of randomized clinical trials. *Journal of Manual and Manipulative Therapy,* 17(4), 237–246.
40. Mintken, P.E., Cleland, J.A., Carpenter, K.J., Bieniek, M.L., Keirns, M., Whitman, J.M. (2010). Some factors predict successful short-term outcomes in individuals with shoulder pain receiving cervicothoracic manipulation: a single-arm trial. *Physical Therapy,* 90(1), 26–42.
41. Strunce, J.B., Walker, M.J., Boyles, R.E., Young, B.A. (2009). The immediate effects of thoracic spine and rib manipulation on subjects with primary complaints of shoulder pain." *Journal of Manual and Manipulative Therapy,* 17(4), 230–236.
42. Littlewood, C., Lowe, A., Moore, J. (2012). Rotator cuff disorders: a survey of current UK physiotherapy practice. *Shoulder & Elbow,* 4, 64-71
43. Kuijpers, T, et al. (2005). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain,* 120(3), 276–285.
44. Croft, P., Pope, D., Silman, A. (1996). The clinical course of shoulder pain: prospective cohort study in primary care. *British Medical Journal,* 313, 601-2.
45. Ostor, A.J., Richards, C.A., Prevost, A.T., Hazleman, B.L., Speed, C.A. (2004). Interrater reproducibility of clinical tests for rotator cuff lesions. *Annals of the Rheumatic Diseases,* 63, 1288–92
46. Ottenheijm, R.P.G., Hesselmans, N.J.J.M., Kemper, A., Moser, A., de Bie, R.A., Dinant, G.J., Cals, J.W.L. (2014). GPs' perspectives on the diagnostic work-up in patients with shoulder pain: a qualitative study. *Journal of Evaluation in Clinical Practice,* 20, 239-245
47. Minagawa, H., et al. (2013). Prevalence of symptomatic and asymptomatic rotator cuff tears in the general population: From mass-screening in one village. *Journal of Orthopaedics,* 10(1), 8-12.
48. Girish, G., Lobo, L.G., Jacobson, J.A., Morag, Y., Miller, B., Jamadar, D.A. (2011). Ultrasound of the shoulder: asymptomatic findings in men. *American Journal of Roentgenology,* 197(4), 13-9.
49. Thoomes-de Graaf, M., et al. (2014). Interprofessional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Manual Therapy,* 19(5), 478-483.
50. Reeves, B. (1975). The natural history of the frozen shoulder syndrome. *Scandinavian Journal of Rheumatology,* 4(4), 193-6
51. Lewis, J.S. (2009). Rotator cuff tendinopathy/subacromial impingement syndrome: is it time for a new method of assessment? *British Journal of Sports Medicine,* 43, 259-264
52. Abdulla, S.Y., Southerst, D., Côté, P., Shearer, H.M., Sutton, D., Randhawa, K., Varatharajan, S., Wong, J.J., Yu, H.,

Marchand, A.A., Chrobak, K., Woitzik, E., Shergill, Y., Ferguson, B., Stupar, M., Nordin, M., Jacobs, C., Mior, S., Carroll, L.J., van der Velde, G., Taylor-Vaisey, A. (2015). Is exercise effective for the management of subacromial impingement syndrome and other soft tissue injuries of the shoulder? A systematic review by the Ontario Prtotocol for Traffic Injury Management (OPTIMa) Collaboration. *Manual Therapy,* 20(5), 646-56

53. Hanchard, N.C., Lenza, M., Handoll, H.H., Takwoingi, Y. (2013). Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *The Cochrane Database of systematic reviews,* 30(4).

Chapter 4

Effect of routine diagnostic imaging for patients with musculoskeletal disorders: A meta-analysis

Published in European Journal of Internal Medicine

2016, 26(8)

Yasmaine Karel
Karin Verkerk
Silvio Endenburg
Sven Metselaar
Arianne Verhagen

ABSTRACT

Purpose

The increasing use of diagnostic imaging has led to high expenditures, unnecessary invasive procedures and/or false-positive diagnoses, without certainty that the patients actually benefit from these imaging procedures. This review explores whether diagnostic imaging leads to better patient-reported outcomes in individuals with musculoskeletal disorders.

Method

Databases were searched from inception to September 2013, together with scrutiny of selected bibliographies. Trials were eligible when: 1) a diagnostic imaging procedure was compared with any control group not getting or not receiving the results of imaging; 2) the population included individuals suffering from musculoskeletal disorders, and 3) if patient-reported outcomes were available. Primary outcome measures were pain and function. Secondary outcome measures were satisfaction and quality of life. Subgroup analysis was done for different musculoskeletal complaints and high technological medical imaging (MRI/CT).

Results

Eleven trials were eligible. The effects of diagnostic imaging were only evaluated in patients with low back pain (n=7) and knee complaints (n=4). Overall, there was a moderate level of evidence for no benefit of diagnostic imaging on all outcomes compared with controls. A significant but clinically irrelevant effect was found in favor of no (routine) imaging in low back pain patients in terms of pain severity at short [SMD 0.17 (0.04-0.31)] and long-term follow-up [SMD 0.13 (0.02-0.24)], and for overall improvement [RR 1.15 (1.03-1.28)]. Subgroup analysis did not significantly change these results.

Conclusion

These results strengthen the available evidence that routine referral to diagnostic imaging by general practitioners for patients with knee and low back pain yields little to no benefit.

Keywords

diagnostic tests, musculoskeletal/connective tissue disorders, back pain, primary care, radiology.

INTRODUCTION

For patients in whom the diagnosis remains uncertain after history taking and physical examination, general practitioners (or clinicians in general) can turn to diagnostic imaging modalities [1]. However, there has been a steady but debatable increase in the use of diagnostic imaging. For example, in the USA, between 1995 and 2005 the frequency of computed tomography (CT) has doubled and for magnetic resonance imaging (MRI) it has more than tripled [2]. The increase of diagnostic tests can lead to a false-positive diagnosis, 'pseudo' disease, or adverse effects, resulting in an unnecessary chain of events [3-6]. Imaging procedures may also lead to incidental findings, which can be found in both symptomatic and asymptomatic individuals [7, 8] indicating that diagnostic imaging findings may not always be responsible for the complaints experienced by the patient. The USA has experienced an larger number of spine surgeries due to an increase in the rate of spinal imaging [9] and others have reported increasing costs due to diagnostic imaging [10-12]. On the other hand the advancements in medical imaging techniques like MRI and other high technological medical imaging techniques can be used to replace older imaging techniques.

A previous systematic review including six randomized clinical trials (RCTs) in low back pain patients reported that immediate, routine lumbar spine imaging did not improve patient-reported outcomes [13]. Several trials have focused on patients with other musculoskeletal disorders, of which two found significant results for the effect of imaging [14-16]. Clinicians generally assume that reassurance must follow from a confident statement that no disease has been found. Nevertheless, negative test results are not always effective in reassuring patients [17]. A recent systematic review of five RCTs concluded that there is very limited evidence from current studies for the reassuring value of diagnostic tests in patients with varying complaints [18].

Although diagnostic imaging procedures are believed to influence patient care in a variety of ways, it remains unclear whether there is sufficient evidence to show that patient outcomes improve due to diagnostic imaging [13, 18]. Until now, no review has studied the effectiveness of diagnostic imaging for patients with musculoskeletal disorders other than low back pain, or has used the GRADE approach to determine the strength of the evidence. Therefore, this review aims to evaluate the role of immediate (after first consultation) diagnostic imaging procedures in patients with musculoskeletal disorders on patient-reported outcome measures (PROMs) using the GRADE approach.

METHODS

Selection criteria

RCTs were eligible when: 1) a diagnostic imaging procedure was compared with a control group not getting diagnostic imaging or not receiving results of imaging; 2) the population included individuals suffering from musculoskeletal disorders, and 3) if one of the following primary outcomes were reported: disability, pain, sick leave, quality of life, satisfaction, mental health, reassurance, or overall improvement/recovery.

Search method

Three review authors (YK,SE,SM) identified RCTs by searching the databases of MEDLINE, Cochrane, EMBASE and PubMed from inception to September 2013 (supplementary material). Relevant reference lists were also reviewed for additional citations. Two review authors (YK,KV) independently performed the study selection. Any disagreements were resolved by discussion, or with a third review author (AV), to reach consensus.

Risk of bias assessment

Two review authors (YK,KV) independently assessed the risk of bias using the Delphi list [19, 20]. In case of discrepancy, discussion was used to resolve any disagreement, or with a third review author (AV), to reach consensus. The Delphi list consists of nine items. For the present review we consider a study to have low risk bias when five or more of the items are answered with "yes"; this is supported by empirical evidence from the Cochrane Back Review Group [21].

Data extraction

Data extraction was first done by one review author (YK) using a standardized form and checked by a second author (KV), independently. When necessary, a third author (AV) resolved discrepancies. Descriptive data included study setting, country, selection criteria, population characteristics, description of intervention(s), outcomes (pain, function, quality of life, recovery and satisfaction) and follow-up. We extracted the number of participants randomized, the number of patients included in each analysis, and the means and standard deviations (SDs) of follow-up measurements.

Data analysis

Short-term follow-up was defined as being closest to 3 months and long-term follow-up as being closest to 12 months. Studies were excluded from analysis if they had insufficient data on means (or within-group differences) and SDs and the original authors could not be contacted. Pooling was done using a random effects model [22]. In case only median scores could be extracted, the median value was used as the mean and the SD was

estimated from the interquartile range. For continuous outcomes the standardized mean differences (SMD) was calculated and a risk ratio (RR) for dichotomous outcomes with the accompanying 95% confidence intervals (CI). A SMD of 0-0.2 was regarded as no effect, 0.2-0.5 as a small effect, 0.5-0.8 as a moderate effect, and >0.8 as a large effect [23]. Results were considered clinically relevant when the difference between groups was ≥ 15% [24]. Wherever possible, subgroup analyses were done (separately) for different musculoskeletal complaints, study setting, and/or imaging methods (high technological imaging techniques like MRI/CT). Pooling the effects of all trials was done when heterogeneity was low ($I^2 \leq 40\%$), otherwise only the subgroup analysis was reported. Sensitivity analysis was done excluding studies with a high risk of bias, in order to control for biased results. A funnel plot evaluated publication bias only if there were ≥ 10 trials for each effect estimate; otherwise, the power of the tests would be too low to distinguish the chance from real asymmetry [25]. All analyses were conducted in Review Manager 5.2.

Strength of the evidence

The Grades of Recommendation, Assessment Development and Evaluation (GRADE) was applied to assess the overall quality of the evidence and strength of recommendations [26]. The quality of the evidence for a specific outcome was downgraded by one level for each of the factors that was encountered: 1) limitations due to study design (>25% of the included studies with a high risk of bias), 2) inconsistency of results [significant statistical heterogeneity (I^2 >40%) or inconsistent findings between the studies (≤75% of the participants report findings in the same direction)], 3) indirectness of evidence (factors affecting the generalizability of results), 4) imprecision (total number of participants <300 for each outcome), and 5) other items (e.g. reporting/publication bias, flawed design). The quality of evidence is considered to be high when RCTs with low risk of bias provide consistent, generalizable and precise results for a particular outcome [27]. Two review authors (YK,AV) scored the levels of evidence. The following levels of the quality of the evidence were applied:

- *High quality:* Further research is very unlikely to change the confidence in the estimate of the effect.
- *Moderate quality:* Further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate.
- *Low quality:* Further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change it.
- *Very low quality:* Great uncertainty about the estimate.

RESULTS

Results of the search and description of studies

Searching the databases resulted in 13,167 references (Figure A). After screening on title and abstract, 32 references remained. Then, screening the full-text article excluded 17 references, leaving 15 references for inclusion [11, 14-16, 28-38]. Three RCTs were published twice [15, 28, 35-38] and one trial had three different publications [11, 33, 34]. Although the DAMASK trial had 6 publications [14, 39-43] only one [14] met the inclusion criteria. One Damask publication [40] presented the trial protocol and was used for the risk of bias assessment. One of the articles [15] reported the results of two trials and was therefore regarded as two separate trials.

Finally, 10 trials were included in the analysis and their characteristics are presented in Table A.

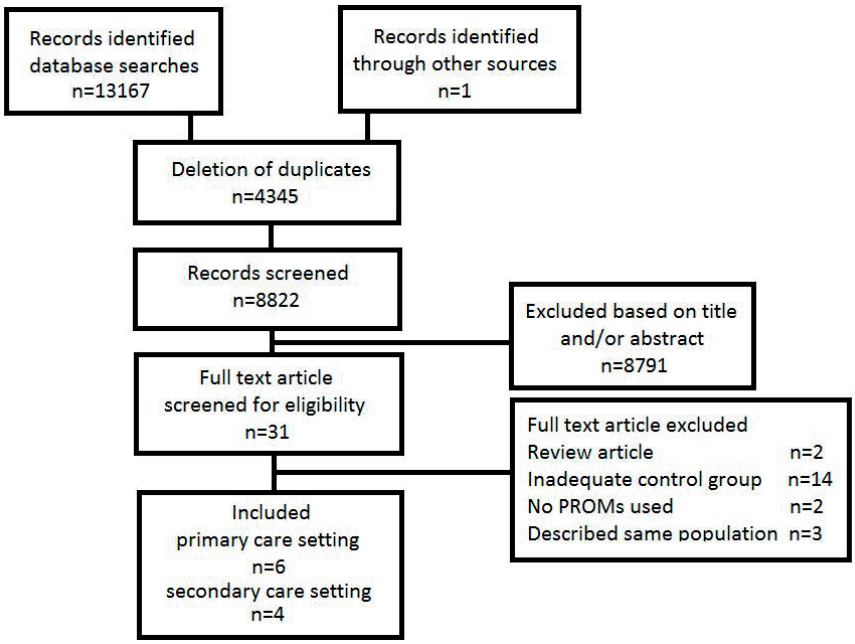


Figure A. Flow diagram

Population

The 11 trials included a total of 2,777 patients (ranging from 50-782 patients per trial); mean age ranged from 28-52 years. Seven trials included a population with acute or subacute low back pain [28-33, 35] and four trials included patients with knee complaints [14-16]. One low back pain study did not report measures of variability and was not pooled in the analysis [30].

Table A. Study Characteristics.

Study	Participants	Interventions	Outcomes	Follow-up	Notes
Ash et al and Modic et al [28,38]	Patients with acute LBP and/or radiculopathy recruited from primary care, Mean age: 43 years; 29% male, Country: US; n=246	1. MRI results provided within 48 hours (n= 131) 2. Patient and physician blinded to MRI imaging results (n= 115).	Pain (SF-36 Bodily pain (0-100)) Function (SF-36 Physical functioning (0-100)) Function (RDQ (0-24)) Pain (VAS (0-10)) absenteeism, self- efficacy scores FABQ (0-24)	2, 4, 6, 8 weeks & 6, 12 and 24 months	All patients received treatment recommendations from guidelines. Patients and physicians in control group were told the imaging results after 6 months
Brealy, DAMASK-trial [14,40]	Patients consulting their GP about continuing knee problems, Mean age: 40 years; 64% male, country: UK; n=553	1. MRI + provisional orthopedic referral (n= 279) 2. Orthopadic referral alone (n= 274)	SF-36 Bodily pain (0-100) Function (SF-36 Physical functioning (0-100)) Function (KQoL-26 (0-100)), Qol (EQ-5D (-0.59-1))	6, 12 and 24 months	112 patients in MRI group had arthroscopy after MRI versus 77 patients in the control group. Whether patients in the control group received imaging was not reported
Bryan (a) et al [15,37]*	Patients with knee complaints attending the orthopedic clinics who were considered for arthroscopy, Mean age: 36 years; 65% male, country: UK; n=118	1. MRI (n= 59) 2. No-MRI (n= 59) Trial arms were immediately listed for arthroscopy, reviewed in clinic (both before and after surgery)	QoL (EQ-5D (-0.59-1)), Pain (SF-36 Bodily pain (0-100)) Function (SF-36 Physical functioning (0-100))	6, 12 and 24 months	9 patients in MRI group did not receive an MRI, and 41% received surgery versus 71% in the control group. Surgery was mainly arthroscopy.
Bryan (b) et al [15,37])*	Patients with acute knee injuries attending the accident and emergency department, Mean age: 28 years; 70% male, country: UK; n=120	1. MRI reviewed with management plan (n= 57) 2. MRI results blinded from clinical management team (n= 63)	QoL (EQ-5D (-0.59-1)), SF-36 Bodily pain (0-100) SF-36 Physical functioning (0-100) Lysholm score (0-25)	6, 12 and 24 months	If patients in the control group still had problems after 6 weeks the protocol allowed for the MRI scan to be reviewed. 30% of MRI patients had surgery compared to 24% of no-MRI patients

Table A. Study Characteristics. (continued)

Study	Participants	Interventions	Outcomes	Follow-up	Notes
Cohen et al [29]	Patients with LBP; clinical candidates, referred by GP, for an epidural steroid injection (ESI), Mean age: 52 years; 66% male, country: US; n=132	1. Physical examination and MRI (n= 67) 2.Physical examination only (n=65)	Leg pain score (0-10), Pain (NRS (0-10)), Function (ODI (0-100%)), GPE, Medication reduction	1 month & 3 months for patients with positive results from ESI	
Deyo et al [30]	Patients with LBP presenting to the walk-in clinic of a public hospital. Mean age: 33; 76% male, country: US; n=101	1. Radiography (n= 43) 2. Education, if no improvement was seen after 3 weeks radiography was done (n= 49)	Pain (1-6) Patient satisfaction (9-27) Function (SIP (0-100)) Self-rated improvement	3 weeks and 3 months	After 3 months 29% of control group had received lumbar radiography
Djais and Kalim [31]	Patients with acute LBP attending rheumatology department, Median age: 40; 55% male, country: Indonesia, n=101	1. Radiography (n=51) 2. Usual care without radiography (n=50)	Function (RDQ (0-24)) Pain (VAS (0-10)) QoL (EQ-5D (-0.59-1)) Overall improvement	3 weeks	Drop-out rate 28%
Gilbert et al [32]	Patients with lumbar spine disorder referred by their GP to an orthopaedic specialist or neurosurgeon, Mean age: 43; 49% male, country: UK; n=782	1. CT or MRI (n=393) 2. Delayed imaging (no MRI or CT unless a clear clinical indication developed) (n=389)	Function (ALBP (0-100)) EQ-5D (-0.59-1)) Pain (SF-36 Bodily Pain (0-100)) Function (SF-36 Physical Functioning (0-100))	8 and 24 months	After 24 months 30% of the control group had received imaging
Kendrick et al and Miller et al [11,33,34]	Patients with LBP consulting the GP, Median age: 39; 87% male, country: UK; n=421	1. Radiography (n= 210) 2. Usual care without radiography (n= 211)	Function (RDQ (0-24)) Pain (VAS (0-10)) QoL (EQ-5D (-0.59-1)) Satisfaction (9-27)	3 and 9 months	12% of patients in the control group had radiography Over 78% drop-out on satisfaction

Table A. Study Characteristics. (continued)

Study	Participants	Interventions	Outcomes	Follow-up	Notes
Kerry et al [35, 36]	Patients consulting their general practitioner with LBP, Mean age: 44; 51% male, country: UK; n=153	1. Referral for radiography (n= 73) 2. Usual care without referral for lumbar radiography (n= 80)	Function (RDQ (0-24)) HADS Pain (SF-36 Bodily pain (0-100)) SF-36 Physical functioning (0-100) QoL (EQ-5D (0-100))	6 weeks and 1 year	Baseline differences, the intervention group had a longer duration of complaints and a higher score on the SF-36. Drop-out >20% After 12 months 14% of the control group had received radiography
Patel et al [16]	Patients with acute knee complaint due to a twisting injury Mean age: 30; 72% male, country: UK; n=50	1. MRI within 2 weeks (n= 25) 2. Conventional management with physiotherapy (n= 25)	Pain (VAS (0-10)), Function (Activity limitation (0-10)) Satisfaction (0-10)	2 weeks and 3 months	1 person in the control group opted for an MRI

LBP Low Back Pain, MRI Magnetic Resonance Imaging, NRS Numeric Rating Scale, GPE Global Perceived Effect, SF-36 Short Form 36 item, Kqol-26 Knee Quality of Life 26 item, QoL Quality of Life, EQ-5D EuroQol 5 dimensions, SIP Sickness Impact Profile, VAS Visual analog scale, ALBP Aberdeen Low Back Pain score, FABQ Fear Avoidance Beliefs Questionnaire GP General Practitioner, HADS Hospital Anxiety and Depression Scale, US United States, UK United Kingdom

* reported data extracted from figures without the actual numbers

Seven trials were performed in the UK [14-16, 32, 33, 35], three in the USA [28-30] and one in Indonesia [31]. The study setting was either primary [14, 28, 29, 32, 33, 35] or secondary health care [15, 16, 30, 31]. Four trials specified the duration of complaints in their inclusion criteria; this ranged from ≤ 1 week to 12 weeks [11, 16, 28, 31].

Interventions

Six trials used MRI as the diagnostic imaging procedure [14-16, 28, 29], one of these used either CT or MRI [32] and four trials used radiography [30, 31, 33, 35]. Five trials compared immediate or early imaging with usual care [16, 30-33]. Four had a control group that could receive imaging based on the usual care trajectory, and two of these trials reported a waiting time for imaging ranging from 29 days [15] to 12 weeks [14]. One trial provided MRI results to the intervention group within 48 h while the control group was blinded to the MRI results [28]. Two trials [15] compared arthroscopy with MRI and arthroscopy alone and in one trial [29] all patients received an epidural steroid injection either based on history and physical examination, or on clinical findings and imaging results.

Five trials [14, 30, 32, 33, 35] reported the percentage of patients receiving imaging in the control group (ranging from 2-30%) as part of usual care.

Outcome measures

All trials assessed both pain and function (Table 1). Five trials examined pain with the Bodily Pain score of the Short Form 36 (SF-36). Four trials reported pain with a visual analogue scale (VAS) or the numerical rating scale (NRS) and another trial rated pain on a six-point scale. To assess function, both generic and disease-specific measurement instruments were used. Disease-specific measurement instruments were the Roland Disability Index, the Aberdeen Low Back Pain Score, the Oswestery Disability Index, the Lysholm score, and the Knee Quality of Life Questionnaire.

Two trials reported median scores and interquartile ranges [31, 33]. For pain and function, all outcome measures were continuous. In five trials overall improvement was measured as dichotomous. Two of these five trials reported satisfaction on an ordinal scale and two on a continuous scale, of which one had a 78% dropout rate on this outcome and was excluded from the analysis (fatal flaw) [33]. Only the results for dichotomous outcomes were pooled.

For two trials we contacted the authors for additional information. For one [33] of these trials we received information from the author about a systematic review including this trial [13, 33]. Another trial did not report data to impute SDs [30]; unfortunately, we did not receive any response from these authors. Because one article [15] only reported data in figures, the data were estimated from these figures.

Risk of bias assessment

Six trials (55%) were considered to have low risk of bias [14, 15, 28, 29, 33]. Overall, risk of bias was threatened by the inability to blind patients (n=10), care providers (n=11) or outcome assessors (n=10), and by the absence of an intention-to-treat analysis (n=8). Concealment of randomization was not adequately reported in three trials. The results of the risk of bias assessment are presented in (supplementary material).

Effects of imaging

All effects estimates are described in the summary of findings (Table B). GRADE scoring is reported for short and long-term follow-up and (separately) for low back pain and knee studies. Only subgroup results are reported when heterogeneity was high for the overall effect estimate.

Pain. Figure B.1-2. shows the improvement in pain on short and long-term follow-up.
Pooling the studies with low back pain patients resulted in a significant effect in favor of no imaging on the short [SMD 0.17 (95%CI: 0.04-0.31)] and long term [SMD 0.13 (95%CI: 0.02-0.24)] but the effect size was below 0.2, while the trials with patients with knee complaints found no difference on the long term [SMD 0.02 (95%CI: -0.14-0.18)]. In the short-term analysis only one study with knee complaints had available results on pain; these results indicated a non-significant effect in favor of imaging (Figure B.1). Heterogeneity was small (I^2=39%) at short-term follow-up and not present at long-term follow-up. When all trials were pooled, no significant and clinically relevant differences were found on the short term [SMD 0.10 (95%CI: -0.08-0.29)]. On long-term follow-up data showed borderline significant results in favor of no imaging [SMD 0.09 (95%CI: 0.00-0.18)] but the effect size remained below 0.2.

In the short-term analysis there were four studies and in the long-term analysis five studies with a primary care population. Effects sizes for both the short term [SMD 0.15 (95%CI: 0.01-0.30)] and long term [SMD 0.11 (95%CI: 0.01-0.20)] resulted in borderline significant effects in favor of no imaging but the effect size was below 0.20.

Pooling only the trials using radiography (n=3) as imaging method resulted in a significant effect in favor of no imaging but a SMD below 0.2 [SMD 0.15 (95%CI: 0.03-0.26)], whereas pooling the trials with MRI (n=8) found no difference [SMD 0.07 (95%CI: -0.05-0.18)] (data not shown).

Overall, we found moderate level of evidence (downgraded based on limitations in study design) for a small clinically irrelevant effect on pain in favor of no imaging on the long term, especially for the low back pain trials and trials using radiography.

Table B. Summary of findings table.

Patient recovery after diagnostic imaging in patients with musculoskeletal disorders
Population: patients with musculoskeletal complaints Intervention: Diagnostic imaging (radiography/MRI) Setting: Primary care/secondary care emergency department/orthopaedic department. Comparison: Usual care/not getting immediate imaging/not receiving results of imaging

Outcomes	Standardized mean difference (95% Confidence Interval)	Number of participants (studies)	Quality of evidence (GRADE score)	Comments
Pain short-term	0.10 (-0.08; 0.29)	890 (6 studies)	Moderate level [2]	Two studies the median was used as mean and SD was calculated from IQR
LBP studies	0.17 (0.04; 0.31)	814 (5 studies)	Low level [1, 2]	
Knee pain studies	-0.51 (-1.10 to 0.07)	46 (1 study)	Very low level [1, 2, 3]	
Primary care studies	0.15 (0.01 to 0.30)	768 (3 studies)	Moderate level [2]	
Pain long term	0.09 (0.00; 0.18)	1875 (7 studies)	Moderate level [2]	One study only reported 24 months follow-up, another only figures and for 1 study the median was used as mean and SD was calculated from IQR
LBP studies	0.13 (0.02; 0.24)	1281 (4 studies)	Moderate level [2]	
Knee pain studies	0.02 (-0.14; 0.18)	594 (3 studies)	High level	
Primary care studies	0.11 (0.01 to 0.20)	1752 (5 studies)	Moderate level [2]	
Function gen. short term	-0.21 (-0.55; 0.12)	348 (3 studies)	Low level [1, 2]	Two studies the median was used as mean and SD was calculated from IQR
LBP studies	-0.12 (-0.49; 0.25)	302 (2 studies)	Low level [1, 2]	
Knee pain studies	-0.57 (-1.16; 0.02)	46 (1 study)	Very low level [1, 2, 3]	
Primary care studies	-0.12 (-0.49 to 0.25)	302 (2 studies)	Low level [1, 2]	
Function gen. long term	0.08 (-0.05; 0.20)	1481 (6 studies)	Moderate level [2]	One study only reported 24months follow up, another only figures and for 1 study the median was used as mean and SD was calculated from IQR
LBP studies	0.10 (-0.03; 0.23)	887 (3 studies)	Moderate level [2]	
Knee pain studies	-0.07 (-0.44; 0.31)	594 (3 studies)	Moderate level [1]	
Primary care studies	0.13 (0.02 to 0.24)	1358 (4 studies)	Moderate level [1]	
Function spec. short-term	0.11 (-0.04; 0.27)	844 (5 studies)	Moderate level [2]	Two studies the median was used as mean and SD was calculated from IQR
Primary care studies	0.09 (-0.06 to 0.23)	678 (4 studies)	Moderate level [2]	
Function spec. long-term	0.01 (-0.34; 0.41)	1281 (4 studies)	Low level [1, 2]	One study only reported 24months follow up, another only figures and for 1 study the median was used as mean and SD was calculated from IQR
tLBP studies	0.04 (-0.38; 0.45)	525 (2 studies)	Moderate level [1]	
Knee pain studies				

Table B. Summary of findings table. (continued)

Outcomes	Standardized mean difference (95% Confidence Interval)			Number of participants (studies)	Quality of evidence (GRADE score)	Comments
QoL short-term LBP studies	-0.00 (-0.07; 0.06)			202 (2 studies)	Very low level [1, 2, 3]	
QoL long-term	0.01 (-0.12; 0.10)			1270 (7studies)	Moderate level [2]	
LBP studies	0.03 (-0.09; 0.14)			1143 (5 studies)	Moderate level [2]	
Knee studies	-0.18 (-0.54; 0.18)			123 (2 studies)	Moderate level [3]	
Primary care studies	-0.03 (-0.14 to 0.09)			1143 (3 studies)	Low level [1, 2]	
	Assumed mean score/ comparative risk (95% CI)		Relative Effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)	
	Imaging	No imaging				
Overall improvement LBP studies	128 of 343 recovered	151 of 335 recovered	RR 1.15 (1.03; 1.28)	855 (5 studies)	Moderate level [2]	
Satisfaction	66 of 150 satisfied	73 of 152 satisfied	RR 1.03 (0.85; 1.24)	302 (2 studies)	Low level [1, 2]	

[1] Downgraded because of inconsistency

[2] Downgraded because of limitations in study design

[3] Downgraded because of imprecision

SD standard deviation, CI confidence interval, IQR interquartile range, LBP Low Back Pain, Gen Generic, Spec Specific, QoL Quality of Life

Figure B.1.

Pain intensity short-term

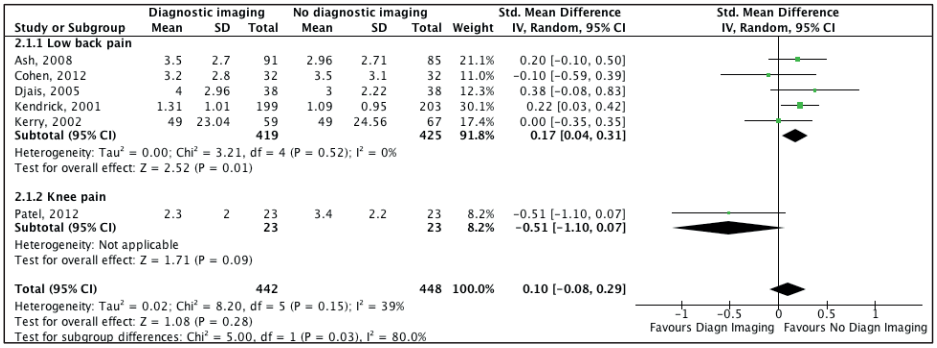


Figure B.2.

Pain intensity long-term

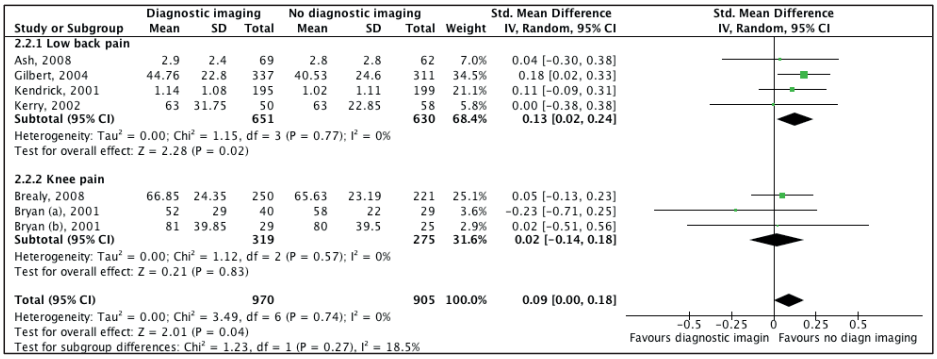


Figure B. Pain intensity long-term and short-term

Function. Figure C.1-4. shows the improvement in function measured with generic and specific measurement instruments for short and long-term follow-up.

Heterogeneity was present at short-term outcome (I^2=55%) and small for long-term outcome. Subgroup analysis for patients with low back pain had non-significant differences at short term [SMD -.021 (95%CI: -0.55-012)] and long term [SMD 0.10 (95%CI: -0.03-0.23)]. Trials with knee complaints were only available for the long-term results and showed a non-significant effect in favor of imaging [SMD -0.07 (95%CI: -0.44-0.31)]. The overall effect estimate for knee and low back pain studies combined at long term found no effects and were not significant [SMD 0.08 (95%CI: -0.05-0.20)].

In the short-term analysis there were two studies and in the long-term analysis four studies with a primary care population. Effects sizes for the short term [SMD -0.12 (95%CI: -0.49-0.25)] were not significant. Long-term analysis resulted in a small borderline significant effect [SMD 0.13 (95%CI: 0.02-0.24)] in favor of no imaging but a SMD below 0.2.

Excluding the only trial using radiography as a method of imaging resulted in a non-significant effect estimate in the MRI subgroup [SMD -0.08 (95%CI: -0.27-0.11)] (data not shown).

We found low level evidence (downgraded based on limitations in study design and inconsistency) that there is no difference on the short term and moderate level of evidence (downgraded based on study design and inconsistency) on the long term for function measured with generic measurement instruments.

Figure C3-4 shows improvement in function with disease-specific instruments. Heterogeneity was very small for the short term (I^2=16%) because no trials with knee complaints were available. Substantial heterogeneity was present at long-term follow-up (I^2=70%). Both outcome measures are reported per subgroup. Subgroup analysis for low back pain trials resulted in a non-significant effect on the short term [SMD 0.11 (95%CI: -0.04-0.27)] and long term [SMD 0.01 (95%CI: -0.23-0.25)].

The short-term analysis included four studies with a primary care population; pooling these studies did not significantly alter the effect size [SMD 0.09 (95%CI: -0.06-0.23)]. All studies in the long-term analysis were primary care populations.

Pooling studies with knee complaints resulted in a non-significant effect [SMD 0.04 (95% CI -0.38; 0.45)].

No differences were found [SMD 0.01 (95%CI: -0.19-0.21)] when analysing trials using MRI (n=6). Pooling trials using radiography (n=3) resulted in a borderline significant difference [SMD 0.13 (95%CI: -0.00-0.25)] in favor of the no imaging group (data not shown) but the SMD was below 0.2. Separate analyses for primary care studies were not possible because of the small number of available studies.

We found moderate level of evidence (downgraded based on limitations in study design) for no differences between both groups at short-term follow-up for patients with low back pain, and low level of evidence (downgraded based on limitations in study design and inconsistency) that there is no difference between imaging and no imaging for disease-specific function at long-term follow-up, irrespective of the subgroups. Subgroup analysis found a small borderline significant effect in favor of the no imaging group in trials using radiography.

Satisfaction. Moderate heterogeneity was present (I^2=44%). Because of the limited number of trials the short and long-term results were combined (data not shown). Overall, we found a low level of evidence for no differences (downgraded based on limitations in study design and inconsistency) between the groups [RR 1.03 (95%CI: 0.85-1.24)].

Figure C.1.

Function measured with generic instruments short-term

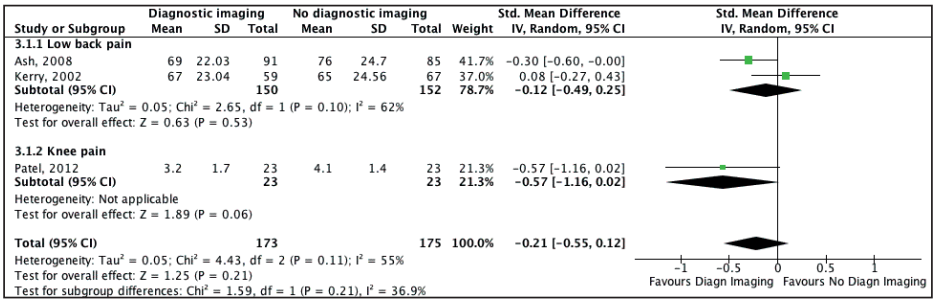


Figure C.2.

Function measured with generic instruments long-term

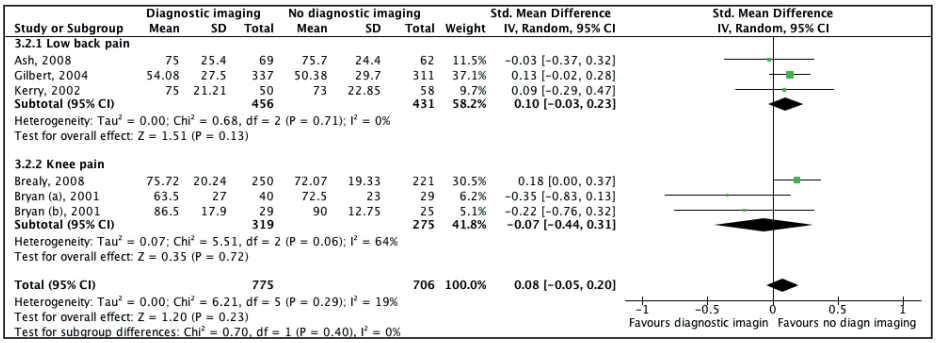


Figure C.3.

Function measured with disease specific instruments short-term

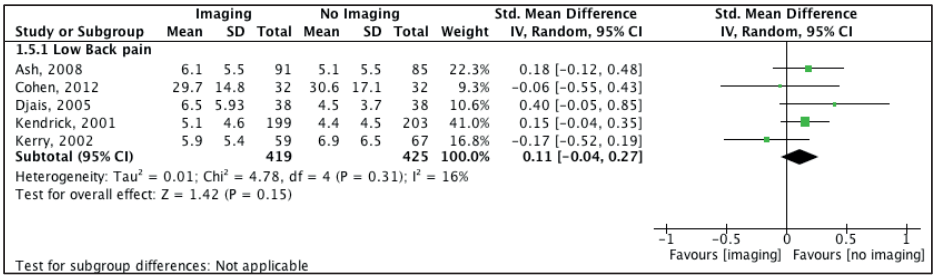


Figure C.4.

Function measured with diseases specific instruments long-term

Figure C. Function short-term & long-term

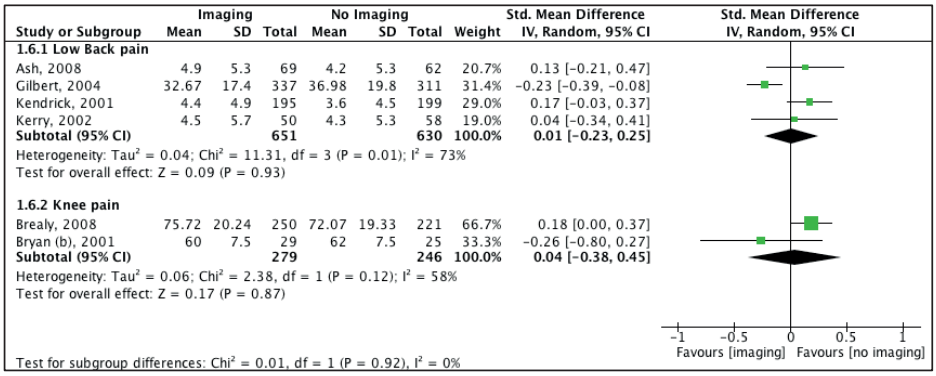


Figure C.1-4.

Function short-term & long-term

Figure C. Function short-term & long-term *(continued)*

4

Quality of Life. Figure D shows the results of 'quality of life' for the short and long-term follow-up.

Substantial heterogeneity was present at short-term (I^2=86%) but not at long-term follow-up (figure D). For the short-term pooled effect estimate, only two low back pain studies were available [SMD -0.07 (95%CI: -0.83-0.68)]. Subgroup analysis for the long-term effect resulted in slightly different non-significant effects between knee [SMD 0.18 (95%CI: -0.18-0.54)] and low back pain studies [SMD -0.03 (95%CI: -0.14-0.09)]. The overall effect at the long term showed no difference for knee and low back pain studies combined [SMD 0.01 (95%CI: -0.10-0.12)]. Pooling the studies performed in primary care did not significantly alter the effect size [SMD -0.03 (95%CI: -0.14-0.09)]. Overall, low level of evidence (downgraded because of limitations in study design and inconsistency) was found for no difference concerning quality of life for patients with knee pain and with low back pain at short-term follow-up and moderate level of evidence at long-term follow-up.

Overall improvement. Figure E shows the results of 'overall improvement'.

Short and long-term results were combined due to the limited number of trials reporting overall improvement. No studies with knee pain presented results for overall improvement. Heterogeneity was not present. Overall improvement showed a significant but clinically irrelevant result in favor of the no imaging group (RR 1.15, 95%CI: 1.03-1.28). Sensitivity analysis showed that excluding two trials [30, 31] with high risk of bias did not change the results (RR 1.13, 95%CI: 1.01-1.27).

Four studies were performed in primary care; pooling these studies did not alter the results (RR 1.15, 95%CI: 1.03-1.28).

Figure D.1.

Quality of Life short-term

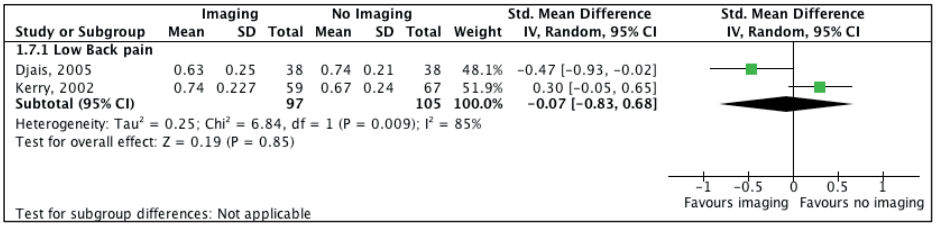


Figure D.2.

Quality of life long-term

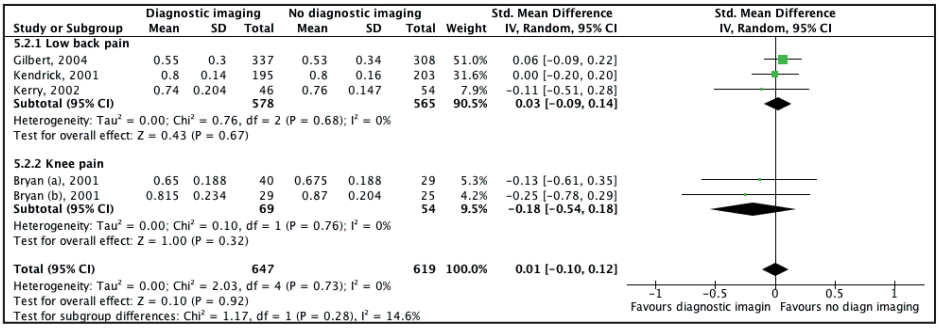


Figure D. Quality of life short- & long-term

We found a moderate level of evidence (downgraded because of limitations in study design) for a small effect in favor of no imaging concerning overall improvement for patients with low back pain.

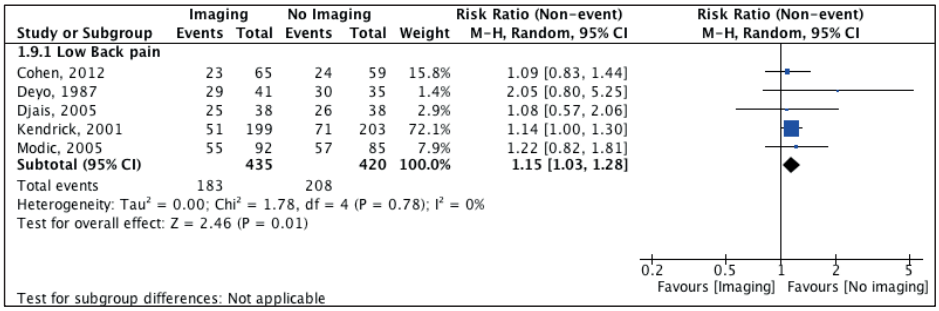


Figure E. Overall improvement long term

DISCUSSION

Overall, our results showed that early imaging strategies do not improve patient-reported outcomes (PROMs) in patients with low back pain or knee complaints. Small differences were found between these subgroups in pain, function and quality of life, in which the low back pain subgroup usually had larger effect sizes in favor of "no imaging". Notably, more trials concerning low back pain were available. Subgroup analysis in low back pain patients led to small significant effect in favor of no routine imaging. The majority of imaging tests used in low back pain show an absence of abnormality; however, this may not reassure patients and can lead to possible negative effects of imaging.

Strengths and limitations

An important strength of the present review is the sensitive search strategy applied to reduce the chance of missing relevant studies and thereby reducing publication bias. In a search strategy, defining 'diagnostic imaging' as the intervention appeared to be somewhat difficult. However, because of the sensitive search strategy (including all possible synonyms) it is unlikely that relevant trials were missed.

We aimed to include patients with all kinds of musculoskeletal disorders, acute as well as subacute, or chronic complaints. Subgroup analysis was possible for patients with knee or low back pain, and no significant differences between these subgroups were found (except for pain). The populations with knee complaints were mostly traumatic or acute knee complaints, probably having a different clinical course than that of low back pain; this might explain the differences found regarding pain. Subgroup analysis for primary care studies did not alter the results.

Another source of heterogeneity could arise from the different types of imaging modalities used in the trials. Subgroup analysis showed a significant difference between imaging modalities for radiography on pain and 'borderline' significant difference on function measured with disease-specific instruments; however, this might be based on chance and here no differences between knee and low back pain were made.

Overall, 45% of the trials scored high risk of bias. Differences in study design could have caused heterogeneity. Only one trial [29] was able to blind their patients for the allocated intervention. Given the nature of the intervention and the clinical setting of most of the trials, blinding of patients was difficult and might have caused some bias [49, 50].

Standardization of treatment after imaging was underreported and could clearly account for bias in the study results. All studies reported that some sort of treatment was provided after imaging. Treatment might have influenced the outcome of interest by increasing or decreasing the contrast between the groups. Two trials [14, 15] even reported having arthroscopy of the knee after imaging, and another study [29] used imaging in the experimental group who received an epidural steroid injection, thereby biasing the

effect of imaging on the outcome. Although these co-interventions will affect outcome, the effect of imaging is to influence treatment decisions such as these. In the future, these treatment decisions will differ between intervention and control groups. Furthermore, usual care was hardly described and might differ between different countries. In several trials the control group could also receive diagnostic imaging as part of usual care, thereby decreasing the contrast between the intervention and control treatment. Although including trials performed in a clinical setting can increase generalizability, it can also affect the validity and reliability of the results.

Because all trials used valid PROMs, pooling of results was possible. Disease-specific and generic measurement instruments for function were pooled separately. Generic instruments tended to be in favor of imaging at short term, while the results generated by the disease-specific instruments tended to find no differences. In contrast, disease-specific instruments might be more responsive compared to generic instruments, or the results might be attributed to measurement error [51, 52]. All outcome measures were patient reported outcome measures. Whether treatment regime changed due to the "intervention" is unknown.

The fact that all trials excluded patients suspected of having a serious underlying condition shows the effect for diagnostic imaging in patients were its still uncertain whether it may have a favorable effect. All trials, but one [18] (who excluded one patient with malignancy), did not report finding any serious underlying conditions.

In clinical trials the patient population is selected using strict selection criteria, which also hampers generalizability of the results. In the present review, caution is needed when drawing conclusions because of the small number of studies in the subgroup analysis, the considerable amount of risk of bias, and the diversity of the study settings.

Comparison with existing literature

Results from our review are comparable with those from an earlier review [13], although we found a clearer tendency towards benefit of no imaging for low back pain patients. Another review [18] studied the effect of diagnostic imaging on reassurance, and included five trials with populations also having chest pain or headache. The latter review included two trials with musculoskeletal complaints, which were also included in our review.

Implications

In patients with a musculoskeletal disorder, imaging did not lead to better PROMs. On the contrary, some results showed a tendency towards better outcomes after no routine imaging. Other factors, like exposure to radiation, increasing costs, and use of unnecessary invasive procedures, might also influence the clinical benefit for patients.

Imaging has its place in health care where serious conditions are suspected or when surgery is considered. The natural history of low back pain is benign, as 90% of patients

recover within 6 weeks, and resolves with little intervention without knowing the anatomic diagnosis [53]. In the first 6 weeks diagnostic imaging should be used in the presence of red flags (smoking, age, history of cancer, diabetes, drug abuse, chronic NSAID, unnatural course of pain, night pain or symptoms of cauda iquina) [54, 55]. Patient with complaints longer than 6 weeks, diagnostic imaging does not necessarily disclose clear pathologic diagnoses [56]. Degenerative findings are common and whether these findings can attribute to the complaints remains unknown. It seems that the results might be limited by our current inability to understand this complex multifactorial condition and future research should focus on the ability to diagnose the condition.

Patients with knee complaints reported 25% recovery after 3 months and 44% after 12 months in primary care [57]. Urgent referral to a specialist is necessary when there are signs of fracture, acute locked knee or severe pain after patella dislocation at the initial consultation that is likely to be attributed to a trauma [58]. Several clinical decision rules are validated that identify patients with a high risk of fracture [59-61]. Diagnostic imaging can also be helpful in establishing the correct diagnosis in non-traumatic knee complaints. In order to prevent excessive imaging, especially the number of images without pathology, patients should be managed conservatively and imaging should be considered when patients show no improvement [58]. According to the America College of Radiography the initial imaging study for non-traumatic knee pain should be radiography [62]. MRI is needed to further examine intra-articular abnormalities (lesions of ligaments, tendons, bone, cartilage and menisci) [63-65].

Future research should focus on trials with low risk of bias, paying special attention to standardization and blinding of trial participants. Also, future trials should try to 'prevent' patients in the usual care group from receiving any type of imaging. Furthermore, reporting the effect of clinical decisions (e.g. the number of patients having surgery or therapy) in the long term is required to study the clinical impact of imaging.

CONCLUSIONS

Routine diagnostic imaging in patients with low back pain or knee complaints did not change the outcome for pain, function, quality of life, recovery nor satisfaction. In patients with low back pain routine imaging may even cause some harm. Our results indicate that it is unlikely that use of routine diagnostic imaging in all patients leads to better patient-reported outcome measures. Imaging has its place in health care where serious conditions are suspected or when surgery is considered. Diagnostic imaging can be considered in patients with low back pain to rule out a serious underlying condition in the presence of red flags and in subactute/chronic low back pain patients who show no improvement. Clinical decision rules should be used by clinicians in patients with

traumatic knee complaints. In non-traumatic knee complaints diagnostic imaging should be used if conservative treatment fails.

Caution is required when drawing conclusions, due to the small number of studies with heterogeneity in patient populations and the presence of risk of bias in a considerable percentage of the studies.

Learning points

Evidence from trials comparing routine diagnostic imaging with usual care or no imaging has yielded conflicting results.

Results from this review show small significant effects on pain and overall improvement, especially for patients with low back pain, not in favour of imaging. No different effects after receiving diagnostic imaging were found among patients with knee pain.

REFERENCES

1. Bussières, A., Peterson, C., Taylor, J.A. (2007). Diagnostic imaging practice guidelines for musculoskeletal complaints in adults an evidence-based approach: introduction. *Journal of Manipulative and Physiological Therapeutics*, 30(9), 617-83.
2. Baker, L., Atlas, S.W., Afendulis, C.C. (2008). Expanded use of imaging technology and the challenge of measuring value. *Health Affairs*, 27(6), 1467-78.
3. Fisher, E., Welch, H.G. (1999). Avoiding the unintended consequences of growth in medical care: how might more be worse? *Journal of the American Medical Association*, 281(5), 446-53.
4. Moyniham, R., Doust, J., Henry, D. (2012). Preventing overdiagnosis: how to stop harming the healthy. *British Medical Journal*, 344, e3502.
5. Wassenaar, M., van Rijn, R.M., van Tulder, M.W., Verhagen, A.P., van der Windt, D.A.W.M., Koes, B.W., de Boer, M.R., Ginai, A.Z., Ostelo, R.W.J.G. (2012). Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic review. *European Spine Journal*, 21(2), 220-227.
6. van Rijn, R., Wassenaar, M., Verhagen, A.P., Ostelo, R.W., Ginai, A.Z., de Boer, M.R., van Tulder, M.W., Koes, B.W. (2012). Computed tomography for the diagnosis of lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic review. *European Spine Journal*, 21(2), 228-239.
7. Boden, S., Davis, D.O., Dina, T.S., Patronas, N.J., Wiesel, S.W. (1990). Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects: A prospective investigation. *The Journal of Bone and Joint Surgery American Volume*, 72(3), 403-8.
8. Jensen, M., Brant-Zawadzki, M.N., Obuchowski, N., Modic, M.T., Malkasian, D., Ross, J.S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *The New England Journal of Medicine*, 331(2), 69-73.
9. Lurie, J., Birkmeyer, N.J., Wernstein, J.N. (2003). Rates of advanced spinal imaging and spine surgery. *Spine*, 28(6), 616-20.
10. Maniadakis, N., Gray, A. (2000). The economic burden of back pain in the UK. *Pain*, 84(1), 95-103.
11. Miller, P., Kendrick, D., Bentley, E., Fielding, K. (2002). Cost-effectiveness of lumbar spine radiography in primary care patients with low back pain. *Spine*, 27(20), 2291-7.
12. Oei, E., Nikken, J.J., Ginai, A.Z., Krestin, G.P., Verhaar, J.A., van Vugt, A.B., Hunink, M.G. (2009). Costs and effectiveness of a brief MRI examination of patients with acute knee injury. *European Radiology*, 19(2), 409-18.
13. Chou, R., Fu, R., Carrino, J.A., Deyo, R.A. (2009). Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet*, 373(9662), 463-72.
14. Brealy, S., DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. (2008). Effectiveness of GP access to magnetic resonance imaging of the knee: a randomized trial. *British Journal of General Practice*, 58(556), e1-e9.
15. Bryan, S., Weatherburn, G., Bungay, H., Hatrick, C., Salas, C., Parry, D., Field, S., Heatley, F. (2001). The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint. *Health Technology Assessment*, 5(27), 1-95.
16. Patel, N., Bucknill, A., Ahearne, D., Denning, J., Desai, K., Watson, M. (2012). Early magnetic resonance imaging in acute knee injury: a cost analysis. *Knee Surgery,*

Sports Traumatology, Arthroscopy, 20, 1152-1158.

17. McDonald, I., Daly, J., Jelinek, V.M., Panetta, F., Gutman, J.M. (1996). Opening Pandora's box: the unpredictability of reassurance by a normal test result. *British Medical Journal*, 313(7053), 329-32.
18. Van Ravesteijn, H., van Dijk, I., Darmon, D., van de Laar, F., Lucassen, P., Hartman, T.O., van Weel, C., Speckens, A. (2012). The reassuring value of diagnostic tests: a systematic review. *Patient Education and Counseling*, 86(1), 3-8.
19. Verhagen, A., de Vet, H.C., de Bie, R.A., Kessels, A.G., Boers, M., Bouter, L.M., Knipschild, P.G. (1998). The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of Clinical Epidemiology*, 51(12), 1235-41.
20. Verhagen, A., de Vet, H.C., de Bie, R.A., Boers, M., van den Brandt, P.A. (2001). The art of quality assessment of RCTs included in systematic reviews. *Journal of Clinical Epidemiology*, 54(7), 651-4.
21. Tulder, M.V., Suttorp, M., Morton, S., Bouter, L.M., Shekelle, P. (2009). Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine*, 34(16), 1685-92.
22. Higgins, J., Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539-1558.
23. Cohen, J. *Satistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.
24. Members, P.P. (2001). Philadelphia Panel Evidence-Based Clinical Practice Guidelines on Selected Rehabilitation Interventions: Overview and Methodology. *Physical Therapy*, 81(10), 1629-40.
25. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011.
26. Guyatt, G., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünneman, H.J. (2008). GRADE: an emerging consensus on ratin quality of evidence and strength recommendations. *British Medical Journal*, 336(7650):924-6.
27. Furlan, A., Pennick, V., Bombardier, C., Tulder, M.V., on behalf of the Editorial Board of the Cochrane Back Review Group. (2008). *Updated method guidelines for systematic reviews in the Cochrane Back Review Group*, C.B.R.G.E. Board Editor.
28. Ash, L., Modic, M.T., Obuchowski, N.A., Ross, J.S., Brant-Zawadzki, M.N., Grooff, P.N. (2008). Effects of diagnostic information, per se, on patient outcomes in acute radiculopathy and low back pain. *American Journal of Neuroradiology*, 29(6), 1098-103.
29. Cohen, S., Gupta, A., Strassels, S.A., Christo, P.J., Erdek, M.A., Griffith, S.R., Kurihara, C., Buckenmaier, C.C., Cornblath, D., Vu, T. (2012). Effect of MRI on Treatment Results or Decision Making in Patients With Lumbosacral Radiculopathy Referred for Epidural Steroid Injections. *Archive of Internal Medicine*, 172(2), 134-42.
30. Deyo, R., Diehl, A.K., Rosenthal, M. (1987). Reducing roentgenography use. Can patient expectations be altered? *Archive of Internal Medicine*, 147(1), 141-5.
31. Djais, N., Kalim, H. (2005). The role of lumbar spine radiography in the outcomes of patients with simple acute low back pain. *APLAR J Rheumatology*, 8, 45-50.
32. Gilbert, F., Grant, A.M., Gillan, M.G., Vale, L., Scott, N.W., Campbell, M.K., Wardlaw, D., Knight, D., McIntosh, E., Porter, R.W. (2004). Does early imaging influence management and improve outcome in patients with low back pain? A pragmatic randomised controlled trial. *Health Technology Assessment*, 8(17), 1-131.
33. Kendrick, D., Fielding, K., Bentley, E., Kerslake, R., Miller, P., Pringle, M. (2001). *Radiography of the lumbar spine in*

primary care patients with low back pain: randomised controlled trial. British Medical Journal, 322(7283), 400-5.

34. Kendrick, D., Fielding, K., Bentley, E., Miller, P., Kerslake R., Pringle, M. (2001). The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial. *Health Technology Assessment*, 5(30), 1-69.
35. Kerry, S., Hilton, S., Dundas, D., Rink, E., Oakeshott, P. (2002). Radiography for low back pain: a randomised controlled trial and observational study in primary care. *British Journal of General Practice*, 52(479), 469-74.
36. Kerry, S., Hilton, S., Patel, S., Dundas, D., Rink, E., Lord, J. (2000). Routine referral for radiography of patients presenting with low back pain: is patients' outcome influenced by GPs' referral for plain radiography? *Health Technology Assessment*, 4(20), 1-119.
37. Bryan, S., Bungay, H.P., Weatherburn, G., Field, S. (2004). Magnetic resonance imaging for investigation of the knee joint: a clinical and economic evaluation. *International Journal of Technology Assessment in Health Care*, 20(2), 222-9.
38. Modic, M., Obuchowski, N.A., Ross, J.S., Brant-Zawadzki, M.N., Grooff, P.N., Mazanec, D.J., Benzel, E.C. (2005). Acute Low Back Pain and Radiculopathy: MR Imaging Findings anf Their Prognostic Role and Effect on Outcome. *Radiology*, 237, 597-604.
39. Garratt, A.M., Brealy, S., Gillespie, W.J., DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. (2004). Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford)*, 43(11), 1414-23.
40. Brealy, S., Atwell, C., Bryan, S., Coulton, S., Cox, H., Cross, B., Fylan, F., Garratt, A., Gilbert, F.J., Gillan, M.G.C., Hendry, M., Hood, K., Houston, H., King, D., Morton, V., Orchard, J., Robling, M., Russell, I.T., Torgerson, D., Wadsworth, V., Wilkinson, C. (2006). The DAMASK trial protocol: a pragmatic randomised trial to evaluate whether GPs should have direct access to MRI for patients with suspected internal derangement of the knee. *BMC Health Services Research*, 6, DOI: 10.1186/1472-6963-6-133.
41. Brealy, S., DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. (2007). Influence of magnetic resonance of the knee on GPs' decisions: a randomised trial. *British Journal of General Practice*, 57(541), 622-9.
42. Brealy, S., DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. (2008). Cost-effectiveness of magnetic resonance imaging of the knee for patients presenting in primary care. *Br J Gen Pract*, 58, e10-e16 DOI: 10.3399/bjgp08X342660.
43. Brealy, S., Andronis, L., Dale, V., Gibbon, A.J., Gilbert, F.J., Hendry, M., Hood, K., King, D., Wilkinson, C., DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. (2012). The effect of waiting times from general practitioner referral to MRI or orthopaedic consultation for the knee on patient-based outcomes. *Br J Radiol,* 85, e1134-9 DOI: 10.1259/bjr/12729937.
44. Laisné, F., Lecomte, C., Corbière, M. (2012). Biopsychosocial predictors of prognosis in musculoskeletal disorders: a systematic review of the literature. *Disability and Rehabilitation*, 34(5), 355-82.
45. Ramond, A., Bouton, C., Richard, I., Roquelaure, Y., Baufreton, C., Legrand, E., Huez, J.F. (2011). Psychosocial risk factors for chronic low back pain in primary care–a systematic review. *Family Practice*, 28(1), 12-21.

46. Pincus, T., Burton, A.K., Vogel, S., Field, A.P. (2002). A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine*, 27(5), E109-20.
47. Steenstra, I., Verbeek, J.H., Heymans, M.W., Bongers, P.M. (2005). Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occupational and Environmental Medicine*, 62(12), 851-60.
48. Jarvik, J., Maravilla, K.R., Haynor, D.R., Levitz, M., Deyo, R.A. (1997). Rapid MR imaging versus plain radiography in patients with low back pain: initial results of a randomized study. *Radiology*, 204, 447-54.
49. Schulz, K., Grimes, D.A. (2002). Blinding in randomised trials: hiding who got what. *Lancet*, 359, 696-700.
50. Kent, W. Randomisation and blinding in clinical trials. WebmedCentral CLINICAL TRIALS 2012; 3(5):WMC003289.
51. Malý, M., Vondra, V. (2006). Generic versus disease-specific instruments in quality-of-life assessment of chronic obstructive pulmonary disease. *Methods of Infromation in Medicine*, 45(2), 211-5.
52. Bombardier, C., Melfi, C.A., Paul, J., Green, R., Hawker, G., Wright, J., Coyte, P. (1995). Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. *Medical Care*, 33(4), AS131-44.
53. Pengel, L.H., Herbert, R.D., Maher, C.G., et.al. (2003). Acute low back pain: systematic review of its prognosis. *British Medical Journal*, 327(7410), 323
54. Rinkus, K.M., Knaub, M.A. (2008). Clinical and Diagnostic Evaluation of Low Back Pain. *Seminars in Spine Surgery*, 20, 93-101
55. Wheeler, A.H. (1995). Diagnosis and management of low back pain and sciatica. *Am Fam Phys*, 52, 1333-41
56. Davis, P.C., Wippold, F.J., Cornelius, R.S., Angtuaco, E.J., Broderick, D.F., Brown, D.C., Garvin, C.F., Hartl, R., Holly, L., McConnell, C.T. Jr., Mechtler, L.L., Rosenow, J.M., Seidenwurm, D.J., Smirniotopoulos, J.G., Expert Panel on Neurologic Imaging. (2011). ACR Appropriateness Criteria® low back pain. [online publication]. Reston (VA): American College of Radiology (ACR), 8
57. Van der Waal, J.M., Bot, S.D., Terwee, C.B., Van der Windt, D.A., Scholten, R.J., Bouten, L.M., Dekker, J. (2005). Course and prognosis of knee copmlaints in general practice. *Arthritis and Rheumatism*, 53(6), 920-30
57. Belo, J.N., Berg, H.F., Klein Ikkink, A.J., Wildervanck-Dekker, C.M.J., Smorenburg, H.A.A.J., Draijer, L.W. (2010). Clinical guideline 'traumatic knee complaints' from the Dutch College of General Practitioners (Dutch). *Huisarts en Wetenschap*, 54, 147-158
58] Bauer, S.J., Hollander, J.E., Fuchs, S.H., Thode, H.C. Jr. (1995). A clinical decision rule in the evaluation of acute knee injuries. *Journal of Emerging Medicine*, 14(5), 611-615
59. Seaberg, D.C., Jackson, R. (1994). Clinical decision rule for knee radiographs. *American Journal of Emerging Medicine*, 12(5), 541-543
60. Stiell, I.G., Wells, G.A., Hoag, R.H., et.al. (1997). Implementation of the Ottawa Knee Rule for the use of radiography in acute knee injuries. *JAMA*, 278(23), 2075-2079
61. American College of Radiology ACR. Appropriatness Criteria; nontraumatic knee pain. *Radiology*, 2000; 215(suppl): 1-1511
62. Oei, E.H., Nikken, J.J., Verstijnen, A.C., Ginai, A.Z., Hunink, M.G. (2003). MR imaging of the menisci and cruciate ligaments: a systematic review. *Radiology*, 226(3), 837-848

63. Crawford, R., Walley, G., Bridgman, S., Maffuli, N. (2007). Magnetic resonance imaging versus arthroscopy in the diagnosis of knee pathology, concentrating on meniscal lesions and ACL tears: a systematic review. *British Medical Bulletin*, 84, 5-23
64. Mackenzie, R., Palmer, C.R., Lomas, D.J., Dixon, A.K. (2006). Magnetic resonance imaging of the knee: diagnostic performance studies. *Clinical Radiology,* 51(4), 251-257

Appendix 1. Search strategy

Pubmed and Medline Mesh terms

("Back Pain"[Mesh] OR Back Pain*[tiab] OR Back ache*[tiab] OR Backache*[tiab] OR "Rheumatic Disease"[Mesh] OR Rheumatic Disease*[tiab] OR Rheumat*[tiab] OR Enthesopath*[tiab] OR Osteoarthrit*[tiab] OR "Neck Pain"[Mesh] OR Neck Pain*[tiab] OR Neck ache*[tiab] OR Neckache*[tiab] OR Cervicalgia*[tiab] OR Cervical Pain*[tiab] OR "Shoulder pain"[Mesh] OR Shoulder pain*[tiab] OR "wrist injuries"[Mesh] OR wrist injur*[tiab] OR "hip injuries"[Mesh] OR Hip injury[tiab] OR "Patellofemoral Pain Syndrome"[Mesh] OR Patellofemoral Pain*[tiab] OR knee Pain*[tiab] OR "Knee injuries"[Mesh] OR Knee injur*[tiab] OR cruciate ligament tear*[tiab] OR Meniscus[tiab] OR "Foot injuries"[Mesh] OR Foot injur*[tiab] OR Achillis tendon[tiab] OR plantar fasciitis[tiab] OR musculoskeletal complaint*[tiab] OR musculoskeletal pain*[tiab] OR skeletal complaint*[tiab] OR skeletal pain*[tiab] OR muscular complaint*[tiab] OR muscular pain*[tiab] OR muscle complaint*[tiab] OR muscle pain*[tiab] OR muscles complaint*[tiab] OR muscles pain*[tiab])

AND

("Diagnostic Imaging"[mh] OR Diagnostic test*[tiab] OR mri[tiab] OR radiograph*[tiab] OR imaging*[tiab] OR Tomogra*[tiab] OR CT OR Ultrasonogra*[tiab] OR sonogra*[tiab])

AND

(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR randomly[tiab] OR trial[tiab] OR groups[tiab]) NOT (animals[mh] NOT humans[mh])

Embase Emtree terms

('musculoskeletal chest pain'/de OR 'musculoskeletal injury'/exp OR 'musculoskeletal pain'/de OR 'musculoskeletal disease'/exp OR 'musculoskeletal stiffness'/exp OR myalgia/exp OR 'wrist pain'/de OR 'wrist injury'/de OR 'elbow injury'/de OR 'shoulder pain'/de OR 'shoulder injury'/de OR 'neck pain'/de OR 'neck injury'/exp OR 'cervical spine injury'/de OR 'spine injury'/de OR 'low back pain'/de OR 'hip injury'/de OR 'hip pain'/de OR 'knee injury'/exp OR 'knee pain'/de OR 'ankle injury'/exp OR 'ankle pain'/de OR 'patellofemoral pain syndrome'/de OR osteoarthritis/exp OR 'foot injury'/exp OR 'plantar fasciitis'/de OR (((musculoskelet* OR skelet* OR muscular* OR muscle* OR wrist* OR elbow* OR shoulder* OR neck OR spine OR 'low back' OR 'lower back' OR lowback OR lumbar OR lumbal OR lumbosacral OR hip OR hips OR knee* OR ankle* OR loin OR cervical OR patellofemor* OR lumbosacroiliac OR ligament* OR foot OR feet) NEAR/3 (pain* OR ache* OR complaint* OR injur* OR syndrome* OR disorder* OR symptom* OR strain* OR rupture* OR menisc* OR achilles OR tendon* OR lesion* OR tear* OR failure*)) OR myalgia* OR neckache* OR cervicalgia* OR lumbago OR lumbagalg* OR lumbodynia* OR osteoarthr* OR arthritis OR arthrosis OR 'degenerative joint disease' OR 'osteo arthritis' OR 'plantar fasciitis'):ab,ti)

AND

('imaging and display'/de OR 'diagnostic imaging'/de OR thermography/de OR spectroscopy/exp OR scintigraphy/exp OR radiography/exp OR 'computer assisted tomography'/exp OR 'nuclear magnetic resonance imaging'/exp OR ultrasound/de OR echography/exp OR myelography/de OR thermography/de OR (imaging OR radioimaging OR thermogra* OR spectroscop* OR scintigra* OR laminoscintigra* OR scintillation* OR scintillogra* OR scintiphotogra* OR radiogra* OR electroradiogra* OR pneumoradiogra* OR radiophotogra* OR roentgen* OR rontgen* OR x-ray OR xray OR tomogram* OR ((cat OR ct) NEXT/1 scan*) OR mri OR nmri OR NMR OR ultraso* OR 'ultra sound' OR echogram* OR echoscop* OR echosound OR sonogram* OR ultrasonogram* OR myelogra* OR medullogra* OR thermogra* OR thermoscan* OR infrared OR 'ophthalmo diaphanoscopy' OR transillumination):ab,ti)

AND

((random* OR factorial* OR crossover* OR (cross NEXT/1 over*) OR placebo* OR ((doubl* OR singl*) NEXT/1 blind*) OR assign* OR allocat* OR volunteer*):ab,ti OR 'crossover procedure'/de OR 'double-blind procedure'/de OR 'randomized controlled trial'/de OR 'single-blind procedure'/de) NOT ([animals]/lim NOT [humans]/lim)

AND

('disease course'/exp OR 'therapy effect'/de OR 'pain assessment'/de OR reassurance/de OR 'daily life activity'/de OR 'ADL disability'/exp OR 'patient satisfaction'/de OR 'psychological aspect'/de OR anxiety/de OR 'cost effectiveness analysis'/de OR (convalescen* OR recover* OR deteriorate* OR 'disease course' OR prognis* OR relapse* OR (therap* NEAR/3 effect*) OR (pain NEAR/3 (assess* OR measure* OR score*)) OR reassur* OR (daily NEAR/3 (activit* OR function*)) OR ADL OR satisf* OR psycholog* OR anxiety*):ab,ti)

Appendix 2. PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
Abstract			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4, 16-17
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4-5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4-5

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analysis	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	6
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	App 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	7-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	7-11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	7-11
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	7-11
Discussion			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	7-11
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	11-13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14

NA Not applicable, Fig Figure, App Appendix, p page

Appendix 3. Risk of bias assessment.

Study	Randomization	Concealment	Comparable Baseline characteristics	Selection criteria	Outcome assessor blinded	Therapist/care provider blinded	Patient blinded	Numerical information sufficient	Intention to treat analysis	Total
Ash et al and Modic et al	1	1	1	1	?	0	0	1	1	6
Brealy et al	1	1	1	1	0	0	0	1	0	5
Bryan (a) et al	1	1	1	1	?	?	?	0	1	5
Bryan (b) et al	1	1	1	1	?	?	?	0	1	5
Cohen et al	1	1	1	1	1	0	1	1	0	7
Deyo et al	1	?	1	1	0	?	0	0	0	3
Djais and Kalim	1	?	1	1	0	0	0	1	0	4
Gilbert et al	1	?	1	1	?	?	?	1	0	4
Kendrick et al and Miller et al	1	1	1	1	?	?	?	1	0	5
Kerry et al	1	1	0	1	?	0	?	1	0	4
Patel et al	1	1	?	1	?	?	?	1	0	4
	100%	73%	82%	100%	9%	0%	9%	73%	27%	

1: yes
0: no
?: don't known

Chapter 5

Does the outcome of diagnostic ultrasound influence the treatment modalities and recovery in patients with shoulder pain in physiotherapy practice?

not yet published

Yasmaine Karel
Audilia Miranda
Marloes Thoomes-de Graaf
Wendy Scholten- Peeters
Ramon Ottenheijm
Bart Koes
Arianne Verhagen

ABSTRACT

Study Design

Prospective cohort study including 389 patients with shoulder pain in primary care physiotherapy.

Background

There is an increased tendency to use diagnostic ultrasound to aid the diagnostic strategy and target treatment. It is a relatively cheap and accessible imaging technique but the implications for practice and patients are unknown.

Objectives

To study the influence of diagnostic ultrasound (US) on diagnostic work-up, treatment modalities and recovery in patients with shoulder pain in physical therapy practice.

Method

Participants with a new episode of shoulder pain were assessed at baseline and followed for 26 weeks. Diagnostic work-up, including the use of diagnostic US, and treatment strategies were reported by the therapists at 3, 6 and 12 weeks. Patients reported on recovery at 6, 12 and 26 weeks follow-up.

Results

Most patients were diagnosed with subacromial impingement/pain syndrome after physical examination or diagnostic US. Diagnostic US was used in 31% of the participants. Tendinopathy was the most found abnormality in this sub-population. The patients who underwent diagnostic US were more frequently treated using exercise therapy. Patients that did not have a diagnostic US were more likely to receive massage therapy, trigger point therapy or mobilisation techniques. In the non-US-group (64%) more patients reported being recovered than in the US group (53%). Logistic regression analyses did not show a significant association between diagnostic US and recovery after 26 weeks (0.88, 95%CI:0.50-1.57).

Conclusion

Diagnostic US as a work-up component does not seem to influence diagnosis or recovery but does influence the choice of treatment modality. High quality randomized trials should study the effect of diagnostic US on recovery.

INTRODUCTION

Shoulder complaints are the third most common musculoskeletal complaint in the Netherlands. [18] Studies have shown an unfavourable recovery for 40-70% of patients with shoulder pain after 6 months and high indirect costs attributed to sick leave. [6, 15, 19, 29] In Dutch general practice about 50% of patients receive medication, 32% a wait-and-see policy and 16% are referred to a physical therapist. [8]

Initial management of patients with shoulder complaints is usually conservative except for younger patients with an acute traumatic rotator cuff tear. [1] When primary care treatment fails to improve the patient's symptoms, a referral to secondary care can be made.

According to the Dutch guidelines, physical therapists (PTs) and general practitioners (GPs) are recommended to classify patients into one of three groups: 1) with reduced passive range of motion (complaints due to glenohumeral deficit), 2) without reduced passive range of motion but with a painful abduction range (subacromial deficit), 3) without reduced passive range of motion and without a painful abduction range (shoulder instability). [9, 13] This classification can give the clinician an indication of the nature of the complaint. Research has shown that based on history taking and physical examination a more detailed classification of diagnostic labels is not reliable and not likely to change the initial therapeutic approach chosen by the GP. [4, 11, 12]

In primary care there is an increased tendency to use diagnostic ultrasound (US) to aid the diagnostic strategy and target treatment. It is a relatively cheap and accessible imaging technique. Some clinicians believe that determining an accurate diagnosis is essential to be able to provide the appropriate treatment. However, there is a lack of correlation between rotator cuff tears and symptoms experienced by the patient. [21] Therefore our aim was to study the influence of diagnostic US on clinical reasoning, treatment modalities and recovery in physical therapy practice.

METHOD

Study design

This study was part of a prospective cohort study with a follow-up of 26 weeks in PT practice including patients with non-specific shoulder complaints: named "X". Details of the study design are published elsewhere. [16] The Medical Ethics Committee of the Erasmus Medical Center approved the study protocol (MEC-2011-414).

Study population

Physical therapists (n=125) from the South West region on the Netherlands participated in the study and recruited patients from November 2011 till November 2012. Patients were either referred by their GP or consulted the PT through direct access.

Patients with shoulder pain were eligible when they were 18 years or over and adequately understood the Dutch language. Exclusion criteria were: patients with serious pathologies (infection, cancer or fracture), shoulder surgery in the past 12 months or diagnostic imaging techniques (musculoskeletal ultrasound, magnetic resonance imaging or radiography) performed on the shoulder in the past 3 months. The PTs using diagnostic US in usual care had to have at least one year of experience with diagnostic US and at least made 100 US scans of the shoulder.

Data collection

Data from PTs were collected at baseline, 3, 6 and 12 weeks after inclusion using digital questionnaires. Patients received a questionnaire at baseline, 6, 12 and 26 weeks after inclusion. Informed consent was received of patients and rights were protected. Clinical characteristics of the PTs (age, sex, work experience and/or specialization) and of the patients (age, gender, pain, duration of complaints and recurrence) were reported at baseline. The Shoulder Pain and Disability Index (SPADI) was used to measure level of disability. The Numeric Rating Scale (NRS-11) was used to score pain intensity. The scale ranges from 0 to 10, with 0 representing "no pain" and 10 "severe disabling pain". [6]

PTs reported the planned management at baseline in terms of initial clinical diagnosis (diagnostic label), the use of US (yes/no), pathological findings on diagnostic US, changes in clinical diagnosis after diagnostic US and initial therapeutic management of the patient. The diagnostic US could either be performed before or after physical examination. Whenever a treatment plan changed during follow-up, the PTs reported the reasons for change and the new treatment goal(s). Possible interventions were categorized into: information/advice, exercise therapy, massage, manual joint mobilization/ manipulation, extracorporeal shockwave therapy (ESWT), transcutaneous electrical nerve stimulation (TENS), trigger point therapy, taping/bracing or posture correction. Exercise therapy was subdivided in a) exercise of (muscle) function (strength/length), b) exercise of activities, c) stabilisation techniques for the rotator cuff/ scapulo-thoracic sliding mechanism.

Outcome measures

Diagnostic US. The following pathological findings were listed: tendinopathy, calcification, full thickness/partial thickness tears, biceps tendon rupture, bursitis, subacromial impingement syndrome, glenohumeral discontinuity, acromion discontinuity, osteoarthritis, labrum tear/SLAP, capsular thickening, and rotator cuff atrophy. One patient could

have more than one US finding but the first diagnosis was considered the most relevant to the complaints.

Recovery. Recovery status of the patient was measured with the Global Perceived Effect scale (GPE). The GPE uses a 7-point Likert scale indicating whether the patient's condition had improved or deteriorated since the start of their treatment. The outcome was dichotomised into "recovered" and "not recovered", with "recovered" defined as "completely recovered" or "much improved". The GPE is validated for patients with musculoskeletal complaints. [14]

Statistical analysis

Descriptive statistics of both baseline characteristics and outcome measures were presented in mean scores for continuous data with a normal distribution. Otherwise, median scores and the interquartile range (IQR) were used. Pearson's chi-square test was used to compare categorical data between groups. The Fisher exact test was used for small samples (n<10). If distribution was non-parametric, medians were compared using the Independent Sample Median Test. Distributions was compared using the Mann-Whitney U test. For the parametric distributions means were compared using the two-sample t-test. A p-value ≤ 0.05 was considered statistically significant. Binary logistic regression analysis was used to estimate the effect of diagnostic US on recovery, controlled for confounders. The variables age, duration of complaints, level of disability and pain were considered as possible confounders from previous literature. Crude and adjusted ORs with 95% confidence intervals (CI) were obtained. Complete case analysis was used on all the analyses. The number of missings is reported with all data. SPSS 22.0 was used for all analyses.

RESULTS

Study population

A total of 389 patients with a mean age of 50 years were included. In total 267 patients received a treatment solely based on history taking and physical examination (non-US-group), and 122 patients underwent a diagnostic US at baseline performed by a PT and were treated based on a post-ultrasound diagnosis (US-group).

Baseline

There was no significant difference in the gender distribution between the US and non-US-group (Table 1). The age of patients ranged from 19 to 83 years, with the majority between 45 and 54 years old. The mean difference of 4.7 years (95% CI 1.8-7.6) between the patients in the US and non-US-group was small but statistically significant.

Table 1. Baseline characteristics

Baseline characteristics	Total N=389	Non-US group N=267	US-group N=122
Female n (%)	206 (55)	145 (56)	61 (53)
Age mean (sd)	49.9 (13.2)	48.5 (12.8)	53.2 (13.6)*
Age groups n (%)			
≤34	50 (13)	37 (14)	13 (11)
35-44	78 (21)	62 (24)	16 (14)
45-54	108 (29)	78 (30)	30 (26)
55-64	79 (21)	53 (21)	26 (23)
≥65	59 (16)	29 (11)	30 (26)*
Duration of complaints in weeks, median (IQR)	12 (6-26)	12 (6-26)	12 (7-28)
Disability SPADI, mean (SD)	47 (21)	45 (22)	52 (20)*
Pain NRS-11, median, (IQR)	6 (4-7)	6 (4-7)	7 (5-7)*
Recurrent complaint yes, n (%)	158 (43)	106 (42)	52 (46)
Cause yes (%)			
Unexpected movement	23 (6)	13 (5)	10 (9)
Overuse	132 (36)	100 (39)	32 (28)
Trauma	24 (7)	16 (6)	8 (7)
Sports injury	32 (9)	21 (8)	11 (10)
Unclear	128 (35)	85 (33)	43(38)
Other	29 (8)	20 (8)	9 (8)

N Number, SD Standard Deviation, IQR Inter Quartile Range, SPADI Shoulder Pain And Disability Index, NRS Numeric Rating Scale
*p-value <0.05

When divided into age groups, there were significantly more patients in the age group between 35 and 44 years and in the age group of 65 years and older in the US-group (Table 1).

The median duration of complaints at inclusion in both groups was 12 weeks. The mean difference in disability score (SPADI) was 6.68 (95%CI 1.98- 11.37). The pain intensity score (NRS-11) at time of inclusion was significantly higher for the US-group (Table 1).

In the non-US-group 39% of the patients stated that their complaints were caused by overuse. This was significantly more compared to the US-group in which overuse accounted for 28% of the cases. There was no difference between the two groups for other probable causes of shoulder pain. An overview of the PT characteristics is presented in table 2.

Table 2. Characteristics of physiotherapists

	Total (n=102)	PTs without diagnostic US machine (n=64)	PTs with diagnostic US machine (n=38)
Sex men, N (%)	91 (77)	43 (68)	35 (92)*
Age in years, mean (SD)	44 (11)	44 (12)	45 (9)
Experience in years, N (%)			
<5	20 (20)	14 (22)	6 (16)
5-10	21 (21)	14 (22)	7 (18)
>11	61 (60)	36 (56)	25 (66)
Specialization, N (%)			
Manual therapist	51 (50)	34 (53)	17 (45)
Sports	21 (21)	11 (17)	10 (26)
Geriatrics	2 (2)	2 (3)	0
Pediatrics	2 (2)	2 (3)	0
Psychosocial	1 (1)	1 (2)	0
Vocational	5 (5)	3 (5)	2 (5)
Lymphatic	6 (6)	3 (5)	3 (8)
Worktime, N (%)			
Parttime	28 (22)	21 (36)	7 (19)
fulltime	67 (54)	38 (64)	29 (81)

N number, SD standard deviation
*p-value <0.05

Diagnostic US findings

Of the 122 patients who underwent diagnostic US 99 had complete data. The number of abnormalities ranged from 0 to 5 per patient. The majority of patients (n=42) had 2 abnormalities, 1 patient had 5.

Tendinopathy was the most found abnormality (30.8%), followed by calcification (19.5%), partial-thickness tendon tears (10.2%) and subacromial impingement (8.8%). The supraspinatus tendon was the most affected tendon. In 7 patients (3.1%) no pathology was found (Figure 1).

Reasons for using diagnostic US

In 34% (42/122) of all patients receiving a diagnostic US the reason was that the PT expected this would lead to a more specific clinical diagnosis, and in 13% (16/122) that it would help the PT in selecting the most appropriate intervention. In 12% of the patients (15/122) the PTs used the US findings to confirm their initial diagnosis and in another 11% (14/122) to better inform the patient about their complaints. Other reasons were 1) that it was a routine procedure in the physical examination, 2) that it would serve

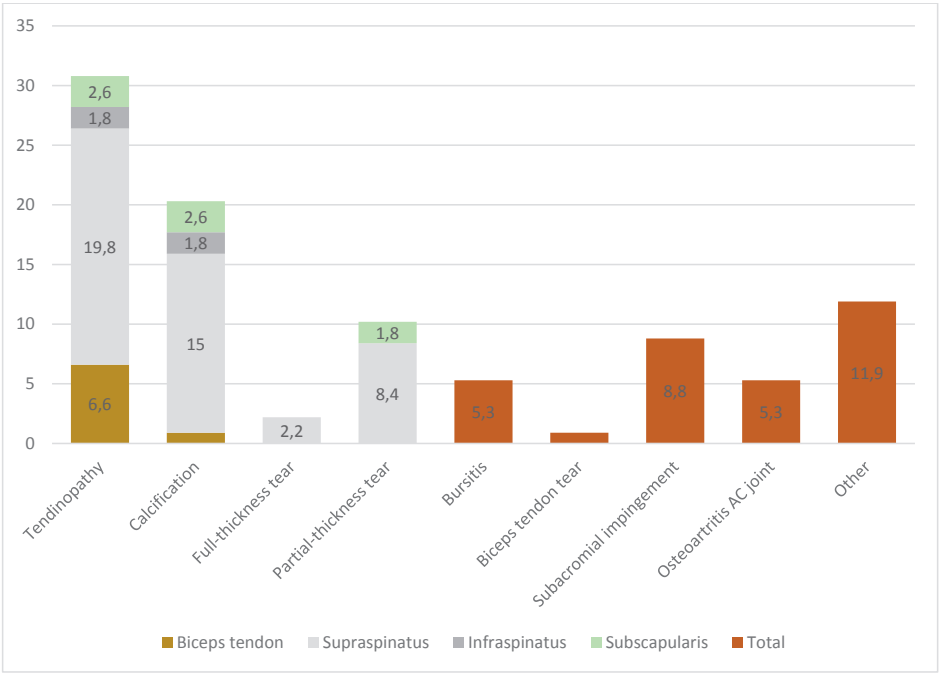


Figure 1. Percentage of US findings per pathology for each anatomical structure in the shoulder (colours) (n=116).

as a baseline measurement, 3) it was a request by a colleague and 4) that it would improve their professional position towards other health professionals. These were not selected frequently. Results suggest that US was most frequently performed when there was a suspicion of subacromial pathology.

Clinical diagnoses

Subacromial impingement syndrome (SIS) was the most reported diagnosis overall (Table 3). In the non-US-group this was followed by a disorder of the cervicothoracic spine (CTS) and costae, frozen shoulder/capsulitis and instability of the glenohumeral joint. In the US-group this was followed by a non-specific diagnosis, sprain or strain and instability of the glenohumeral joint (Table 2). In the US-group 75 patients also had a pre-US diagnosis, based on history taking with or without physical examination. SIS was the most occurring pre-US diagnosis (57.3%), followed by sprain or strain (12%), another non-specific diagnosis (6.7%) and acromioclavicular (AC) or sternoclavicular (SC) joint disorder. The clinical diagnosis changed in 29% (35/122) of patients after diagnostic US. In 31% (11/35) the clinical diagnosis changed from various diagnoses to a sprain (trauma) or strain.

Table 3. Clinical diagnosis for each group

Diagnostic groups (n, %)	*Total (n=340)*	*Non-US-group* (n=241)*	*US-group* (n=99)*
Subacromial impingement syndrome	139 (40.9)	79 (32.8)	60 (60.6)
Disorder of cervicothoracic spine (CTS) and costae	53 (15.6)	51 (21.2)	2 (2)
Frozen shoulder/capsulitis	29 (8.5)	27 (11.2)	2 (2)
Instability of the glenohumeral joint	27 (7.9)	22 (9.1)	5 (5.1)
Sprain or strain in neck/shoulder region	19 (5.6)	12 (5.0)	7 (7.1)
Internal (posterior) impingement syndrome	18 (5.3)	17 (6.4)	1 (1)
Acromioclavicular (AC) or sternoclavicular (SC) joint disorder	15 (4.4)	13 (5.4)	2 (2)
Biceps tendinopathy	10 (2.9)	6 (2.5)	4 (4)
Myofascial trigger point in neck/ shoulder	2 (0.6)	2 (0.8)	0
SLAP lesion (Superior Labral tear from Anterior to Posterior)	1 (0.3)	0	1 (1)
Muscular hypertonia in neck/shoulder	0	0	0
Other non-specific	21 (6.2)	11 (4.6)	12 (12.1)
Unclear/ Not possible to specify a clear diagnosis	4 (1.2)	1 (0.4)	3 (3)

* Non-US-group= diagnosis set after history and/or physical examination; US-group= diagnosis set after ultrasound. Missings non-US-group: 26; US-group: 23

Treatment

Patients were usually treated with a selection of different treatment modalities. In the non-US-group the maximum number of different modalities (including the different forms of exercise therapy) was 7. In the US-group there was a maximum number of 6 different modalities. In both groups the median of different modalities was 3 (p= 0.13).

The median number of treatment sessions in both groups was 7 and did not differ statistically significant between the US and non-US groups.

In the non-US-group 8.3% of the patients were referred (back) to their GP, 3.8% were referred to another healthcare professional (HP). In the US-group 13.2% were referred to their GP and 8.3% to another HP. The difference between the two groups for referral to GP or another HP was not statistically significant.

Informing, advising, counselling and coaching were the most used approaches regardless of the clinical diagnosis (Table 4). Table 4 shows the number of patients receiving a treatment modality per clinical diagnosis (left side of the table) and overall between the US and the non-US group (right side of the table).

Patients labelled with SIS received statistically significant more often stabilisation of the rotator cuff in the non-US group compared with the US group. There were significantly more patients treated with trigger point therapy through stretching and/or dry needling in the non-US-group (12.7 % vs 1.7%).

Table 4. Treatment modalities per clinical diagnosis and for the non-US-group and the US-group

Treatment modalities (n, %)	Total n=387	Sub-acromial impinge-ment n=139	Disorder CTS/ costae n=53	Frozen shoulder n=29
Informing, advising, counselling and coaching	332 (85.8)	128 (92.1)	49 (92.5)	29 (100)
Exercise therapy:	320 (82.7)	133 (95.7)	37 (69.8)	23 (79.3)
exercise of (muscle) function	230 (59.4)	88 (63.3)	27 (50.9)	22 (75.9)
exercise of activities	76 (19.6)	22 (15.8)	12 (22.6)	8 (27.6)
stabilisation rotator cuff/ scapula	212 (54.8)	*99 (71.2)	13 (24.5)	11 (37.9)
Massage	33 (8.5)	6 (4.3)	13 (24.5)	5 (17.2)
Manipulation and mobilisation techniques	215 (55.6)	73 (52.5)	49 (92.5)	19 (65.5)
Extracorporeal shock wave therapy	41 (10.6)	29 (20.9)	1 (1.9)	1 (3.4)
Passive modalities	5 (1.3)	0	0	2 (6.9)
Trigger point therapy (stretching/ dry needling)	32 (8.3)	*11 (7.9)	**5 (9.4)	3 (10.3)
Stabilisation shoulder (tape/bandaging)	54 (14)	26 (18.7)	0	0
No treatment	2 (0.5)	1 (0.7)	0	0
Other	37 (9.6)	*7 (5)	8 (15.1)	5 (17.2)

*p-value ≤0.05= statistically significant in favour of the non-US-group *within that specific diagnostic group.*
**p-value ≤0.05= statistically significant in favour of the US-group *within that specific diagnostic group.*
*** 2-sided p-value for comparison between non-US-group and US-group (the last two columns)
ESWT= extracorporeal shock wave therapy, CTS= cervicothoracic spine, AC= Acromioclavicular, SC= sternoclavicular.
Treatment for SLAP lesion and Unclear diagnosis are not shown in this table, due to small sample sizes.
No patients were diagnosed with muscular hypertonia

For patients labelled with a disorder of the cervicothoracic spine statistically significant more patients were treated with triggerpoint therapy in the non-US-group compared to the US-group.

A statistical significantly higher number of patients in the US-group received advice, counselling and coaching or extracorporeal shockwave therapy. Patients that did not have a diagnostic US were more likely to receive massage therapy, trigger point therapy or manipulation and mobilisation techniques.

Also, more patients in the US group were treated with ESWT. Pts could use ESWT in case of calcifications but only 33.6% of patients with calcifications were treated with ESW.

Recovery

The proportion of missing data on recovery was high for both the non-US-group and the US-group, ranging from 23% to 33% (Figure 2). At 6 weeks there were statistical significantly more patients in the non-US-group (46.2%) that reported being recovered

Instability n=27	Sprain/ Strain n=19	Internal impinge-ment n=18	AC/SC Joint n=15	Biceps tendino-pathy n=10	Other non-specific n=21	non-US group n=265	US-group n=122
23 (85.2)	15 (78.9)	17 (94.4)	11 (73.3)	10 (100)	20 (95.2)	222 (83.8)	110 (90.2)
26 (96.3)	18 (94.7)	17 (94.4)	12 (80)	9 (90)	17 (81)	209 (78.9)	111 (91)***
20 (74.1)	17 (89.5)	7 (38.9)	8 (53.3)	8 (80)	16 (76.2)	155 (58.5)	75 (61.5)
8 (29.6)	5 (26.3)	4 (22.2)	3 (20)	2 (20)	5 (23.8)	54 (20.4)	22 (18)
22 (81.5)	14 (73.7)	14 (77.8)	6 (40)	4 (40)	7 (33.3)	142 (53.6)	70 (57.4)
1 (3.7)	1 (5.3)	1 (5.6)	3 (20)	0	1 (4.8)	28 (10.6)	5 (4.1)***
7 (25.9)	6 (31.6)	13 (72.2)	13 (86.7)	*5 (50)	11 (52.4)	159 (60)	56 (45.9)***
0	0	5 (27.8)	1 (6.7)	0	0	23 (8.7)	18 (14.8)***
0	1 (5.3)	0	1 (6.7)	0	0	3 (1.1)	2 (1.6)
2 (7.4)	2 (10.5)	3 (16.7)	2 (13.3)	0	2 (9.5)	27 (10.2)	5 (4.1)***
11 (40.7)	2 (10.5)	4 (22.2)	**3 (20)	0	2 (9.5)	32 (12.1)	22 (18)
1 (3.7)	0	0	0	0	0	2(0.8)	0
0	1(5.3)	2 (11.1)	3 (20)	3 (30)	3 (14.3)	31 (11.7)	6 (4.9)***

compared to the US-group (30.2%). The difference in recovery was not statistically significant at 12 and 26 weeks but still the proportion of patients reporting recovery was higher for the non-US-group.

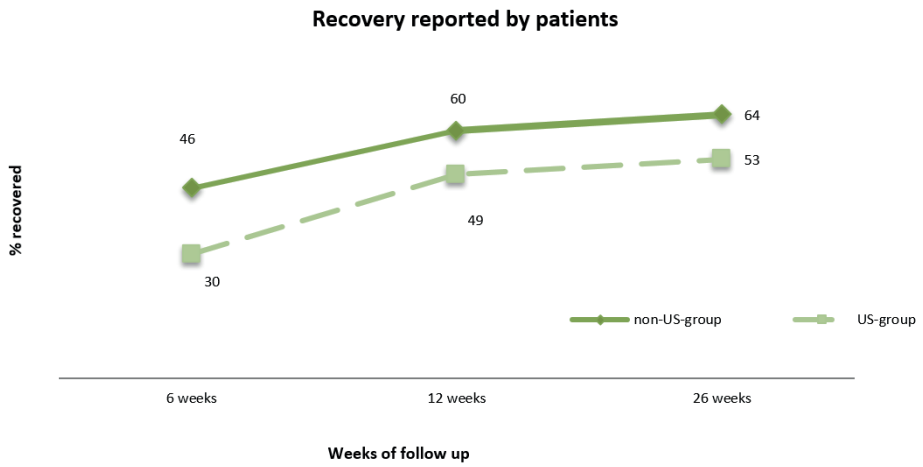


Figure 2. Patients that reporterd 'strongly improved' or 'completely recovered' on the Global Perceived Effect (GPE) scale.
Missings non-US-group and US-group at week 6: 26% and 30%, week 12: 33% and 25%, week 26: 33% and 23% resp.

Association between diagnostic US and recovery

Binary logistic regression analysis found a statistically significant crude OR of 0.53 (CI 0.30-0.92), meaning a negative association between diagnostic US and recovery at 6 weeks. The estimate changed after adjusting for confounders to 0.64 (0.36-1.14) and was not statistically significant anymore. Both adjusted ORs after 12 weeks (0.73, CI:0.42-1.28) and 26 weeks (0.88, CI:0.50-1.57) were also not statistically significant.

DISCUSSION

The most common clinical diagnosis was SIS and for the US-group the clinical diagnosis did not change after the diagnostic US. The referral rate was slightly higher in the US-group but not statistically significant. The use of diagnostic US did seem to have some influence on the applied treatment modalities by the PTs. There were slightly more patients treated with exercise therapy in the US-group, but when subdivided in different subgroups of exercise therapy, no statistically significant differences were found. In the non-US-group statistically significant more patients were treated with manipulation and mobilisation techniques, massage and triggerpoint therapy. No major differences were found in other diagnostic groups. In the non-US-group more patients reported being recovered. The difference was only significant at 6 weeks follow-up. The use of diagnostic US seemed to have a negative effect on recovery at 6 weeks but this effect might be confounded by indication: i.e. patients with a worse prognosis based on for example age, duration of symptoms, level of disability and pain and/or variable which we did not measure have a higher chance to receive a diagnostic US.

Comparison with the literature

Baseline characteristics were similar to other studies done in primary care. This study had slightly more (56%) female patients, which was in line with other literature. [5, 25, 28] Most patients were between 45 and 64 years of age; this age group consults their PT most often for all kind of musculoskeletal complaints in the Netherlands. [3] Similar to the results in our study, SIS, in particular rotator cuff tendinopathy, is the most frequently diagnosed disorder. [28]

In our study the PT that made a diagnostic US found a tendinopathy in the majority of patients, and only 5.2% of the patients had a full-thickness tear. A retrospective observational study under 240 patients who were referred by GPs to make a diagnostic US, concluded that in most cases there was a calcific tendonitis (29%), a tendinopathy in 11% of cases and a full-thickness tear in 8%. [24] A prospective study where patients with acute shoulder pain were referred to a radiologist for a diagnostic US also found calcific tendonitis to be the most frequent pathology (50.4%) followed by tendinopathy

(28.7%) and full thickness tears (3.1%). [22] A systematic review with secondary care studies, showed that tendinopathy (30-39%) and full thickness tears (24-70%) were the most observed disorders. [23] The differences of pathologies on diagnostic US between studies can be attributed to the different criteria used for obtaining a diagnostic US or selection criteria of patients. PTs with sufficient experience were selected but no explicit criteria were set for performing a diagnostic US; it was left to the discretion of the PT. This might influence the validity of the pathological findings. The majority of PTs in this study used US to identify a more specific clinical diagnosis.

In our study no pathology on diagnostic US was found in 6% of the patient, which is in contrast to the 40% described in previous literature where US was performed by radiologists in a primary care population. [24] This might indicate that PTs already use diagnostic US in a patient group where they suspect to find pathology. In line with other literature the supraspinatus tendon was the most frequently affected tendon. [17, 24]

Research shows that after the 5th decade an increase in asymptomatic rotator cuff tendon tears are found, linearly increasing every decade. [20] In the US-group in our study 26% of the patient were 65 years or older confirming the earlier results. In this group less patients had US diagnosed tendinopathy and calcifications. These US findings may have been due to degeneration and may not have been the cause of the symptoms described by the patients. Furthermore, more than one abnormality was frequently found in patients, but they may not have had any clinical implications. Research performed in 51 men without complaints of the shoulder, showed that in 96% asymptomatic abnormalities were found. [10] Subacromial bursal thickening was found in 78%, osteoarthritis of the AC joint in 65% and supraspinatus tendinosis in 39%. [10]

A cross-sectional study has shown that MRI and diagnostic US have equally high accuracy for identifying biceps pathology and rotator cuff tears, while physical examination has modest accuracy. [2] In addition, US could not detect glenoid labral tears and bone erosion. [2] This confirms that the choice for the use of additional imaging should be based on clinical information and might not be helpful as a standard method of assessment. Otherwise no assessment of the relevance of the abnormalities found trough US can be made.

Strengths and limitations

Our study was set in a primary care patient population. Little is known about US findings in primary care populations. Most studies on diagnostic US are performed in secondary care where US is usually used for the work-up to a surgical intervention. [23, 24]

Our study was first to evaluate diagnostic US performed by PTs. In most literature on the accuracy of diagnostic US the scan was performed by a radiologist. PTs in contrast to radiologists tend to find more tendinopathy and partial-thickness tears. [27] Furthermore, the reliability between PTs and radiologists in this study is borderline substantial for

full-thickness tears only. [27] These results suggests that the diagnosis after diagnostic US performed may have questionable validity.

The PTs who participated in this study had knowledge of diagnostic US and showed interest in determining its value in the diagnostic process. The PTs decided which patients were to have an US, therefore there might be a selection in the patients that received an US. Baseline factors between these groups did not differ.

Furthermore, there is no uniformity in the definition of the various diagnostic labels used in different studies and the labels only have fair to moderate inter-observer reproducibility. This challenges the ability to compare various study results. [26]

Implications

The need for a specific diagnosis is mainly driven by the desire to influence the outcome of a patient by a specific treatment modality and thereby establish a more efficient and cost-effective treatment plan. Where patients with calcification should be treated with rest and analgesics due to its self-limiting nature, ESWT can be considered or a referral in younger patients with an acute traumatic rotator cuff tear. [1, 13] In our study only 33.6% of patients with calcifications were treated with ESWT. Of all patients with full thickness tears only one was referred to the GP and none were referred to other health care professionals. As full thickness tears may not heal and may require surgery, especially in the younger athlete, the orthopaedic surgeon will have to consider which management would be appropriate. This advice is also recommended in the evidence statement for PTs. [13] The evidence statement recommends exercise therapy, which most PTs used in their treatment regime. Whether the diagnostic US provided more information to choose exercise therapy more often remains unknown. Trigger point therapy was still used in a small number of patients while the evidence statement discourages this.

CONCLUSION

Diagnostic US as a work-up component does not seem to influence diagnostic work-up, and recovery but the choice of treatment differed between the groups. The patients who underwent diagnostic US were more frequently treated using exercise therapy. Patients that did not have a diagnostic US were more likely to receive massage therapy, trigger point therapy or manipulation and mobilisation techniques. High quality randomized trials should study the effect of diagnostic US on recovery.

REFERENCES

1. Arce, G., Bak, K., Bain, G., Calvo, E., Ejnisman, B., Di Giacomo, G., Gutierrez, V., Guttmann, D., Itoi, E., Kibler, B.W., Ludvigsen, T., Mazzocca, A., de Castro Pochini, A., Savoie, F., Sugaya, H., Uribe, J., Vergara, F., Willems, J., Yoo, Y.S., McNeil, J.W., Provencher, M.T. (2013). Management of disorders of the rotator cuff: proceedings of the ISAKOS upper extremity committee consensus meeting. *Arthroscopy*, 29, 1840-50
2. Ardic, F., Kahraman, Y., Kacar, M., Kahraman, M.C., Findikoglu, G., Yorgancioglu, Z.R. (2006). Shoulder Impingement Syndrome: relationship between clinical, functional and radiological fundings. *American Journal of Physical Medicine & Rehabilitation*, 1(85), 53-60.
3. Barten, D.J.A., Verberne, L.D.M., Koppes, L.L.J. *Care provided by the physiotherapist - gender and age. NIVEL Primary Care Registration*. [Online] 2013. www.nivel.nl/node/3236.
4. Beaudreuil, J., Nizard, R., Thomas, T., Peyre, M., Liotard, J.P., Boileau, P., Marc, T., Dromard, C., Steyer, E., Bardin, T., Orcel, P., Walch, G. (2009). Contribution of clinical tests to the diagnosis of rotator cuff disease: A systematic literature review. *Joint Bone Spine*, 76(1), 15-19.
5. Bot, S.D., van der Waal, J.M., Terwee, C.B., van der Windt, D.A., Schellevis, F.G., Bouter, L.M., Dekker, J. (2005). Incidence and prevalence of complaints of the neck and upper extremity in general practice. *Annals of Rheumatic Diseases*, 64, 118-23.
6. Bot, S.D., van der Waal, J.M., Terwee, C.B., van der Windt, D.A., Scholten, R.J., Bouter, L.M., Dekker, J. (2005). Predictors of outcome in neck and shoulder symptoms: a cohort study in general practice. *Spine*, 30(16), E459-70.
7. Breivik, E., Björnsson, G., Skovlund, E. (2000). A Comparison of Pain Rating Scales by Sampling From Clinical Trial Data. *The Clinical Journal of Pain*, 16(1), 22-28.
8. Dorrestijn, O., Greving, K., van der Veen, W.J., van der Meer, K., Diercks, R.L., Winters, J.C., Stevens, M. (2010). Patients with shoulder complaints in general practice: consumption of medical care. *Rheumatology*, 2(50), 389-395
9. Geraets, J.J., de Jongh, A.C., Boeke, A.J., Buis, P.A., Spinnewijn, W.E., Geijer, R.M., Goudswaard, A.N. (2008). Dutch College of General Practitioners: practice guideline for shoulder complaints. *Huisarts in Wetenschap*, 11(51), 565-566.
10. Girish, G., Lobo, L.G., Jacobson, J.A., Morag, Y., Miller, B., Jamadar, D.A. (2011). Ultrasound of the shoulder: asymptomatic findings in men. AJR Am J *Roentgenol*, 197(2), W713-9.
11. Hegedus, E.J., Goode, A., Campbell, S., Morin, A., Tamaddoni, M., Moorman, C.T., Cook, C. (2007). Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine*, 42(2), 80-92.
12. Hughes, P.C., Taylor, N.F., Green, R.A. (2008). Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. *Australian Journal of Physiotherapy*, 3(53), 159-170.
13. Jansen, M.F., Brooijmans, F., Geraets, J.J.X.R., Lenssen, A.F., Ottenheijm, R.P.G., Penning, L.I.F., de Bie, R.A. (2001). KNGF Evidence Statement Subacromiale klachten. *Supplement bij het Nederlands Tijdschrift voor Fysiotherapie*, 121(1).
14. Kamper, S.J., Ostelo, R.W., Knol, D.L., Maher, C.G., de Vet, H.C., Hancock, M.J. (2010). Global perceived effect scales provided reliable assessments of health transition on people with musculoskeletal disorders, but ratings are strongly influences

by current status. *Journal of Clinical Epidemiology,* 63(7), 760-6.

15. Karel, Y.H., Verhagen, A.P., Thoomes-de Graaf, M., Duijn, E., van den Borne, M.P., Beumer, A., Ottenheijm, R.P., Dinant, G.J., Koes, B.W., Scholten-Peeters, G.G. (2016). Development of a Prognostic Model for Patients With Shoulder Complaints in Physical Therapist Practice. *Physical Therapy*.
16. Karel, Y.H., Scholten-Peeters, W.G., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R.P., van den Borne, M.P., Koes, B.W., Verhagen, A.P., Dinant, G.J., Tetteroo, E., Beumer, A., van Broekhoven, J.B., Heijmans, M. (2013). Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. *BMC Musculoskelet Disord,* 14(62).
17. Kim, H., Kim, S., Seo, Y. (2007). Ultrasonographic findings of painful shoulders and correlation between physical examination and ultrasonographic rotator cuff tear. *Modern Rheumatology*, 14(1), 213-219.
18. Kooijman, M.K., Verberne, L.D.M., Barten, J.A., Leemrijse, C.J., Veenhof, C., Swinkels, I.C.S. Jaarcijfers 2012 en trendcijfers 2008-2012 fysiotherapie. Landelijke Informatievoorziening Paramedische Zorg/ NIVEL Zorgregistraties eerste lijn. [Online] 2013. http://www.nivel.nl/NZR/zorgregistraties-eerstelijn.
19. Kuijpers, T., van Tulder, M.W., van der Heijden, G.J.M.G., Bouter, L.M., van der Windt, D.A.W.M. (2006). Costs of shoulder pain in primary care consulters: a prospective cohort study in The Netherlands. *BMC Musculoskelet Disord,* 83(7).
20. Milgrom, C., Schaffler, M., Gilbert, S., van Holsbeeck, M. (1995). Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender. *J Bone Joint Surg Br.*, *77*(2), 296-8.
21. Minagawa, H., Yamamoto, N., Abe, H., Fukuda, M., Seki, N., Kikuchi, K., Kijima, H., Itoi, E. (2014). Prevalence of symptomatic and asymptomatic rotator cuff tears in the general population: From mass-screening in one village. *Journal of Orthopaedics*, 10(1), 8-12.
22. Ottenheijm, R.P., Cals, J.W., Weijers, R., Vanderdood, K., de Bie, R.A., Dinant, G.J. (2015). Ultrasound Imaging for Tailored Treatment of Patients With Acute Shoulder Pain. *The Annals of Family Medicine,* 13(1), 53-55.
23. Ottenheijm, R.P., Jansen, M.J., Staal, J.B., van den Bruel, A., Weijers, R.E., de Bie, R.A., Dinant, G.J. (2010). Accuracy of Diagnostic Ultrasound in Patients With Suspected Subacromial Disorders: A Systematic Review and Meta-Analysis. *Archives of Phisical Medicine and Rehabilitation*, 91(10), 1616-1625.
24. Ottenheijm, R.P., van't Klooster, I.G., Starmans, L.M., Vanderdood, K., de Bie, R.A., Dinant, G.J., Cals, J.W. (2014). Ultrasound-diagnosed disorders in shoulder patients in daily general practice: a retrospective observational study. *BMC Family Practice*, 15(1), 115.
25. Picavet, H., Schouten, J. (2003). Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC3-study. *Pain*, 102(1), 167-178.
26. Schellingerhout, J.M., Verhagen, A.P., Thomas, S., Koes, B.W. (2008). Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach. *Manual Therapy*, 13(6), 478-83.
27. Thoomes-de Graaf, M., Scholten-Peeters, G.G., Duijn, E., Karel, Y.H., van den Borne, M.P., Beumer, A., Ottenheijm, R.P., Dinant, G.J., Tetteroo, E., Lucas, C., Koes, B.W., Verhagen, A.P. (2014). Inter-professional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Manual Therapy*, 19(5), 478-483.
28. van der Windt, D.A., Koes, B.W., de Jong, B.A., Bouter, L.M. (1995). Shoulder disor-

ders in general practice: incidence, patient characteristics, and management. *Annals of Rheumatic Diseases*, 54(12), 959-964.

29. Virta, L., Joranger, P., Brox, J.I., Eriksson, R. (2012). Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. *BMC Musculoskeletal disorders*, 13(17)

Chapter 6

Development of a Prognostic Model for Patients With Shoulder Complaints in Physiotherapy

Published in Physical Therapy

2016, 97(1)

Yasmaine Karel
Arianne Verhagen
Marloes Thoomes-de Graaf
Edwin Duijn
Maaike van den Borne
Annechien Beumer
Ramon Ottenheijm
Geert-Jan Dinant
Bart Koes
Wendy Scholten-Peeters

ABSTRACT

Background

Health care providers need prognostic factors to distinguish between patients who are likely to recover compared to the ones that do not.

Objective

To describe the clinical course and identify prognostic factors of recovery, in patients with shoulder pain at 26 weeks follow-up.

Design

A prospective cohort study was carried out in the Netherlands including 389 patients consulting a physiotherapist with a new episode of shoulder pain.

Method

Patients were followed for 26 weeks. Potential predictors were selected from the literature, together with the use of diagnostic ultrasound and working alliance and evaluated in multivariable regression analysis. Multiple imputation was used to handle missing data and bootstrap methods for internal validation.

Results

Recovery rate was 60% for the total population and 65% for the working population after 26 weeks. Short duration of complaints, lower disability scores, having a paid job, better working alliance and no feelings of depression/anxiety were associated with recovery. In the working population only duration of complaints and disability remained in the final model. The area under the receiver operator curve (AUC) was 0.67 for the final model of the total population and 0.63 for the working population. After internal validation the AUC was corrected to 0.66 and 0.63.

Limitations

External validation should be done prior to the use in clinical practice.

Conclusion

Results from this study indicate that several factors can predict recovery.

INTRODUCTION

Shoulder complaints are common in western societies and belong to the top 3 of most occurring musculoskeletal complaints. [1] Prevalence rates in the Netherlands range from 6.9 to 48% in primary care. [2-4] About 13% of the patients with shoulder pain who visit the general practitioner are referred to physiotherapy. [4] In the Netherlands patients can visit the physiotherapist without a referral since 2006 and 41% of patients in physiotherapy care used direct access in 2013. [5]

Examining patients with shoulder pain is complex because history taking and physical examination have limited validity for diagnosing the patho-anatomical origin of symptoms. Knowledge about prognostic factors can help the physiotherapist by informing the patient about the expected prognosis and, when indicated, in treatment decisions or referral to other health care professionals. [6, 7] Duration of symptoms, high levels of pain and the presence of co-morbidities have been identified as predictors of poor recovery by patients consulting a General Practitioner (GP). [7-11] Because of the difficulty in diagnosing patients with shoulder pain, physiotherapists are increasing the use of diagnostic ultrasound to assist their clinical decision-making. Nevertheless, the diagnostic and prognostic consequences of using diagnostic ultrasound remains unknown. [12, 13] Furthermore, recent literature suggest patient's prognosis to be influenced by the therapeutic relationship, frequently referred to as "working alliance". [14]

Health care providers need prognostic factors to distinguish between patients who are likely to recover compared to the ones that do not, i.e. the patients which have a high risk of developing chronic shoulder pain. Prognostic factors for shoulder pain have been identified in general practice and only duration of complaints, disability score and age have been identified in a physiotherapy setting. [7, 15] Although patients visiting general practice might be similar in type and severity of complaints compared to the patients in physiotherapy practice, the moment of seeking health care and the treatment provided in both settings is different for most patients. In this study we aim to identify prognostic factors of recovery, including the use of diagnostic ultrasound and working alliance, for patients with shoulder pain in physiotherapy practice.

METHODS

Study Design

This study was a prospective cohort study with a follow-up of 26 weeks in physiotherapy practice of patients with non-specific shoulder complaints. Details of the study design were published in 2013. [16] The Medical Ethics Committee of the Erasmus Medical Center approved the study protocol (MEC-2011-414).

Study Population

From November 2011 to November 2012 physiotherapists recruited consecutive patients. Patients that consulted the physiotherapist were eligible for the study when they suffered from shoulder pain, were aged≥18 years and had adequate understanding of the Dutch language. Patients were excluded if they had serious pathologies (infection, cancer or fracture), previous surgery of the shoulder in the last 12 months, or received diagnostic imaging techniques such as musculoskeletal ultrasound, magnetic resonance imaging or X-ray of the shoulder in the 3 months prior to start of the study. All patients provided written informed consent.

Procedures

During first consultation patients received study information and signed the consent form. This was sent to the researchers together with patients' name and e-mail address. Next, baseline questionnaires were sent to the e-mail address or post address when patients did not have e-mail. Follow-up questionnaires were sent 6, 12 and 26 weeks after the start of the treatment. A maximum of 2 reminders were sent when no response was received after 3 and 5 days.

Candidate predictors

Prognostic factors for recovery for patients with shoulder pain were extracted from the literature and consisted of sociodemographic variables and clinical characteristics. [7, 10, 17-19] Sociodemographic variables were age (continuous), gender, level of education (low = no education, primary school or lower vocational school, medium = lower general secondary school or middle vocational school, high = higher general secondary school, higher vocational school or university), employment status (paid job yes/no) and job description (physically heavy work, static repetitive work or work with awkward postures; yes/no).

Clinical characteristics were duration of complaints (months), previous episode of shoulder pain (yes/no), pain intensity at baseline (11-point numeric rating scale, NRS-11), and co-morbidity of arm (elbow/wrist/hand), back or neck (yes/no), sick leave due to shoulder complaint (yes/no), and increase of complaints during work (yes/no).

The shoulder complaint was considered work related when patients with a paid job answered "yes" to one of the following three questions: (1) Do the complaints worsen or return during activities at work? (2) Have you adapted or reduced your activities at work because of your complaints? (3) Do the complaints diminish after several days off work? [20]

The Dutch Shoulder Pain and Disability Index (SPADI) consist of five items assessing pain and eight items assessing disability. The score ranges from 0 to 100% with a high score indicating more functional disability. The questionnaire has good validity and reliability. [21]

Additionally, we assessed working alliance, the use of diagnostic ultrasound (yes/no) and the anxiety/depression dimension of the EuroQOL five dimensions as possible prognostic factors. Working alliance was measured with the Flemish (Dutch) version of the Working Alliance Inventory (WAV-12) and was assessed after 6 weeks. This questionnaire has three subscales designed to assess three primary components of the working alliance: 1) how closely client and therapist agree on and are mutually engaged in the goals of treatment, 2) how closely client and therapist agree on how to reach the treatment goals and 3) the degree of mutual trust, acceptance, and confidence between client and therapist. Patients score on a 5-point scale ranging from rarely to always. This scale is validated in patients receiving psychotherapy in Belgium. [22, 23]

The EuroQOL 5 dimensions-3L (EQ-5D) was used to measure health related quality of life. Little is known about the prognostic value of psychosocial factors. Therefore we used one dimension focusing on the emotional and social functioning, questioning the patient whether he or she was anxious or depressed (not, moderate or extremely). The EQ-5D is a valid and reliable generic instrument for measuring health related quality of life. [24, 25]

Outcome measures

The primary outcome measure was the Global Perceived Effect (GPE) scale and measures whether the patient rates it's condition as improved or deteriorated since the start of the physiotherapy treatment. It uses a 7-point Likert scale scoring and ranges from 'worse than ever' to 'fully recovered'. Patients were to be considered recovered when they scored 'strongly improved' or 'completely recovered'. [24, 26]

The secondary outcome measure were: 1) pain severity and was measured with the 11 point Numeric Rating Scale (NRS) ranging from no pain (0) to intolerable pain (10) and 2) disability measured with the Shoulder Pain And Disability Index (SPADI) ranging from no disability (0) to complete disability (100).

Sample size

Based on the literature about 40% of the patients with shoulder pain will recover within 6 months. [9, 27, 38] We aimed to include 12 prognostic variables in our prognostic model. Based on the 1 in 10 rule of 10 events per variable, a total of 120 events are needed in the smallest outcome (recovered or not). [28] Adjusting for about 20% missing values, the total population should comprise a minimum of 360 subjects.

Statistical Analysis

First we performed a descriptive analysis by calculating frequencies for categorical variables and means with standard deviations (SD) for continuous variables at 6, 12 and 26 weeks. In case the data was not normally distributed median scores and the interquartile range were reported. Multiple imputation was used in case of missing data. Predictor

variables and the outcome were included in the multiple imputation and was done separately for primary and secondary outcome measures. [29-31] A total of 20 datasets were created and regressions analysis was done in all datasets. Pooled estimates were calculated according to Ruben's rule. [32] All assumptions (linearity between independent variables and log odds and multicollinearity (>0.80) for continuous variables) were checked before model building. Univariable and multivariable regression were reported for the total population and working population separately, because several work related variables (job demands and psychosocial factors at work like low decision authority and low control) are found to be related to recovery in the working population specifically. [20, 33] Unadjusted associations were checked between each candidate predictor and the outcome for significant contribution to the outcome (P>0.2). All candidate predictors derived from the literature were included in the multivariate regression analysis (full model). Multiple logistic regression analysis was used to determine which baseline variables were predictors of recovery at 26 weeks (using the GPE). Next, a backward selection procedure was used to determine which variables were kept in the model (final model). A variable was selected when the variable appeared statistically significant in 12 out of 20 imputed models. [34] A p-value of <0.05 was considered statistically significant. The reliability of the multivariable model was determined with the Hosmer-Lemeshow goodness-of-fit statistic. [35] Discriminative ability of the models was assessed using the area under the receiver-operating characteristic curve (AUC-ROC). An area under the curve (AUC), of 0.5 indicates poor discrimination above chance, 0.7 indicates fair discrimination, 0.8 indicates acceptable discrimination, whereas an AUC of 1.0 indicates perfect discrimination. [35] Optimal models were classified as those that yielded the highest AUC. Calibration of the model predictions was assessed by the amount of overlap between the predicted individual probabilities against the observed recovery. The same 12 predictors used for logistic regression modeling were used for linear regression modeling with pain as outcome to evaluate if the model would be similar for a secondary outcome measure. Only one secondary outcome (pain) was used as a secondary outcome measure in the regression model because the SPADI and NRS scores were highly correlated (α=0.87).

We performed internal validation for the primary outcome measure by bootstrapping in order to correct for overfitting. A total of 1000 new datasets were created by random drawing samples from the dataset and we assessed the AUC. [36] The performance in the bootstrap sample represents estimation of the apparent performance, and the performance in the original sample represents test performance. The difference between these is an estimate of the optimism in the apparent performance. The optimism is subtracted from the apparent performance to estimate the internally validated performance. [37] All imputed datasets were bootstrapped and the AUCs were averaged to get the apparent performance. Statistical analyses were performed by using SPSS 22.0 software. Bootstrap analyses were done with R software. [38]

RESULTS

Study population

In total 412 patients fulfilled the eligibility criteria of which 389 gave informed consent and thus entered the cohort. From the 389 patients 366 (94%) returned the baseline questionnaire. After 26 weeks 272 (70%) returned the questionnaire (figure 1). There were 11% missing values. There were no statistically significant differences in baseline characteristics in patients with or without missing data.

Baseline characteristics of the study population were described in table 1 together with missing data. The population consisted of 170 men (45%), the mean age was 49.9 (SD=13.2), 261 (71%) had a paid job and the median duration of their complaints was 12 weeks (IQR=6-26). The working population did not significantly differ from the total population except concerning disability (SPADI). All patients received physiotherapy treatment.

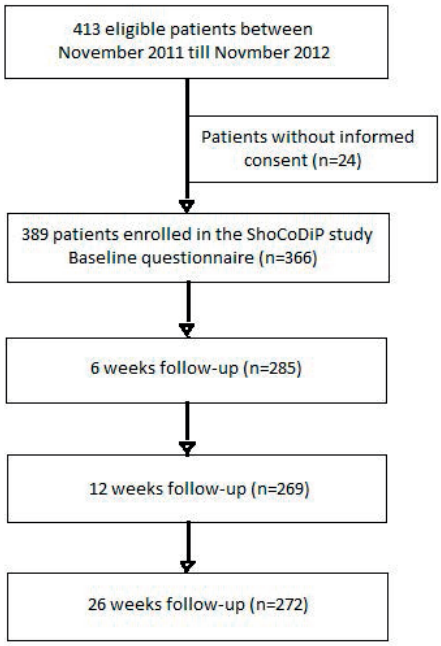


Figure 1. Flow Diagram

Clinical course

After 6 weeks follow-up 118 (41%) patients were recovered; 152 (57%) after 12 weeks and 164 (60%) after 26 weeks. Recovery rates in the working population were slightly higher; 91 patients recovered after 6 weeks (46%), 110 (60%) after 12 weeks and 119 (65%) after 26 weeks.

Median (IQR) SPADI score decreased from 49.5 (29-65) at baseline to 16.9 (3.9-43.0) at 26 weeks (Figure 2) and the NRS median score (IQR) decreased (Figure 3) from 6 (4-7) to 2 (1-5). For the working population, the disability score decreased from 44.9 (27-61) at baseline to 12.7 (3-35) at 26 weeks and pain score decreased from 6 (4-7) to 2 (0-5)

Table 1 Baseline characteristics

Baseline characteristics	Total population (n=389)	Working population (n=261)	Available data (%)
	Sociodemographic		
Age (years) mean (SD)	49.9 (13.2)	45 (10.7)	374 (96)
Male, n (%)	170 (45)	121 (46)	376 (97)
Educational level, n (%)			
Low	40 (11)	16 (6)	366 (94)
Medium	199 (54)	142 (56)	
High	127 (35)	98 (38)	
Paid work, n (%)	261 (71)	-	368 (95)
Full time, n (%)	-	136 (53)	257 (98)
Job description, n (%)			
Physically heavy work	-	64 (25)	258 (99)
Static repetitive work	-	88 (34)	
Work in awkward postures	-	11 (37)	
Work related complaints, n (%)	-	167 (69)	238 (91)
Sick leave, n (%)	-	40 (16)	257 (98)
	Clinical characteristics		
Duration in weeks, med (IQR)	12 (6-26)	12 (5-26)	371 (95)
Recurrent episode, n (%)	158 (43)	111 (44)	364 (94)
Dominant side affected, n (%)	224 (61)	159 (62)	369 (95)
Comorbidity, n (%)	236 (65)	156 (60)	364 (94)
Pain score NRS, med (IQR)	6.0 (4-7)	6.0 (4-7)	373 (96)
SPADI, med (IQR)	49.5 (29-65)	44.9 (27-61)	367 (94)
	Psycho-social characteristic		
Fear/depression EQ5D, n (%)			
not anxious/depressed	300 (83)	209 (83)	360 (93)
moderately	59 (16)	42 (16)	
anxious/depressed			
extremely	1 (0)	0 (0)	
anxious/depressed			
	Other		
Diagnostic US performed, n (%)	122 (31)	67 (26)	389 (100)
Working alliance, mean (SD)	45.3 (9.1)	46.7 (9.6)	87 (22)

N number, SD standard deviation, IQR Interquartile range, med median, NRS Numeric Rating Scale, SPADI Shoulder Pain and Disability Index, EQ-5D EuroQOL 5 Dimensions, US Ultrasound

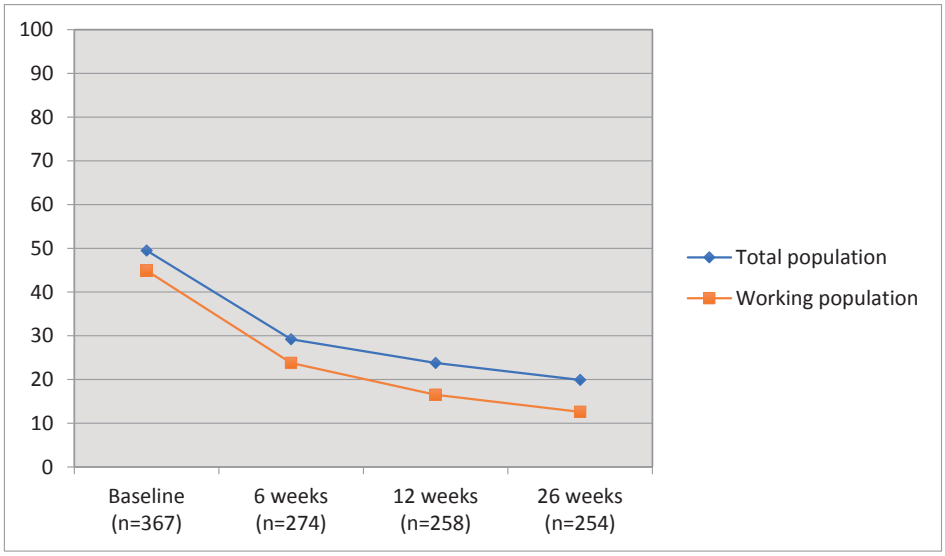


Figure 2. Median scores of disability (SPADI) at baseline, 6, 12 and 26 weeks follow-up.

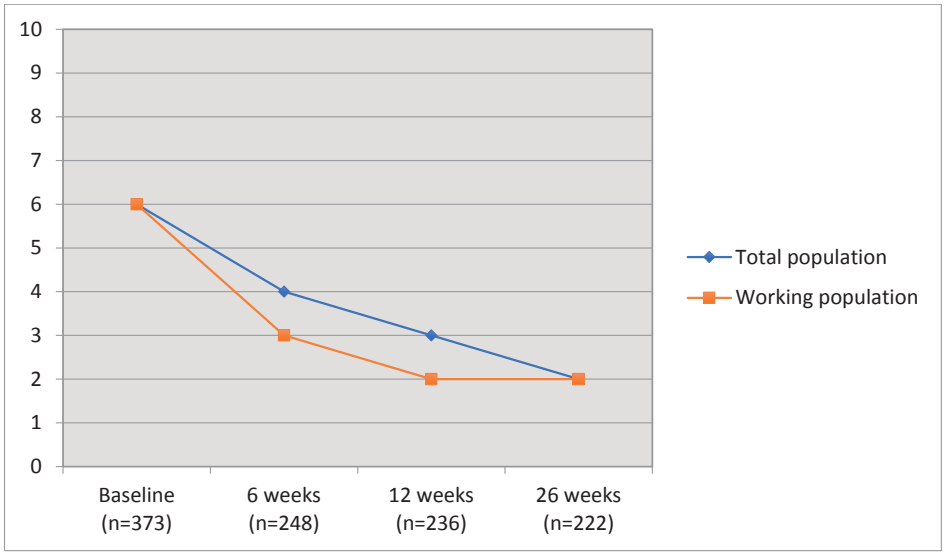


Figure 3. Median scores of pain severity (NRS-11) at baseline, 6, 12 and 26 weeks follow-up.

Table 2. Univariable & multivariable associations with recovery at 26 weeks.

	Total population (n=389) OR [95% CI] Beta		Working population (n=261) OR [95% CI] Beta	
Prognostic factors	Univariable	Multivariable	Univariable	Multivariable
Sociodemographic variables				
Age (years)	0.98[0.96-1.00]*† -0.017	0.99 [0.96-1.02] † -0.008	0.99 [0.97-1.02] † -0.006	1.01 [0.98-1.05] † 0.009
Female	0.9 [0.6-1.6] -0.058	1.1 [0.6-2.0] 0.307	0.9 [0.5-1.7] -0.072	2.0 [0.7-5.3] 0.690
Educational level				
Low	1.0	1.0	1.0	1.0
	0.7 [0.3-1.8]	0.4 [0.2-1.1]	0.6 [0.1-2.6]	0.5 [0.1-2.2]
Medium	-0.348	0.486	-0.451	-0.696
	0.9 [0.4-2.2]	0.5 [0.2-1.2]	0.8 [0.2-3.5]	0.7 [0.1-3.1]
High	-0.078	0.499	-0.101	-0.391
Clinical characteristics				
Duration in weeks	0.99 [0.99-1.00]** † -0.006	0.99 [0.99-0.99]** † -0.006	0.99 [0.99-1.00]** † -0.005	0.99 [0.99-1.00]** † -0.007
Recurrent episode (no)	1.7 [1.0-2.7]** 0.506	1.4 [0.8-2.5] 0.329	1.8 [0.9-3.4]** 0.562	1.5 [0.8-3.1] 0.435
Comorbidity (no)	1.3 [0.7-2.4] 0.270	1.0 [0.5-2.1] 0.012	1.1 [0.6-2.1] 0.111	0.9 [0.4-2.0] -0.084
Pain score NRS	0.9 [0.8-1.0]** -0.133	1.0 [0.8-1.2] 0.010	0.9 [0.8-1.0]* -0.120	1.0 [0.8-1.3] -0.004
Disability score, SPADI	0.98 [0.97-1.00]** † -0.017	0.99 [0.97-1.00] † -0.014	0.98 [0.97-1.00]** † -0.018	0.98 [0.96-1.01] † -0.017
Work related characteristics				
Paid work (no)	0.5 [0.3-0.9]** -0.667	0.6 [0.3-1.2] -0.583		
Full time (no)			0.6 [0.3-1.2]* -0.472	0.5 [0.2-1.2] -0.799
Job description				
Physically heavy work			0.8 [0.3-1.7]	0.9 [0.4-2.3]
			-0.276	-0.091
Static repetitive work			1.1 [0.5-2.4]	1.4 [0.6-3.4]
			0.142	0.352
Work in awkward postures			1.0 [0.2-4.4]	2.0 [0.3-12.1]
			0.094	0.710
Other			1.0	1.0

Table 2. Univariable & multivariable associations with recovery at 26 weeks. (continued)

	Total population (n=389) OR [95% CI] Beta		Working population (n=261) OR [95% CI] Beta	
Prognostic factors	Univariable	Multivariable	Univariable	Multivariable
Work related complaints (no)			0.5 [0.2-1.8] -0.538	0.4 [0.1-1.6] -0.834
Sick leave (no)			0.9 [0.3-2.4] 0.225	1.3 [0.5-3.9] 0.295
Psycho-social characteristics				
Fear/depression, EQ5D,				
No feelings of	1.9 [1.0-3.3]**	2.0 [0.9-4.0]	1.9 [0.9-4.0]*	1.8 [0.7-4.3]
anxiety/depression	0.518	0.655	0.532	0.566
Other				
Diagnostic US performed (no)	1.5 [0.9-2.4]* 0.394	1.2 [0.7-2.2] 0.174	1.4 [0.8-2.7] 0.340	1.3 [0.6-2.8] 0.264
Working alliance	1.0 [1.0-1.1] 0.010	1.0 [0.9-1.1] 0.010	1.0 [1.0-1.1] 0.010	1.0 [0.9-1.1] 0.009

OR: Odds Ratio, CI: Confidence Interval, SPADI: Shoulder Pain and Disability Index, NRS: Numeric Rating Scale, EQ-5D: EuroQOL 5 Dimensions
** P <0.10
* P <0.20
† rounded off with 2 decimals because of small CI

Predictors and model evaluation

All predictors

For all variables included in the model the variance inflation factors were < 1.5 and correlation coefficients <0.8, suggesting that linearity and multicollinearity was not a problem. In the univariable regression analysis, 8 factors were related (P<0.20) with recovery at 26 weeks (Table 2). There was only one patient who scored "very anxious/ depressed" on the depression score of the EQ-5D and therefore this answer option was combined with 'moderately depressed' and the EQ-5D was thus dichotomized in the regression analysis.

First we tested a model that included all prognostic variables (n=12) selected from the literature (Table 2). The R^2 was 0.17 and the ROC curve demonstrated a fair discriminating ability for the regression model with an AUC of 0.70 (95% CI 0.36-1.03) and correctly classified 66% of patients. The model in the working population resulted in similar results (see table 2). The R^2 for the working population was 0.19 and the AUC was 0.72 (95% CI 0.37-1.10) and the model correctly classified 69% of patients.

Backward regression analysis

Results from the backward regression resulted in a model where: a short duration of complaints, lower disability score, having a paid job, no feelings of depression/anxiety and high working alliance were related to recovery (table 3). The R^2 was 0.12 and the AUC was 0.67 (95% CI 0.34-1.0) and the model correctly classified 65% of patients.

In the working population we found identical results (table 3). The final model showed a short duration of complaints and low disability scores were related to recovery. The R^2 was 0.05 and the AUC was 0.63 (95% CI 0.25-1.00) and the model correctly classified 67% of patients.

Table 3 Final model; results from backward logistic regression

Final model after Backward Wald regression for recovery				
	Total population (n=389)		Working population (n=261)	
	OR [95% CI]	Beta	OR [95% CI]	Beta
Duration in weeks	0.99 [0.99-1.00]* †	-0.007*	0.99 [0.99-1.00]* †	-0.006*
Disability score, SPADI	0.99 [0.97-1.00]* †	-0.014*	0.98 [0.97-1.00]* †	-0.017*
Paid work (no)	0.6 [0.3-1.0]*	-0.592*		
Fear/depression, EQ5D,				
No Feelings of anxiety/depression	1.8 [0.9-3.6]	0.588		
Working Alliance	1.0 [0.9-3.6]	0.004		
Performance measures				
R^2	0.12		0.05	
AUC	0.67		0.63	
Bootstrapped AUC	0.66		0.63	
Final model after Backward Wald regression for pain				
Recurrent episode (no)	NA	0.738*	NA	0.779*
Duration in weeks	NA	0.004*	NA	0.005
Disability score, SPADI	NA	0.031*	NA	0.034*
Performance Measures				
R^2	0.13		0.15	

OR odds ratio, CI confidence interval, SPADI Shoulder Pain And Disability Index, EQ5D EuroQol 5 dimensions, AUC Area Under the Curve, R^2 R Squared

* p-value <0.05

† rounded off with 2 decimals because of small CI

Secondary outcome

Using pain as outcome resulted in a model including duration of complaints, recurrent episode and disability score in both the total (R^2=0.13) and working population (R^2=0.15).

Internal validation

Bootstrap method to assess optimism was checked in all prediction models (full and final model after backward elimination) for the primary outcome measure. Discriminative ability decreased in all models after bootstrap. The apparent performance (bootstrap corrected AUC) of the full model in the total population decreased from 0.70 to 0.67. The expected optimism for the AUC of the total population in the full model was 0.024 and 0.0409 in the working population. Optimism of the final model in the total population was 0.008 and 0.002 in the working population (table 3).

DISCUSSION

Our study showed that a short duration of complaints, not having feelings of depression or anxiety, having a paid job, a better working alliance and a low disability score were predictors of recovery after 6 months. Duration of complaints and disability were also predictors of recovery in the working population. In the prediction model for pain a recurrent episode of shoulder pain, short duration of complaints and low disability scores, were the predictors in the final model.

In this prognostic cohort study 60% of patients reported to be recovered after 6 months. This is slightly higher than the 21-51% reported by studies in GP practice. [9, 27, 39]

In line with previous research we found that a shorter duration of symptoms and lower disability scores were significantly associated with recovery. [7, 10, 15, 40-42]

Other prognostic models found the predictors; age, gender, [10] repetitive movement [9] and co-morbidities, [9, 20, 27, 43] which we included as possible predictor but did not remain in the final model. The reason that we did not find co-morbidity to be a predictor might be due to the difference in defining co-morbidity. Like this study, one study formulated co-morbidity as musculoskeletal (yes/no) [20] but others only measured concomitant low back pain [9] or concomitant neck pain [27]. Furthermore, we only asked for the co-morbidities around the shoulder region. Several studies have shown that other co-morbidities (like obesity, headache) also has an impact on an individual's ability to recover. [44-46]

Contrary to our findings, previous studies have not found a significant association of psychosocial factors and shoulder complaints. [7] However, in studies including patients with complaints of the arm, neck and shoulder psychosocial factors appear to have a predictive effect on patient outcome. [20] This effect has not been found in the literature specific for patients with only shoulder pain. We included only one item about depression and anxiety from the EQ-5D. This variable was dichotomized which might contribute to a loss of information. However the variable remained in the final model. One other study found catastrophizing at baseline to be a predictor of function. [44]

Working alliance remained in the final model as well.

It has been suggested that patient reported outcome measures, such as recovery and pain, are sensitive to the effect of interactions between patients and treatment providers. [47] One review has shown that a good working alliance can improve treatment outcomes. [14] Also, good working alliance scores might result in higher levels of adherence. [48] Treatment adherence is important to achieve optimal treatment outcomes and it is widely accepted that a lack of adherence to long-term therapies result in poor treatment outcomes and high costs of health care. The argument is that a good working alliance could help patients adhere to the treatment regime. [48] A good working alliance is partially determined by the communication between the patient and therapist. For that reason effective communication should be an essential skill that therapists need to master in order to improve health care.

Various other studies suggest that working alliance is associated with recovery in physical rehabilitation settings, but more research is needed to determine the strength of the possible relationship between the therapeutic alliance and recovery. [14]

Strength of this study is that we evaluated the prognostic value of two new variables, working alliance and the use of diagnostic ultrasound, upon variables that were described before. Furthermore the number of potential prognostic variables was not large, leading to more valid statistical derivations. [49, 50] There is a possibility that variables not mentioned in the literature were left out of this model but might have been significant predictors in our population.

In the model the use of diagnostic US was added as a dichotomous variable. This is because we assumed that a more specific diagnosis, as found using diagnostic US, leads to a more specific treatment and should lead to better patient outcomes. The low number of patients with an US diagnosis limited our ability to perform any additional analysis.

The percentage of missing values for the outcome was 30% after 6 months follow-up. Missing data was handled adequately with multiple imputations, although the large amount of missing data for working alliance might influence the validity of the data.

The model's performance is likely to be overestimated in the developmental dataset. Therefore we assessed the amount of optimism and corrected by using bootstrapping techniques to internally validate the model. The expected optimism after internal validation was small in all but one model. The optimism in the full model of the working population was substantial, probably due to the relatively small sample size. Similar levels of optimism have been observed earlier in smaller sample sizes. [50, 51] Furthermore the performance of the final model was not very good. Several 95% CI's around the AUC estimates crossed the 0.50 threshold indicating a high likelihood of poor discrimination.

All patients received physiotherapy treatment but it consisted of several treatment modalities resulting in heterogeneity. Besides heterogeneity in treatment, patients with

more severe complaints are more likely to receive more treatment sessions thus possibly influencing recovery status.

Future research.

Based on the relatively low AUC scores the prognostic model could be improved by possibly adding other psychosocial factors besides depression/anxiety and evaluate if the physiotherapy treatment and the number of treatment sessions could cause interaction effects. Hardly any prognostic models are routinely used in clinical practice, probably because most have not been externally validated. [52] It is crucial to quantify the performance of a prognostic model in different populations before applying it in daily practice. Since prognostic models in primary care for patients with shoulder pain seem to have similar performance estimates the next step might be to externally validate a high quality model with appropriate performance/discrimination in a new dataset. [9, 53, 54]

CONCLUSION

We developed and internally validated a model predicting recovery of patients with shoulder complaints in physiotherapy practice. Other variables should be evaluated to improve predictive capacity of the model and next the model should be externally validated before it can be used in clinical practice. In daily practice physiotherapists constantly predict the risk or probability of an individual to recover. Based on the predicted prognosis they inform individual patients about the course of the disease or the choice for further treatment. Knowledge of the predictors described in literature can be informative for the physiotherapist for their prognostic potential. When a model performs well at external validation it will probably be a useful tool, as it may enhance communication. Nevertheless its impact on patient outcomes should be assessed using a clinical trial design.

Acknowledgements

We thank Caspar Looman for statistical support

We thank all the physiotherapists and patients for participating in this study

Funding

Authors declare no conflict of interest.

This study is funded by SIA-RAAK. The ministry of education has made this funding available for the innovation and promotion of research. This study is partly funded by a program grant of the Dutch Arthritis Foundation.

REFERENCES

1. Picavet, H., & Schouten, J. (2003). Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC3-study. *Pain,* 102(1-2), 167-178.
2. Luime, J., Koes, B., Hendriksen, I., Burdorf, A., Verhagen, A., Miedema, H., & Verhaar, J. (2004). Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scandinavian Journal of Rheumatology*, 33, 73-81.
3. Feleus, A., Bierma-Zeinstra, S., Miedema, H., Bernsen, R., Verhaar, J., & Koes, B. (2008). Incidence of non-traumatic complaints of arm, neck and shoulder in general practice. *Manual Therapy*, 13(5), 426-33.
4. Kooijman, M., Swinkels, I., van Dijk, C., de Bakker, D., & Veenhof, C. (2013). Patients with shoulder syndromes in general and physiotherapy practice: an observational study. *BMC Musculoskeletal Disorders*, 14(128).
5. Barten, D.J.A., Verberne, L.D.M., Koppes, L.L.J. Zorg door de fysiotherapeut – aanmelding [Internet]. NIVEL Zorgregistraties eerste lijn; 2015 [cited 2015 august 23]. Available from: www.nivel.nl/node/3097
6. Dinant, G.J., Buntinx, F., Butler, C.C. (2007). The necessary shift from diagnostic to prognostic research. *BMC Family Practice,* 13(8), 53.
7. Kuijpers, T., Van der Windt, D., Van der Heijden, G., & Bouter, L. (2004). Systematic review of prognostic cohort studies in shoulder disorders. *Pain*, 109, 420-431.
8. Keijsers, E., Feleus, A., Miedema, H., Koes, B., & Bierma-Zeinstra, S. (2010). Psychosocial factors predicted nonrecovery in both specific and nonspecific diagnoses at arm, neck, and shoulder. *Journal of Clinical Epidemiology*, 63, 1370-79.
9. Kuijpers, T., van der Windt, D., Boeke, A., Twisk, J., Vergouwe, Y., Bouter, L., & van der Heijden, G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain*, 120(3), 276-285.
10. Kennedy, C., Manno, M., Hogg-Johnson, S., Haines, T., Hurley, L., McKenzie, D., & Beaton, D. (2006). Prognosis in soft tissue disorders of the shoulder: predicting both change in disability and level of disability after treatment. *J Am Phys Ther Assoc.*, 86(7), 1018-1032.
11. Bruls, V.E., Bastiaenen, C.H., de Bie, R.A. (2015). Prognostic factors of complaints of arm, neck, and/or shoulder: a systematic review of prospective cohort studies. *Pain*, 156(5), 765-99.
12. Hanchard, N., Lenza, M., Handoll, H., & Takwoingi, Y. (2013). Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *The Cochrane Database of systematic reviews*, 30(4).
13. Wright, A., Wassinger, C., Frank, M., Michener, L., & Hegedus, E. (2013). Diagnostic accuracy of scapular physical examination tests for shoulder disorders: a systematic review. *British Journal of Sports Medicine*, 47(13), 886-92.
14. Hall, A., Ferreira, P., Maher, C., Latimer, J., & Ferreira, M. (2010). The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Physical Therapy*, 90(8), 1099-110.
15. Chester, R., Shepstone, L., Daniell, H., Sweeting, D., Lewis, J., Jerosch-Herold, C. (2013). Predicting response to physiotherapy treatment for musculoskeletal shoulder pain: a systematic review. *BMC Musculoskeletal Disorders,* 14(203).
16. Karel, Y., Scholten-Peeters, W., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R., van den Borne, M., .et.al. (2013). Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study.

BioMedCentral Musculoskeletal Disorders, 14(62).

17. Holtermann, A., Hansen, J., Burr, H., & Sogaard, K. (2010). Prognostic factors for long-term sickness absence among employees with neck-shoulder and low-back pain. *Scand J Work Environ Heal.*, 36(1), 34-41.
18. Zengh, X., Simpson, J., van der Windt, D., & Elliott, A. (2005). Data from a study of effectiveness suggested potential prognostic factors related to the patterns of shoulder pain. *Journal of Clinical Epidemiology*, 58(8), 823-830.
19. Ginn, K., & Cohen, M. (2003). Conservative treatment for shoulder pain: prognostic indicators of outcome. *Archives of physical medicine and. Rehabilitation*, 85(8), 1231-1235.
20. Karels, C., Bierma-Zeinstra, S., Burdorf, A., Verhagen, A., Nauta, A., & Koes, B. (2007). Social and psychological factors influenced the course of arm, neck and shoulder compaints. *Journal of Clinical Epidemiology*, 60, 839-848.
21. Thoomes-de Graaf, M., Scholten-Peeters, G., Duijn, E., Karel, Y., Koes, B., & Verhagen, A. (2014). The Dutch Shoulder Pain and Disability Index (SPADI): a reliability and validation study. *Quality of Life Research*, 24(6), 1515-9.
22. Stinckens, N., Ulburghs, A., & Claes, L. (2009). De werkalliantie als sleutelelement in het therapie gebeuren; meting met behulp van de WAV-12, de nederlandstalige verkorte versie van working alliance inventory. *Tijdschrift Klinische Psychologie*, 39, 44-60.
23. Hatcher, R., & Gillaspy, J. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, 16, 12-15.
24. Kamper, S., Ostelo, R., Knol, D., Maher, C., de Vet, H., & Hancock, M. (2010). Global perceived effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *Journal of Clinical Epidemiology*, 63(7), 760-766.
25. Brooks, R. (1996). EuroQol: the current state of play. *Health Policy*, 37(1), 53-72.
26. Bekkering, G., Hendriks, H., & Van Tulder, M. (2005). Prognostic factors for low back pain patients referred for physiotherapy. *Spine*, 30(16), 1881–1886.
27. Van der Windt, D., Koes, B., Boeke, A., Deville, W., De Jong, B., & Bouter, L. (1996). Shoulder disorders in general practice: prognostic indicators of outcome. British *Journal of General Practitioners*, 46, 519-523.
28. Peduzzi, P., Concato, J., Kemper, E., Holford, T.R. & Feinstein, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12), 1373-9.
29. Moons, K., Donders, R., Stijnen, T., & Harrell, F. J. (2006). Using the outcome for imputation of missing predictor values was preferred. *Journal of Clinical Epidemiology*, 59, 1092-1101.
30. Janssen, K., Vergouwe, Y., Donders, A., Harrell, F. J., Chen, Q., Grobbee, D., & Moons, K. (2009). Dealing with Missing Predictor Values When Applying Clinical Prediction Models. *Clinical Chemistry*, 55(5).
31. van Buuren, S., Boshuizen, HC., Knook, DL. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine,* 18(6), 65-92.
32. Marshall, A., Altman, D., Holder, R., & Royston, P. (2009). Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Medical Research Methodology*, 9(57). doi:10.1186/1471-2288-9-57
33. Hesselman Borg, J., Westerståhl, M., Lundell, S., Madison, G., Aasa, U. (2016). Longitudinal study exploring factors associated with neck/shoulder pain at 52 years

of age. *Journal of Pain Research*, 24(9), 303-10

34. Heymans, M., van Buuren, S., Knol, D., van Mechelen, W., & de Vet, H. (2007). Variable selection under multiple imputation using bootstrap in a prognostic study. *BMC Medical Research Methodology*, *7*(33). doi:10.1186/1471-2288-7-33
35. Hosmer, D., Lemeshow, S., & Sturdivant, R. *Applied Logistic Regression*. New Jersey: John Wiley & Sons; 2013.
36. Steyerberg, E., Harrell Jr, F., Borsboom, J., Eijkemans, M., Vergouwe, Y., & Habbema, J. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, *54*(8), *774–781*.
37. Efron, B., & Tibshirani, R. *An introduction to the bootstrap*. Monographs on statistics and applied probability. New York: Chapman & Hall; 1993.
38. R Core Team. *R: A language and environment for statistical computing*; 2012. Available from R Foundation for Statistical Computing: http://www.R-project.org/
39. Croft, P., Pope, D., & Silman, A. (1996). The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. *British Medical Journal*, 313, 601-602.
40. Mintken, P., Cleland, J., Carpenter, K., Bieniek, M., Keirns, M., & Whitman, J. (2010). Some factors predict successful short-term outcomes in individuals with shoulder pain receiving cervicothoracic manipulation: a single-arm trial. *Physical Therapy*, 90(1), 26-42.
41. Engebretsen, K., Grotle, M., Bautz-Holter, E., Ekeberg, O., & Brox, J. (2010). Predictors of shoulder pain and disability index (SPADI) and work status after 1 year in patients with subacromial shoulder pain. *BMC Musculoskeletal Disorders*, 11(218).
42. Deutscher, D., Horn, S., Dickstein, R., Hart, D., Smout, R., Gutvirtz, M., & Ariel, I. (2009). Associations between treatment processes, patient characteristics, and outcomes in outpatient physical therapy practice. *Clinical Rheumatology*, 90(8), 1349-63.
43. Thomas, E., van der Windt, D.A., Hay, E.M.,Smidt, N., Dziedzic, K., Bouter, L.M., Croft, P.L. (2005). Two pragmatic trials for shoulder disorders in primary care: generalisability, course and prognostic indicators. *Annals of Rheumatic Diseases*, 64(7), 1056-61.
44. Bingefors, K., Isacson, D. (2004). Epidemiology, co-morbidity, and impact on health-related quality of life of self-reported headache and musculoskeletal pain – a gender persepective. *European Journal of Pain*, 8(5), 435-50
45. Coronado, R.A., Alappattu, M.J., Hart, D.L., George, S.Z. (2011). Total number and severity of comorbidities do not differ based on anatomical region of musculoskeletal pain. *Journal of Orthopaedis Sports and Physical Therapy*, 41(7), 477-85
46. George, S.Z., Beneciuk, J.M., Bialosky, J.E., Lentz, T.A., Zeppeiri, G. jr., Pei, Q., Qu, S.S. (2015). Development of a Review-of-Systems Screening Tool for Orthopaedic Physical Therapists: Results from the Optimal Screening for Prediction of Referral and Outcome (OSPRO) Cohort. *Journal of Orthopaedic Sports and Physical Therapy*, 45(7), 512-26
47. De Bruijn, C., de Bie, R., Geraets, J., Goossens, M., van den Heuvel, W., van den Heijden, G., et.al. (2007). Effect of an education and activation programme on functional limitations and patient-perceived recovery in acute and sub-acute shoulder complaints – a randomized clinical trial. *BMC Musculoskeletal Disorders*, 8(112).
48. Abdel-Tawaba, N., Roter, D. (2002). The relevance of client-centered communication to family planning settings in developing

countries: lessons from the Egyptian experience. *Soc Sci Med.*, 54, 1357–1368

49. Schönberger, M., Humle, F., Zeeman, P., Teasdale, TW. (2006). Working alliance and patient compliance in brain injury rehabilitation and their relation to psychosocial outcome. *Neuropsychol Rehabil*, 16, 298–314.
50. Peduzzi, P., Concato, J., Kemper, E., Holford, T., & Feinstein, A. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49, 1373-1379.
51. Steyerberg, E., Bleeker, S., Moll, H., Grobbee, D., & Moons, K. (2003). Internal and external validation of predictive models: A simulation study of bias and precision in small samples. *Journal of Clinical Epidemiology*, 56, 441-447.
52. Steyerberg, E., Harrell Jr, F., Borsboom, J., Eijkemans, M., Vergouwe, Y., & Habbema, J. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 54(8), 774–781.
53. Collins, GS., Reitsma, JB., Altman, DG., Moons, KGM. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Medicine*, 13(1).
54. Vergouw, D., Heymans, M.W., de Vet, H.C.W., van der Windt, D.A.W.M., van der Horst, H.E. (2011). Prediction of persistent shoulder pain in general practice: Comparing clinical consensus from a Delphi procedure with a statistical scoring system. *BMC Family Practice*, 12(63).

Chapter 7

Validity of the Flemish Working Alliance Inventory in a Dutch physiotherapy setting in patients with shoulder pain.

Published in Physiotherapy Theory and Practice

2018, 34(5)

Yasmaine Karel
Marloes Thoomes-de Graaf
Wendy Scholten-Peeters
Paulo Ferreira
Dimitris Rizopoulos
Bart Koes
Arianne Verhagen

INTRODUCTION

In physiotherapy practice patients usually follow a treatment regimen provided in coherence with the physiotherapist. This interaction between patient and therapist is referred to as a working alliance (WA). WA is first described in psychotherapy as the extent to which a client and therapist work collaboratively, purposefully and connect emotionally. WA is defined as a combination of 3 factors; agreement about the goals of treatment, the tasks of treatment and the bond between client and therapist [1].

For a treatment to be effective one important factor is that the patient complies with the regimen, after which health outcomes are more likely to improve [2]. Therefore it is essential for the therapist to provide a proper transfer of information about the goals and tasks of treatment for the patient in order to carry out the treatment regimen [3, 4]. Besides agreement about treatment goals and tasks, co-operation and compliance are achieved by means of bonding and trust between the therapist and the patient. Patients consult a physiotherapist because they seek help and they are in that case vulnerable. Help must therefore be offered and accepted based on trust. How this relationship will develop during the treatment period can have a significant impact on treatment outcome.

Several reviews have found that WA is a strong predictor of improvement in psychotherapy and psychology practices [5, 13]. Later research has established the importance of a good alliance also in other medical settings, such as in patients with ulcer disease, hypertension and diabetes [6, 7]. One review included 14 studies examining the patient-therapist relationship in physical rehabilitation setting [8]. In 9 studies a registered physiotherapist delivered the interventions. Results of the individual studies indicated that WA has a consistent positive correlation to treatment outcomes of pain, disability, physical/mental health and patient satisfaction [8]. A recent observational study of therapeutic alliance in patients with chronic low back pain confirmed these findings and found WA to be a consistent predictor of function, pain and disability measures [9]. WA might be more important in some therapies especially in those where treatment adherence represents an important component for treatment effect [10].

The Working Alliance Inventory (WAI) is one of the most commonly used and validated questionnaires to measure the working alliance [8]. It has been originally developed as a 36-item questionnaire based on Bordin's model measuring three domains; goal, task and bond [11, 12]. The WAI exists of one questionnaire for the client (WAI (C)) and one for the therapist (WAI (T)). Evidence suggests that the clients WA rating at the beginning of treatment is superior over the therapist rated version in predicting outcome [13].

The WAI was translated to Flemish, which is closely related to Dutch, named the "werk alliantie vragenlijst" (WAV). The 12 most indicative items were selected using confirmatory factor analysis to form the WAV-12 short form [14]. The WAV-12 has been used and validated in patients receiving psychotherapy in Belgium [15]. This study found a good

internal consistency for the three-factor model according to Bordin (task scale; correlation coefficient α=0.85, bond scale α=0.82, goal scale α=0.83). Correlations between the task and goal scales were good (correlation coefficient r=0.80) but correlations between the other scales were both lower (Cronbach's α=0.49). The WAV-12 used a 5-point likert scale instead of a 7-point likert scale in the original WAV-36. Therefore it is difficult to compare results from this validation study with other data. Literature does describe slightly higher correlation coefficients for the English and French short versions [14, 16]. A review has shown that translated versions of a measurement instrument for the neck do not guarantee similar measurement properties compared with the original instrument [17]. Cross-cultural validation in the Dutch population and physiotherapy setting is an important step to evaluate whether the underlying construct still holds for the WAV-12.

Therefore this study aims to investigate whether the WAV-12 is a valid measurement instrument in terms of the construct and discriminative abilities for a population of patients with shoulder pain in physiotherapy care.

METHODS

Study design

The study population consisted of patients with shoulder pain that participated in a prospective cohort study in patients consulting a physiotherapist for shoulder pain [18]. Recruitment period was from November 2011 till December 2012. The Research Committee of the Erasmus Medical Centre in Rotterdam approved the project (MEC-2011-414). After signing an informed consent patients were included and followed up for 6 months.

Participants

A total of 125 physiotherapists were invited to enrol patients. Patients consulting a physiotherapist were included if they suffered from shoulder pain, were aged≥18 years and had adequate understanding of the Dutch language. Patients were excluded if they had serious pathologies (infection, cancer or fracture), surgery of the shoulder in the previous 12 months, or had received diagnostic imaging techniques such as musculoskeletal ultrasound, magnetic resonance imaging or X-ray of the shoulder in the 3 months prior to start of the study. Patients included in the cohort study were followed for 6 months and received usual physiotherapy care. Questionnaires were sent by email at 6, 12 and 26 weeks and 2 reminders were sent after 2 and 4 days whenever the patient had not responded to the questionnaire.

Working Alliance (WA)

WA was measured 6 weeks after baseline for both the patient and physiotherapist, because earlier assessment would not clearly reflect the WA. We used the Flemish version of the WAI (WAV-12). It contains 12 items scored on a 5-point scale ranging from 1 ("never") to 5 ("always") and scoring is done for the total score and each subscale (goal, task and bond). The total score ranges from 12 (low WA) to 60 (high WA), and subscales range from 4 to 20. Where the patient had to fill in the name of the therapist we replaced the empty space with the words: "my therapist".

Statistical analysis

Descriptive data for demographic and symptom severity are presented as percentages for nominal variables (gender, level of education, cause of injury, first episode, reasons for stopping treatment) and as means for continuous variables (age, symptom duration). T-tests were used to test for differences in demographics between participants scoring all WAV-12 items and those who did not. Cronbach's alpha was used to assess the internal consistency of the WAV-12 and we assessed the correlation between patient and therapist scores using Pearsons' correlation coefficient. Coefficients equal or more than 0.7 were regarded as acceptable. R and SPSS v20.0 were used to conduct the analysis.

Validation

Performance of the items in the WAV-12 questionnaire was assessed with a partial credit Rasch model [19]. The response patterns from the set of available items in the questionnaire were tested against what is expected by the model that works according to a probabilistic form of Guttman scaling [20]. This scale assumes a deterministic pattern with a hierarchical ordering of items (low and high level of item scale). When a higher level of the item is affirmed, there must be a high probability that lower items will also be affirmed. The analysis gives the probability that a person will affirm an item of the difference between the person's level of working alliance and the level of working alliance expressed by the item.

The Rasch model was used to test; 1) internal validity of the construct, 2) whether specific items exhibit different properties in different subgroups in the population (differential item functioning) and 3) whether item redundancy can be considered [21]. Analysis was done using the ltm package in the statistical programing language R [22].

Firstly a one partial credit model with the discrimination parameter fixed at one was tested to check whether it fits the data. If this model did not fit the data an extended partial credit model with a common discrimination parameter not constrained at one or separate discrimination parameters for each parameter was considered. Uni-dimensionality could further be examined to investigate if the test variance is attributable to the principal factor or construct, estimated with Cronbach's alpha. Due to the fact that some patient responses were missing, multiple imputations were utilized to calculate Cronbach's alpha.

Differential item functioning was examined based on a likelihood ratio X^2 test implemented in the lordif package in R. Expected scores for each item should remain the same whether, an older or younger person (<50, which was the mean age) and a man or women scores the same item.

Rasch analysis can be useful and psychometrically sound in modifying measurement instruments [23]. Different criteria could be considered for item redundancy: High Item Characteristic Curve (ICC), low ICC or items having similar calibrations.

RESULTS

Study population

Sixty-six physiotherapists enrolled in total 389 patients. Physiotherapists were 72% male and had a mean working experience of 15 years.

Of the 389 patients 43% were male, average age was 50 years with a mean duration of shoulder pain of 33 weeks (see table 1). At baseline only 4% of the patients did not fill out the baseline questionnaire. At 6 weeks 30% of the responses were lost to follow up.

Working alliance

Seventy-eight patients (22%) filled in all the WAV-12 questions, enabling us to calculate a total score. The mean WAV score was 45 on a total range of 24 to 60, which is slightly above 50% of the maximum score. Most patients did not answer one or more questions of the WAV-12. The population that had responded to all WAV-12 questions did not significantly differ at baseline with the patients that did not (see table 1). Even though not statistically significant, the difference for duration of complaints appeared to be large. Selective responses can therefore not be excluded. The questions with the most missing values are questions 1, 3, 7 and 9 (see figure 1). Question 3, 7 and 9 are part of the "bond" subscale and question 1 is part of the "goal" subscale. The working alliance score of therapists was 52 and for patients 45. WAV-12 scores between patient and therapist had a poor correlation (r=0.30).

Validity of WAV-12

Of all patients, 274 had at least filled in one or more items of the WAV-12. Three models were fitted to the data. The first model (RASCH) assumes the discrimination parameter is equal for all items and fixed at one. The second model (1PL) assumes the discrimination parameter is equal for all items but is estimated from the data and the third model (gpcm) assumes the discrimination parameter is free to vary across items. Likelihood ratio tests between these models showed that the third model provided the best fit to the data (p=<0.001).

Table 1. Baseline characteristics

Characteristics of cohort	Total n=389	Participants filling in all items of WAV-12; n=87	Participants, missing 1 or more items of WAV-12; n=302
Male (%)	170 (43)	41 (49)	129 (44)
Age (SD)	50 (13)	50 (14)	50 (13)
Duration of complaint in weeks (SD)	33 (82)	27 (58)	34 (88)
Comorbidity (%)			
No	128 (35)	25 (29)	103 (34)
Yes	236 (65)	62 (71)	199 (66)
Medication use (%)	183 (47)	40 (49)	144 (50)
Highest education (%)			
Primary school	40 (10)	12 (15)	28 (10)
High school	199 (51)	44 (54)	155 (54)
University or applied sciences	127 (33)	25 (31)	102 (36)
Paid job (%)	261 (67)	53 (65)	208 (72)
Profession (%)			
Physically intensive job	65 (17)	13 (25)	52 (25)
Static repetitive job	88 (23)	14 (27)	74 (35)
Job with awkward positions/postures			
Other	11 (3)	3 (6)	8 (4)
	99 (25)	22 (42)	77 (36)
NRS median (IQR)	6.0 (3.0)	6.0 (2.0)	6.0 (3.0)
SDQ (SD)	62 (23)	63 (24)	62 (23)
EQ-5D (SD)	0.83 (0.08)	0.82 (0.07)	0.83 (0.09)

NRS Numeric Rating Scale, SDQ Shoulder Disability Index, EQ-5D EuroQol 5 Dimensions, SD standard deviation

Item properties

All but two items (item 1 and 2), showed ceiling effects, meaning that most of the patients scored a good working alliance. Appendix 1 displays the item characteristic curves for the 12 items from the WAV-12. Items 5, 6 and 8 have a high slope and are endorsed at higher levels of working alliance. Items 1, 2 and 4 have a low slope (discrimination) and are endorsed at lower levels of WA. Considerable variation exists between item discrimination indicating the WAV-12 questionnaire includes items measuring the whole construct and items discriminating at lower and higher levels of working alliance (table 2). The item information curve showed the amount of information given by the questionnaire is highest between an ability of -2 and 0, implying that the item set is most useful in discriminating among individuals at the lower end of the working alliance trait.

Table 2. Discrimination values of WAV-12 items

Item	Discrimination	Standard error	Z value
1	0.496	0.103	4.793
2	0.443	0.088	5.066
3	1.286	0.225	5.716
4	0.761	0.118	6.424
5	2.212	0.457	4.842
6	2.067	0.338	6.114
7	1.377	0.234	5.895
8	2.266	0.369	6.139
9	1.151	0.208	5.537
10	1.068	0.158	6.742
11	1.414	0.224	6.319
12	1.107	0.167	6.613

Unidimensionality

Five imputed datasets were created. Cronbachs alpha's were calculated for the 12 items in each dataset and led to a pooled cronbach's alpha coefficient of 0.89. Indicating that items correlate highly and measure the same explanatory concept.

Differential Item Functioning (DIF)

The X^2 tested three models. Model 1 is a standard model where the ability for each person remains the same. Model 2 tests whether levels of ability differ among groups and model 3 adds an interaction term for the level of ability and the group in order to test whether discrimination parameters differ among groups.

Age was dichotomized in younger patients (under the mean age of 50) and older patients (50 and over). The X^2 tested flagged item one for differential item functioning where all models were statistically significant. No differential item functioning was found between men and women. Slightly higher factor scores (mean difference = 0.0385) for the WA in patients being treated by a physiotherapist with less than 13 years of experience but was not statistically significant (p=0.73).

Rasch analysis for the WAV-12 questionnaire indicates that items have good discriminative abilities for the lower end of the construct. High correlations coefficients indicate items measure one construct and other factors like age and experience of the physiotherapist did not influence item scoring. Validity for the items in the questionnaire appears to be sound but due to the difference in the percentage of missing data among the items and observed ceiling effects we advise linguistic (Dutch) and contextual (physiotherapeutic setting) adjustments.

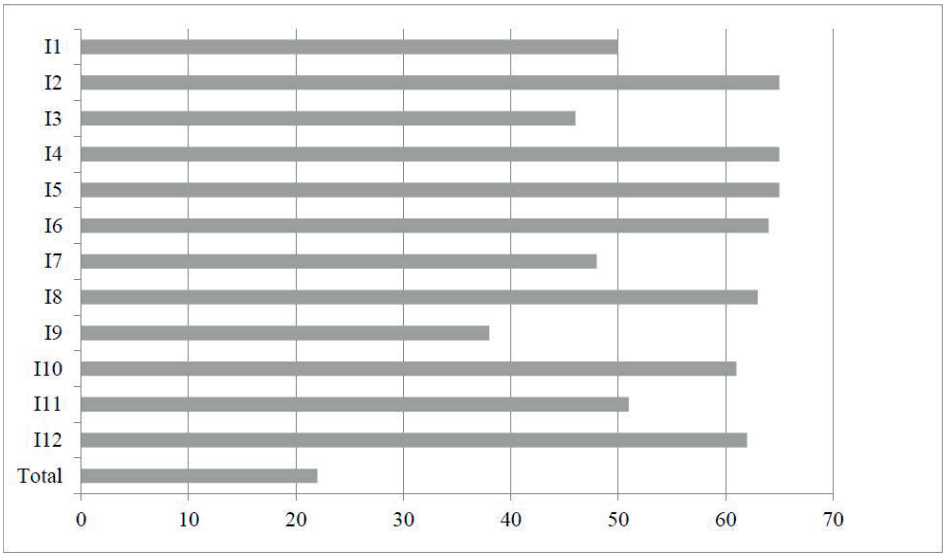


Figure 1. Relative response rate per item of WAV-12

Modification of the WAV-12

We believed rewording was necessary due to the selective number of missing responses in some items of the questionnaire and because the researchers had received comments from several patients and physiotherapists about items 3,7 and 9 of the WAV-12. Therefore we decided to make adjustments in the questionnaire and did a Delphi study. A 2 round survey was employed to ask the panels opinion on the adjustments in the WAV-12. The panel consisted of 11 members (6 clinical/research experts and 5 patients). Panel members were sent a questionnaire via email and these were sent separately to ensure panel members were unaware each of other's identity. For each item the panel member had to give his/her opinion about the adjustments with a 5-point likert scale. If the score was below 3 (neutral, disagree, totally disagree) the panel member were asked to give their reasoning and/or a suggestion for adjustment. If consensus for one item was < 80% after the first round it was included in the second round containing the suggestions of all panel members (anonymous). Full consensus (100% response rate) was reached after the second round and the adjusted questionnaire can be found in the attachments.

DISCUSSION

Main findings

Just a small proportion of patients filled in the complete WAV-12 compared to other questionnaires at 6 weeks follow-up. A large number of participants only completed a

limited number of items. This might indicate that the measurement instrument is not appropriate either in terms of language, setting, or participants had other specific reasons not to complete the questionnaire. The principal investigator also received comments from several patients and therapists, involved in the study, about items 3, 7 and 9 in the WAV-12 questionnaire. The construct theory of the WAV appeared to be sound but ceiling-effects were found in 10 items. Rewording was necessary for the WAV-12.

Comparison with the literature

Items correlated highly and measured the same explanatory concept which is found by several other translated versions of the WAI [11, 13, 14]. A French validation study found a very high correlation between the three subscales indicating that we cannot significantly distinct these subscales [16].

The poor correlation between patient and therapist WA score is consistent with other studies indicating that the two perspectives are not associated, which is confirmed by other studies as well [24, 25]. To ensure unbiased results the patient and the physiotherapist completed the rating forms independently of each other. Nevertheless, contact between the therapist and patient could not have been avoided, resulting in the possibility of deliberation between them.

WA was measured at 6 weeks when alliance might already have evolved into a stable situation whereas the first clinical experience between patient and therapist could determine more valid WA scores [26]. The literature is still inconsistent about what the optimal timing would be for measuring WA and some studies report that early WA predicted recovery after controlling for symptom change [27-30], while others have found a reduction of the predictive value of WA [31-33]. In this study WA was measured at six weeks as the first questionnaire was filled in before the first treatment. Nevertheless, we believe multiple measurements during the treatment period might yield more insight into the concept of WA.

Although WA is a valid construct within psychological interventions and research, whether it predicts recovery in a patient population in physiotherapy setting remains unknown. Psychological interventions are usually based on behavioural therapy that physiotherapists mostly use in chronic patients. The patient population in this study all have a new episode of shoulder pain where WA might be less relevant for the therapeutic process.

Strengths and limitations

This is the first study to perform a validation analysis on the Flemish version of the working alliance inventory in a physiotherapy setting. The measurement tool was able to discriminate between patients that experience a good or poor alliance. In ten items we observed ceiling effects, which might have been due to the fact that patients give socially

desirable answers or that the items do not properly assess the total construct. There appeared to be a pattern in missing items, where 4 items showed more missings than others, indicating that these might need adjustment. The questionnaire was developed in Belgium and applied in a Dutch setting which might not be appropriate given some linguistic characteristic differences of the Belgian Dutch (Flemish) and the Dutch language in the Netherlands. Due to the high number of missings in specific items (item 1, 3 and 9) and low discriminative values (item 1 and 2) we made changes in terms of adjustments in language and specific to the context of physiotherapy.

Implications for future research

The new questionnaire from our Delphi study has not been tested and therefore future research should test the psychometric properties of this questionnaire and evaluate the possible predictive value of the WA throughout the whole process of treatment in patients with musculoskeletal complaints. Whether measuring WA at the beginning or later in therapy is more predictive remains unknown. Studying a relationship between WA and recovery is complex because other factors, like self-adherence, compliance, might influence the relationship and therefore a mediation analysis might find more valid results.

CONCLUSIONS

The WAV-12 measurement tool is not suitable for implementation in clinical or research practice yet. However WA is a concept that needs attention within the field of physiotherapy and therefore we made adjustments to the questionnaire. Previous research has shown a positive correlation between working alliance and recovery in physiotherapy setting. Since shoulder pain can become a chronic condition in more than 50% of patients, interventions from physiotherapy need to be effective and a good WA can possibly contribute to optimal treatment effects.

Funding

The authors report no declarations of interest. This study is financed by the SIA-RAAK grant. The ministry of education has made this funding available for the innovation and promotion of research. This study is partly funded by a program grant of the Dutch Arthritis Foundation.

REFERENCES

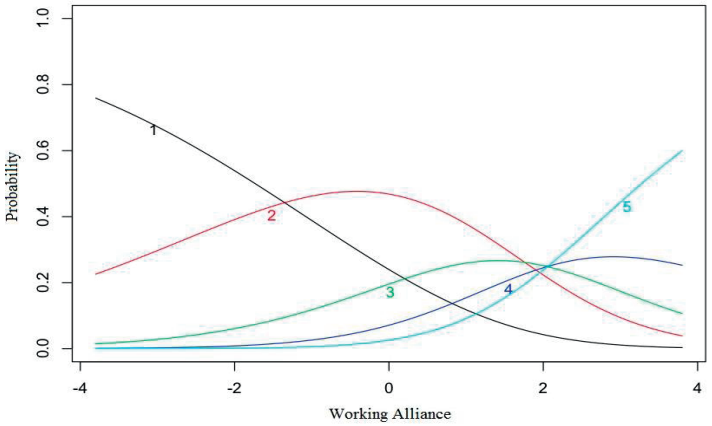
1. Bordin, E. (1979). The generalizability of the psychoanlaytic concepts of the working alliance. Psychotherapy: *Theory, Research and Practice,* 16, 252-60.
2. Bennett, J., Fuertes, J., Keitel, M., Phillips, R. (2011). The role of patient attachement and working alliance on patient adherence, satisfaction and health related quality of life in lupus treatment. *Patient education and counseling*, 85, 53-9
3. Sluijs, E., Kok, G., van der Zee, J. (1993). Correlates of Exercise Compliance in Physical Therapy. *Physical Therapy*, 73, 771-82.
4. Crow, R., Gage, H., Hampson, S., Hart, J., Kimber, A., Thomas, H. (1999). The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technology Assessment*, 3, 1-96.
5. Martin, D., Garske, J., Davis, K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 68, 438-50.
6. Kaplan, S., Greenfield, S., Ware, J. (1989). Assessing effects of physician-patient interactions on the outcomes of chronic disease. *Medical Care*. 27, S110-27.
7. Lee, Y., Lin, J. (2009). The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Social Science & Medicine*, 68, 1060-8.
8. Hall, A., Ferreira, P., Maher, C., Latimer, J., Ferreira, M. (2010). The Influence of the Therapist-Patient relationship on Treatment Outcome in Physical Rehabilitation: A Systematic Review. *Physical Therapy*, 90, 1099-110.
9. Ferreira, P., Ferreira, M., Maher, C., Refshauge, K., Latimer, J., Adams, R. (2013). The Therapeutic Alliance Between Clinicians and Patients Predicts Outcome in Chronic Low Back Pain. *Physical Therapy*, 93, 470-8.
10. Siev, J., Huppert, J., Chambless, D. (2009). The dodo bird, treatment technique, and disseminating empirically supported treatments. *The Behavior Therapist*, 32, 69-75.
11. Horvath, A., Lubrowsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology*, 61, 561-73.
12. Horvath, A., Greenberg, L. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223-33.
13. Horvath, A., Symonds, D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139-49.
14. Tracey, T., Kokotovic, A. (1989). Factor strructure of the working alliance inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207-10
15. Stinckens, N., Ulburghs, A., Claes, L. (2009). De werkalliantie als sleutelelement van de WAV-12, de Nederlandstalige verkorte versie van de Working Alliance Inventory. *Tijdschrift voor Klinische Psychologie*, 39, 44-60.
16. Corbière, M., Bisson, J., Lauzon, S., Ricard, N. (2006). Factorial validation of a French short-form of the Working Alliance Inventory. *International Journal of Methods in Psychiatric Research*, 15, 36-45.
17. Schellingerhout, J., Heymans, M., Verhagen, A., de Vet, H., Koes, B., Terwee, C. (2011). Measurement properties of translated versions of neck-specific questionnaires: a systematic review. *BMC Medical Research Methodology*, 11, 87.
18. Karel, Y., Scholten-Peeters, W., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R., van den Borne, M., Koes, B.W., Verhagen, A.P., Dinant, G.J., Tetteroo, E., Beumer, A., van Broekhoven, J.B., Heijmans, M. (2013).

Current management and prognostic facotrs in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. *BMC Musculoskeletal Disorders*, 14, 62

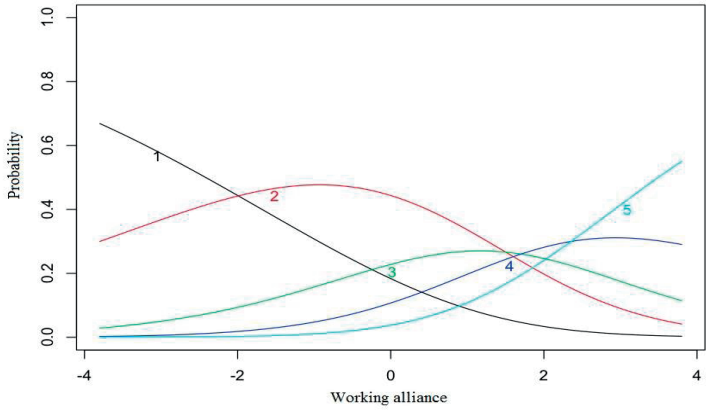
19. Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-74.
20. Guttman, L. (1950). The basis for Scalogram analysis, 4th edn. (I. S. Prediction, Ed.) New York: Wiley: The American Soldier.
21. Tennant, A., Conaghan, P. (2007) The Rasch Measurement Model in Rheumatology: What is it and When Use It? When Should It Be Applied and What Should One Look for in a Rasch Paper? *Arthritis & Rheumatism (arthritis Care & Research), 57*, 1358-62.
22. Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, 17.
23. Velozo, C., Lai, J., Mallinson, T., Haselman, E. (2001). Maintaining instrument quality while reducing items: application of Rasch analysis to a self-report of visual function. *Journal of Outcome Measurement*, 4, 667-80.
24. Huddy, V., Reeder, C., Kontis, D., Wykes, T., Stahl, D. (2012). The Effect of Working Alliance on Adherence and Outcome in Cognitive Remediation Therapy. *The Journal of Nervous and Mental Disease*, 200, 614-9.
25. Bachelor, A. (2013). Clients' and therapists' views of the therapeutic alliance: Similarities, differences and relationship to therapy outcome. *Clinical Psychology and Psychotherapy*, 20, 118-35.
26. Horvath, A. (2001). The Alliance. *Psychotherapy*, 38,365-372.
27. Barber, J., Connolly, M., Crits-Christoph, P., Gladis, L., Siqueland, L. (2000). Alliance Predicts Patient's Outcome beyond In-Treatment Change in Symptoms. *Journal of Consulting and Clinical Psychology*, 68, 1027-32.
28. Anker, M., Owen, J., Duncan, B., Sparks, J.A. (2010). The alliance in couple therapy: Partner influence, early change, and alliance patterns in a naturalistic sample. *Journal of Consulting and Clinical Psychology*, 78, 635-645.
29. De Bolle, M., Johnson, J., De Fruyt, F. (2010). Patient and clinician preceptions of therapeutic alliance as predictors of improvement in depression. *Psychotherapy and Psychosomatics*, 79, 378-385.
30. Klein, D., Schwartz, J., Santiago, N., Vivian, D., Vocisano, C., Castonguay, L., Arnow, B., Blalock, J.A., Manber, R., Markowitz, J.C., Riso, L.P., Rothbaum, B., McCullough, J.P., Thase, M.E., Borian, F.E., Miller, I.W., Keller, M.B. (2003). Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology*, 71, 997-1006.
31. Barber, J., Luborsky, L., Crits-Christoph, P., Thase, M., Weiss, R., Frank, A., Onken, L., Gallop, R. (1999). Therapeutic alliance as a predictor of outcome in treatment of cocaine dependance. *Psychotherapy Research,* 9, 54-73.
32. DeRubeis, R., Feeley, M. (1990). Determinants of change in cognitive therpy for depression. *Cognitive Therapy and Research*, 14, 469-482.
33. Puschner, B., Wolf, M., Kraft, S. (2008). Helping alliance and outcome in psychotherapy: What predicts what in routine outpatient treatment? *Osychotherapy Research*, 18, 167-178.

Appendix 1:

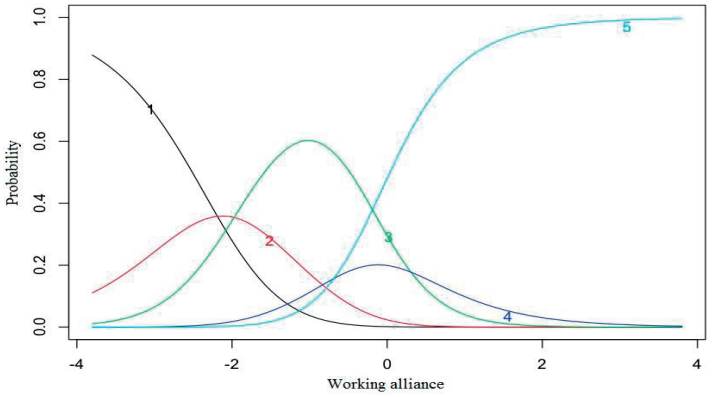
Item characteristic curves for the items in the WAV-12 questionnaire. Probability of working alliance score on the total construct for each response category of the item in different colours (1-5 likert scale).



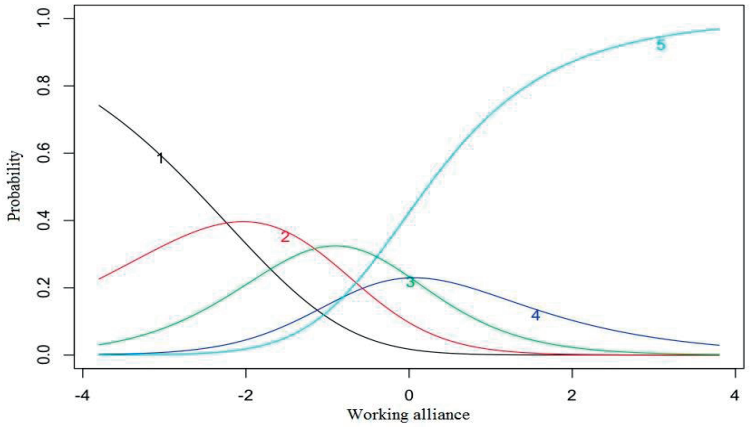
Item response category characteristic curve item 1



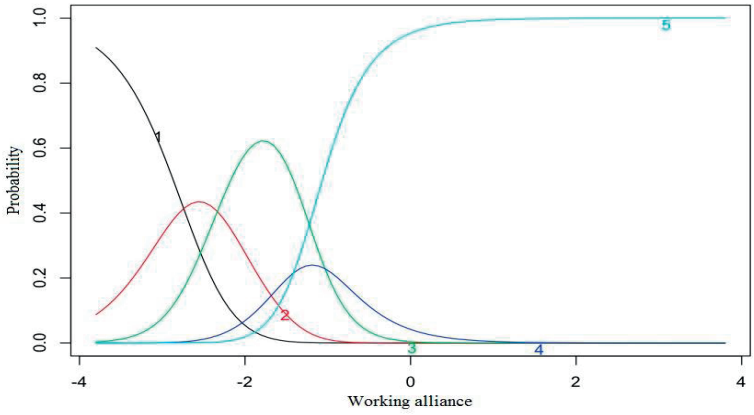
Item response category characteristic curve item 2



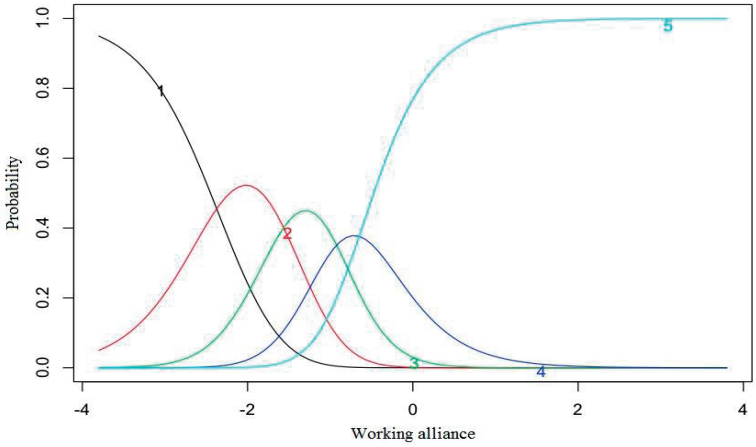
Item response category characteristic curve item 3



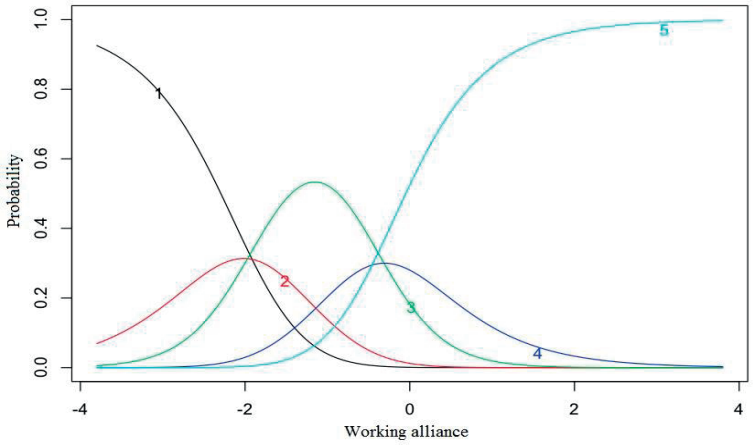
Item response category characteristic curve item 4



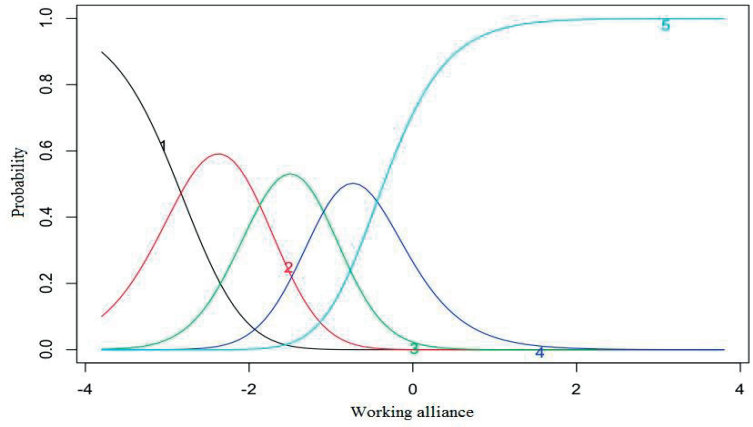
Item response category characteristic curve item 5



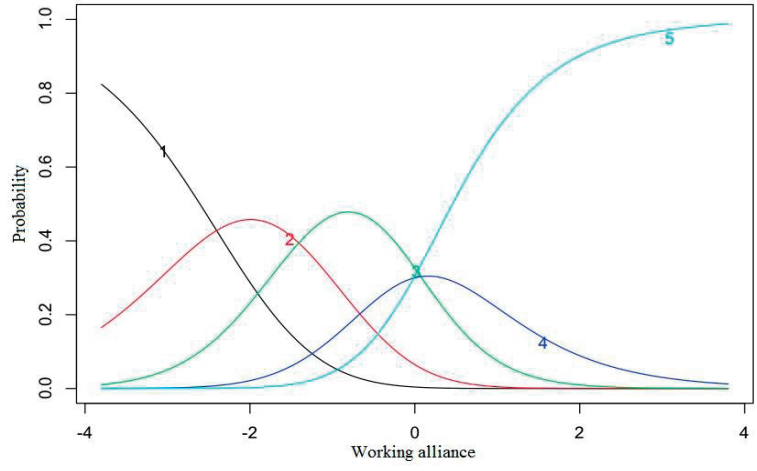
Item response category characteristic curve item 6



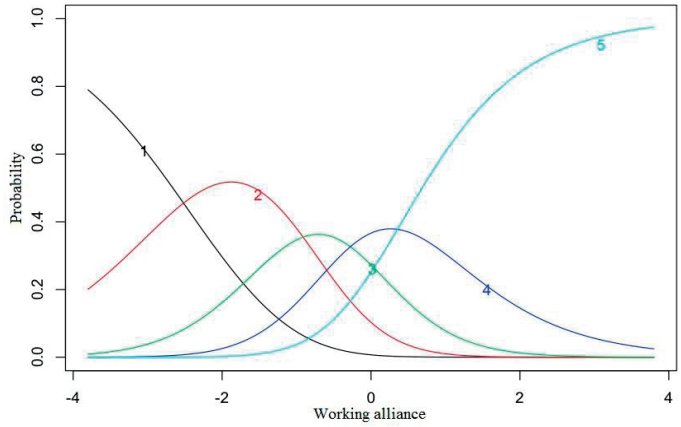
Item response category characteristic curve item 7



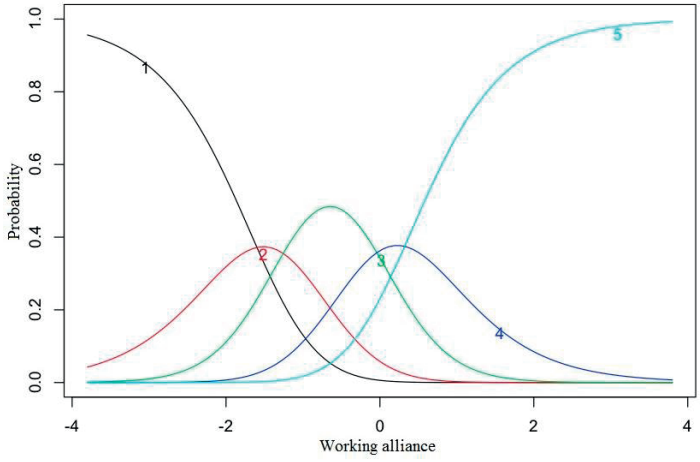
Item response category characteristic curve item 8



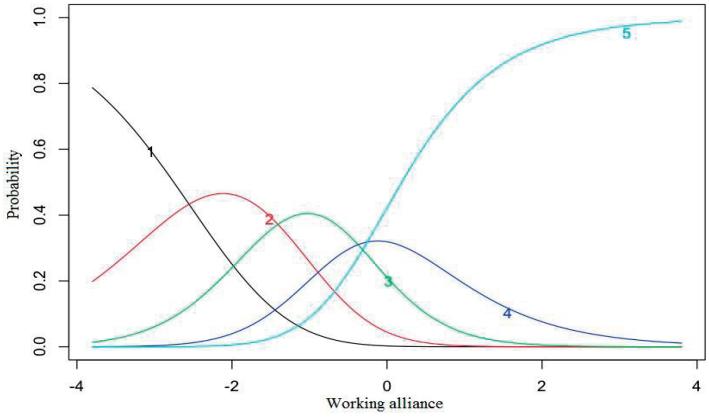
Item response category characteristic curve item 9



Item response category characteristic curve item 10



Item response category characteristic curve item 11



Item response category characteristic curve item 12

Chapter 8

General Discussion

GENERAL DISCUSSION

Primary aims of this thesis were a) to describe the current management in relation to diagnostic work-up (including the use of diagnostic ultrasound) and treatment strategies of physiotherapy care for patients with shoulder pain and b) to identify prognostic factors (including the use of diagnostic ultrasound and working alliance) of recovery for patients with shoulder pain. Firstly, we will discuss the study design, main findings and methodological considerations for current management and diagnostic ultrasound for patients with shoulder pain in physiotherapy care (chapters 2-5). Secondly, we will discuss the development of the prognostic model and the adjusted measurement instrument for the working alliance (chapter 6 and 7). Lastly, we will address implications for practice and recommendations for future research.

Study design

A considerable part of this thesis is based on a prospective cohort study performed in physiotherapy practice with a follow-up of 26 weeks in patients with shoulder pain. The study was conducted in the Southwest region of the Netherlands, between November 2011 and November 2012. We aimed to include as many physiotherapists (PTs) as possible to insure a successful patient recruitment. After sending out several emails to the physiotherapist network of the Avans University of Applied Sciences and organizing several recruitment meetings, 125 physiotherapists participated in the study. Although all 125 PTs did consent to participate, not all PTs did recruit patients for the study. Ultimately all participants were enrolled by 43% of the PTs with an average enrollment of 7 patients per PT. There was a wide variety in the characteristics of PTs participating in the study making selection bias unlikely. However, due to the study sub-question, related to diagnostic ultrasound, bias towards an increased selection of PTs using diagnostic US cannot be disregarded.

The participating PTs received a laminated card to quickly check patients' eligibility during the first consultation and 10 freepost envelopes with the information letter and informed consent. This made it fairly easy for the PTs to select and include the patients for this cohort.

During the recruitment period we continuously checked the number of patients who were recruited, as it is known that only 50% of the Dutch primary care studies succeed in recruiting their target number of patients [1]. An estimated 400 patients were needed based on a 40% recovery rate, 15 prognostic factors in the prognostic model and adjustments for 20% missing values. Eventually, our recruitment was successful as 412 patients were enrolled in the study. In total 389 patients provided us their informed consent. We took great care of designing the enrollment procedures as it is known that researchers are usually overly optimistic regarding recruitment [2]. To reach our target

we carried out several activities/interventions to stimulate patient recruitment. Firstly, participating PTs were regularly reminded of the study by sending out monthly newsletters about the number of patients that were a recruited, the average recruitment rate and how many patients the PT recruited themselves (mirror information). These newsletters also contained some new relevant scientific facts on shoulder pain or interesting conferences. Nevertheless, the number of patients that were enrolled throughout the year decreased and appropriate actions were taken to address these. Secondly, we organized a conference, where PTs that recruited >3 patients were offered to register for free and lastly, we aimed to increase the recruitment rate by offering the PTs accreditation points for their membership to the Dutch central quality register for PTs. An increase of the enrollment rate was observed after each stimulating intervention.

We tried to minimize selection bias through adequate participant selection and therefore designed recruitment methods that resulted in the most representative samples of clinicians and patients. This resulted in a cohort of patients with similar baseline characteristics compared to other studies conducted in primary care.

Another issue in a prospective cohort study can be the loss to follow-up rate. Loss to follow-up was minimized by sending out reminders to patients and PTs by email and telephone. A general rule of thumb requires that the loss to follow-up rate not surpasses 20% [3]. The loss to follow-up rate varied on the several follow-up moments between 31% and 28%; at 12-weeks the amount of missing data was highest (31%). Loss to follow-up mainly becomes a problem when there is a selective loss to follow-up. Fortunately, in our cohort we did not find any indication of selective loss to follow-up. The missing value analysis showed the data was missing (completely) at random and the necessary steps were taken to account for these missing data.

Current management: main findings

Diagnosis. The PTs rated most patients to suffer from a suspected subacromial impingement as primary hypothesis after history taking. This hypothesis commonly coexisted with the assumption of pathology of the glenohumeral joint or the cervico-thoracic spine. Nearly all PTs formulated multiple initial diagnostic hypotheses for each patient, reflecting that the diagnostic process in patients with shoulder pain is complex. After physical examination only a small number of PTs changed the primary initial hypothesis (which was based on history taking alone), indicating that additional physical examination did not provide additional information for the PT to change the initial hypothesis.

Diagnostic ultrasound (US) was performed in 31% of patients and was mostly done before the physical examination and in a substantial number of patients (38%) diagnostic US was performed instead of the physical examination. The PTs believed using diagnostic US would lead to a more specific clinical diagnosis or a more appropriate intervention compared to using physical examination. Semi-structured interviews with GPs showed

that general practitioners (GPs) believe that diagnostic US can lead to more accurate diagnoses as well [52]. Based on the systematic review of the literature no conclusions can be made for the different imaging procedures for shoulder pain patients on the efficacy and influence on patient recovery. Unfortunately only experimental studies were found that included patients with low back pain and knee pain and these studies show that a more specific clinical diagnosis did not lead to better patient reported outcome measures.

Tendinopathy of the rotator cuff was the most occurring pathology assessed by diagnostic US followed by calcification. A full thickness tear based on diagnostic US was found in 5% of patients. The initial hypothesis changed in 31% of the patients after receiving diagnostic US and usually changed to hypotheses such as sprain (trauma) or strain, suggesting that the pathology (tear or tendinopathy) found on diagnostic US determined the final clinical diagnosis. Of patients that had a diagnostic US, 16% were referred back to their GP compared to 8% in those without a diagnostic US. In most of the referred patients the diagnosis was calcific tendinitis or tendinopathy.

The high number of diagnostic US seems to reflect that PTs prefer to use a patho-anatomical diagnosis in the management of shoulder pain. However, previous research has shown that treatment based on a patho-anatomical diagnosis is not more effective than treatment based on signs and symptoms [5]. Results from our study indicate that diagnostic ultrasound does not need to be a standard diagnostic procedure in primary care physiotherapy. The changes in clinical diagnosis observed in our study (to the specific pathologies tendinopathy or calcification) do not increase the need of immediate care from a medical specialist. It might be appropriate in patients where a full thickness tear is suspected, as these cases might need surgical repair. Moreover, assessment of a full thickness tear by a PT shows adequate agreement compared to assessment by radiologists [6]. Furthermore, necessary actions should be taken to improve accuracy of operators and ultrasound findings should always be considered in the clinical context, as asymptomatic findings may be frequently found.

Diagnostic management for the physical examination and diagnostic US was not standardized as it aimed to reflect usual care. PTs that used diagnostic US were only trained in a standardized scanning protocol to ensure that all PTs reviewed the same anatomical structures. Standardizing the diagnostic process would bias the results of an observational study, but on the other hand it might lack validity of the hypotheses and pathologies found on diagnostic ultrasound.

We aimed to observe usual care, but at the same time one of our study questions concerned the use of diagnostic US. Consequently, there might have been more PTs that regularly use diagnostic US that participated in our study because of their specific interest. This could have resulted in an overestimation of the proportions of patients receiving diagnostic US in this study. PTs were considered to be experienced ultrasonographers,

but the criteria for determining the pathology on the diagnostic US are not yet fully developed and might differ between therapists and may have negatively influenced the diagnostic labels from diagnostic US.

Treatment. When studying usual care we can describe the variability between practitioners and assess consistency with recommendations from evidence based practice guidelines. The descriptive goal required measurements of a wide array of potential treatment processes as opposed to narrowly specified measurements. Specific features on treatment intensity (i.e. the specific exercise regimen or specificities of other interventions) were not measured because of feasibility. We also hypothesized that the choice of treatment was primarily based on the clinical diagnoses. However, the patient preferences, contextual factors or insurance policies may have been important factors which have influenced treatment choices, but these were not measured.

Interventions that were most frequently used for patients with a suspected subacromial impingement were advice and exercise therapy. For patients with a suspected cervical/thoracic dysfunction or frozen shoulder, advice and manual mobilization/manipulation of the spine were the most provided interventions. Using several interventions is common practice for PTs. The evidence statement for the management of subacromial complaints from the Royal Dutch Society of Physiotherapists (KNGF) also recommends a combination of advice and exercise therapy in the treatment of a suspected subacromial impingement [7]. Furthermore, the evidence statement suggests that extracorporeal shock wave therapy (ESWT) might be used in patients with a suspected calcification. When PTs suspected a patient to suffer from a calcification, based on diagnostic US findings, 32% (n=15) of patients indeed received ESWT. Overall, less than 10% of all patients were treated with a passive approach (massage, trigger point therapy or electrotherapy); these interventions are not recommended in the evidence statement.

Moreover, the evidence statement recommends 6-12 weeks of PT treatment. After 3 weeks 12% had ended treatment, 29% after 6 weeks and 59% after 12 weeks. A high number of patients still received treatment after 12 weeks (41%).

The evidence statement considers a referral to the GP when patients do not improve after 6 to 12 weeks. The high number of patients without improvement at 12 weeks (41%), as observed in this study, would have increased the number of referrals to the GP or medical specialist enormously. Although we do not know exactly how many patients were referred back after twelve weeks, we can assume that 41% would be an extremely high percentage that would be referred back when following the recommendations from the evidence statement. Furthermore, the higher referral rate (16%) in the diagnostic US group (compared to 8% in the non-US group), might also reflect an unjustified extra number of visits to medical experts or GPs.

Data collection. To collect relevant data of the diagnostic- and therapeutic process of the PTs regarding their management of patients with shoulder pain we used several

questionnaires, for patients and PTs. The questionnaire used by patients was developed using validated questionnaires. We developed a questionnaire for the data collection of the PTs in close collaboration with PTs. Even though the questionnaire was developed through several consensus steps and extensive piloting, the PTs might have misclassified some of the variables of diagnostic criteria or treatment modalities.

Patient therapist relationship (working alliance). The interaction between the patient and PT is considered to be a crucial part of the therapeutic process. In order to measure the working alliance, we used a Belgian-Dutch (Flemish) version of the working alliance inventory (WAV-12). Unfortunately, we found a high number of missing responses for specific items (especially on the bonding scale), probably due to the linguistic characteristics of these items. Therefore, we subsequently made changes in terms of adjustments in language and the context of physiotherapy in the WAV-12 using a Delphi consensus study involving patients, researchers and practitioners. This new version of the WAI-12; the Dutch Physio Alliance Scale (D-PAS), however, needs further validation.

Prognosis and recovery: main findings

Recovery. After 6 months 60% of patients were completely recovered according to the GPE (global perceived effect) scores. In the working population the recovery rate was slightly higher: 65%. The recovery rates found in our study were slightly higher than other studies in general practice [56, 57, 58]. It might be questioned whether this difference is due to the therapeutic interventions, the measurement instrument or because there might be a different population in PT practice compared to general practice. We excluded patients that had surgery in the past 6 months and all patients that received previous physiotherapy treatment for the same complaint, which is an important difference compared to other studies. An observational study in the Netherlands found that patient who were not referred to the PT by their GP are younger, more often have recurrent complaints and the complaints are more often related to sports or leisure activities [11]. On the other hand, our population did not seem to differ concerning baseline characteristics from other observational studies done in primary care.

Prognosis/prediction. We found the following prognostic factors for recovery after 6 months; a short duration of complaints; not having feelings of depression or anxiety; having a paid job; a better working alliance with their PT and a low disability score. Duration of complaints and disability were also predictors of recovery in the working population. Having a paid job and not having feelings of depression or anxiety were the strongest predictors. The predictors age, sex, repetitive movements and comorbidities, that were reported as predictors in the literature [12, 13, 14, 15], did not remain in our final prognostic model. This might be due to differences in measurement. We defined only upper limb co-morbidity, while other studies take into account all comorbidities or also measure concomitant low back- and neck pain. Because the prognostic model only

showed moderate performance it will be necessary in the future to include additional factors like psychosocial and emotional factors to improve the performance and discrimination of the model. Due to the lack of evidence of these factors in shoulder pain patients at the time we designed this study we did not include these variables in our project.
There are different ways to develop prognostic models, statistically and methodologically, all of which could lead to differences in the final prognostic models [16]. We selected prognostic factors based on the literature and presented all important performance statistics [13, 12, 17, 18, 19]. We added diagnostic US as a possible prognostic factor because imaging procedures might influence recovery [20] and we assumed that US would lead to a more specific diagnosis, a subsequent more specific treatment and thus to better patient outcomes. This variable was dichotomized however, the performance of diagnostic US might differ for different subgroups. The low number of patients with an ultrasound diagnosis limited our ability to perform any additional analysis for subgroups (e.g. rotator cuff tears)

Likewise, we added working alliance to the model because it showed significant associations in other musculoskeletal populations [21]. Working alliance remained in the final prognostic model, meaning that the relationship between the physiotherapist and the patient is an important factor for the treatment process.

Implications for practice

We suggest that the 12 week time frame for referral is not helpful in clinical practice, since surgical management and conservative management show similar results and the time needed for conservative management (mostly exercise therapy) to work might be longer than 12 weeks in patients with high levels of pain or disability [22, 23, 24]. For the patients that progress into chronic complaints, other factors (like psychosocial-emotional factors or central nervous system in pain behavior) should be taken into account [14].

Several studies showed that pathologies are found in asymptomatic shoulders and therefore the pathology seen on the US might not be the cause of the symptoms experienced. However, in case of full thickness tears (early) surgical repair is sometimes required and therefore it is necessary to identify these tears early in primary care [25, 26].

The use of prognostic models links to a shift towards stratified care, where the individual's profile and the presence of prognostic factors help guide individual treatment decisions. Although the prediction model is not yet valid for use in clinical practice, prognostic factors like duration of complaints, level of disability, having a paid job, working alliance and feelings of depression/anxiety should be addressed or taken into account when making a treatment decision in patients with shoulder pain.

Implications for future research

Almost all PTs in our study used exercise in all treatment sessions. However, it is not clear what specific type or dose of exercise can be recommended and future studies should investigate these or whether exercise therapy might be more effective for a specific subgroup of patients.

In our study, the use (yes/no) of diagnostic US (and not the outcome) was dichotomized to predict recovery. The effect of receiving a diagnostic ultrasound on recovery and reassurance might differ for different subgroups of patients and should be studied in a large controlled trial. The working alliance between PT and patient predicted recovery but further research is needed to explore the concept and impact of working alliance in physiotherapy care.

The explained variance of the prognostic model for recovery in this study was still only moderate, meaning that we cannot yet validate a useful prognostic tool to select patients that are at risk for chronicity. More factors should be taken into account to improve the explained variance in future prognostic models. Some of the prognostic factors are modifiable and future research should investigate whether changes in these factors can contribute to patient recovery.

REFERENCES

1. van der Wouden, J., Blankenstein, A., Huibers, M., van der Windt, D., Stalman, W.V.A. (2007). Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research. *J Clin Epidemiol,* 60.
2. van der Windt, D., Koes, B., van Aarst, M., Heemskerk, M., Bouter, L. (2000). Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment. *Br J Gen Pract,* 50(454), 371-4.
3. Fewtrell, M., Kennedy, K., Singhal, A., Martin, R., Ness, A., Hadders-Algra, M., Koletzko, B., Lucas, A. (2013). How much loss to follow-up is acceptable in long-term radomised trials and prospective studies? *Archives of Diseases in Childhood,* 93, 458-461.
4. Ottenheijm, R., Hesselmans, N., Kemper, A. Moser, A., de Bie, R., Dinant, G., Cals, J. (2014). GPs' perspectives on the diagnostic work-up in patients with shoulder pain: a qualitative study. *J Eval Clin Pract.,* 20(3), 239-45.
5. Dinant, G., Buntinx, F., Butler, C. (2007). The necessary shift from diagnostic to prognostic research. *BMC Family Practice,* 8(53).
6. Thoomes-de Graaf, M., Scholten-Peeters, G., Duijn, E., Karel, Y., van den Borne, M., Beumer, A., Ottenheijm, R., Dinant, G., Tetteroo, E., Lucas,C., Koes, B., Verhagen, A. (2014). Inter-professional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Manual Therapy,* 19(5), 478-83.
7. Jansen, M., Brooijmans, F., Geraets, J., Lenssen, A., Ottenheijm, R., Penning, L., de Bie, R. (2011). KNGF Evidence Statement Subacromiale klachten. *Supplement bij het Nederlands Tijdschrift voor Fysiotherapie,* 121(1).
8. Croft, P., Pope, D., Silman, A. (1996). Primary CareRheumatology Society Shoulder StudyGroup. The clinical course of shoulderpain: prospective cohort study in primarycare. B*MJ,* 313, 601-602.
9. Kuijpers, T., van der Windt, D., Boeke, A., Twisk, J., Vergouwe, Y., Bouter, L., van der Heijden, G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain,* 120(3), 276-85.
10. Van der Windt, D., Koes, B., Boeke, A., Devillé, W., De Jong, B., Bouter, L. (1996). Shoulder disorders in general practice: prognostic indicators of outcome. *Br J Gen Pract.,* 46(410), 519-23.
11. Kooijman, M., Swinkels, I., van Dijk, C., de Bakker, D., Veenhof, C. (2013). Patients with shoulder syndromes in general and physiotherapy practice: an observational study. *BMC Musculoskelet Disord.,* 14, 128.
12. Kuijpers, T., van der Windt, D., van derHeijden G., Bouter, L. (2004). Systematic review of prognostic cohort studies in shoulder disorders. *Pain,* 109, 420-431.
13. Kennedy, C., Manno, M., Hogg-Johnson, S., Haines, T., Hurley, L., McKenzie, D., Beaton, D. (2006). Prognosis in soft tissue disorders of the shoulder: predicting both change in disability and level of disability after treatment. *Phys Ther.,* 86(7), 1033-7.
14. Karels, C., Bierma-Zeinstra, S., Burdorf, A., Verhagen, A., Nauta, A., Koes, B. (2007). Social and psychological factorsinfluenced the course of arm, neck andshoulder compaints. *J Clin Epidemiol.,* 60, 839-848.
15. Thomas, E., van der Windt, D., Hay, E., Smidt, N., Dziedzic, K., Bouter, L., Croft, P. (2005). Two pragmatic trials for shoulder dis-orders in primary care: generalisability,course and prognostic indicators. *Ann Rheum Dis.,* 64(7).
16. Steyerberg, E., Moons, K., van der Windt, D., Hayden, J., Perel, P., Schroter, S., Riley,

R., Hemingway, H., Altman, D., Group, P. (2013). Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research, *PLoS Med,* 10(2), e1001381.

17. Ginn, K., Cohen, M. (2003). Conservative treat-ment for shoulder pain: prognostic indica-tors of outcome. *Arch Phys Med Rehabil,* 85, 1231–1235.
18. Zengh, X., Simpson, J., van der Windt, D., Elliott, A. (2005). Data from a study of effectiveness suggested potential prognostic factors related to the patterns of shoulder pain. *J Clin Epidemiol,* 58, 823-830.
19. Steyerberg, E., Vickers, A., Cook, N., Gerds, T., Gonen, M., Obuchowski, N.,Pencina, M., Kattan, M. (2010). Assess-ing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology.,* 21(1), 128-38.
20. Chou, R. Fu, R., Carrino, J., Deyo, R. (2009). Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet,* 373(9662), 463-472.
21. Hall, A., Ferreira, P., Maher, C., Latimer, J., Ferreira, M. (2010). The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a system-atic review. *Phys Ther.,* 90(8), 1099-110.
22. Brumitt, J., Cuddeford, T. (2015). Current concepts of muscle and tendon adaptation to strenght and conditioning. *Int J Sports Phys Ther,* 10(6), 748–759.
23. Reeves, N., Maganaris, C., Narici, M. (2003). Effect of strength training on human patella tendon mechanical properties of older individuals. *J Physiol,* 548, 971–981.
24. Kubo, K., Kanehisa, H,. Fukunaga, T. (2002). Effects of resistance and stretching training programmes on the viscoelastic properties of human tendon structures in vivo. *J Physiol,* 538, 219–226.
25. Millett, P., Wilcox, R., O'Holleran, J., War-ner, J. (2006). Rehabilitation of the rotator cuff: an evaluation-based approach. *J Am Acad Orthop Surg,* 14(11), 599-609.
26. Lähteenmäki, H., Hiltunen, A., Virolainen, P., Nelimarkka, O. (2007). Repair of full-thickness rotator cuff tears is recommended regardless of tear size and age: a retrospec-tive study of 218 patients. *J Shoulder Elbow Surg,* 16(5), 586-90.
27. Holtermann, A., Hansen, J.V., Burr, H., Sogaard, K. (2010). Prognostic factors for long-term sick-ness absence among employ-ees withneck-shoulder and low-back pain. *Scand J Work Environ Health.,* 36, 34-41.
28. Thoomes-de Graaf, M., Scholten-Peeters, G., Duijn, E., Karel, Y., van den Borne, M., Beumer, A., Ottenheijm, R., Dinant, G., Tetteroo, E., Lucas, C., Koes, B., Verhagen, A. (2014). Inter-professional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Man Ther.,* 19(5), 478-83.
29. Lewis, J. (2009). Rotator cuff tendinopathy/ subacromial impingement syndrome: is it time for a new method of assessment? *Br J Sports Med,* 43(4), 259-64.

Chapter 9

Summary

Samenvatting

Dankwoord

About the author

PhD Portfolio

List of Publications

SUMMARY

Patients with shoulder pain concern a large group of patients that visit the physiotherapist in the Netherlands. Diagnosing patients with shoulder pain is complex and there is debate around the diagnostic labels. Chapter 1 is the general introduction describing the management of shoulder pain in primary care from previous studies. Current trends show an increase of diagnostic ultrasound by physiotherapists, aiming to provide the physiotherapist with a better patho-anatomical diagnosis. However, the pathologies found with diagnostic imaging does not always necessarily explain the patients symptoms. It may lead to asymptomatic findings or unnecessary referrals. On the other hand it may guide for a more targeted treatment plan for the physiotherapist and thus a better prognosis. At the moment the prognosis for patients with shoulder pain is not very optimistic. Physiotherapists should consider prognostic factors that may aid the clinical decision-making process. This study aimed to describe current physiotherapy management and evaluate several prognostic factors that could improve recovery.

Chapter 2 presents the protocol of the prospective ShoCoDiP cohort study, including an interrater reliability study, in primary care physiotherapy for patient with shoulder pain. The observational study was primarily designed to evaluate physiotherapy care and study the prognostic factors in patients with shoulder pain. The working alliance and the use of diagnostic ultrasound were considered possible prognostic factors, besides the described prognostic factors from literature.

Chapter 3 describes the current diagnostic and therapeutic management from physiotherapists for patients with shoulder pain. Physiotherapists reported 1) hypotheses after patient history and physical examination, 2) the use of specific tests, 3) whether diagnostic ultrasound was used and the pathologies found and 4) the treatment plan based on the final clinical diagnosis. Patients with shoulder pain in physiotherapy practice frequently show signs of subacromial impingement/pain syndrome. Diagnostic ultrasound was used in 31% and of these patients the clinical diagnosis changed in 29%. The interventions used by the physiotherapists were generally in line with the evidence statement for subacromial impingement/pain syndrome however a small proportion of physiotherapists used massage and tape/bracing techniques. A large proportion of patients were still receiving treatment after 12 weeks even when no improvement was observed.

Chapter 4 presents the effects of a systematic review of routine diagnostic imaging for patients with musculoskeletal disorders. This review explores whether diagnostic imaging leads to better patient-reported outcomes in individuals with musculoskeletal disorders. Eleven trials were found including only patients with knee pain and low back pain. No studies including patients with shoulder pain were found. Overall, there was a moderate level of evidence for no benefit of diagnostic imaging on all outcomes compared with

controls. A significant but clinically irrelevant effect was found in favor of no (routine) imaging in low back pain patients in terms of pain severity at short [SMD 0.17 (0.04-0.31)] and long-term follow-up [SMD 0.13 (0.02-0.24)], and for overall improvement [RR 1.15 (1.03-1.28)].

Chapter 5 reports the influence of using diagnostic ultrasound on the clinical diagnosis, treatment modalities and recovery. Patients that received an ultrasound were more frequently treated with exercise therapy and patients without an ultrasound with massage therapy, triggerpoint therapy or mobilization techniques. More patients reported being recovered in the group that did not receive a diagnostic ultrasound. However, logistic regression analysis did not find a significant association between diagnostic US and recovery after 26 weeks.

Chapter 6 describes the development of a prognostic model for patient with shoulder pain. Potential predictors were selected from the literature together with two new variables (the use of diagnostic ultrasound and working alliance) and were evaluated in multivariable regression analysis. Missing data was handled with multiple imputation and the prognostic model was bootstrapped for internal validation. Short duration of complaints, lower disability scores, having a paid job, better working alliance and no feelings of depression/anxiety were factors associated with recovery. Only duration of complaints and disability were associated with recovery in the working population. The area under the receiver operator curve (AUC) was 0.67 for the final model of the total population and 0.63 for the working population. After internal validation the AUC was slightly lower.

Chapter 7 presents the validity and adjustments for the Flemish Working Alliance Inventory. A total of 274 patients filled in one or more items of the werk alliantie vragenlijst (WAV-12). A RASCH analysis showed good discriminative abilities of the items and that they all contributed to a one-dimensional construct. Although results from the analysis were good we believed rewording was necessary due to the selective nature of missing items. A Delphi study including researchers, patients and physiotherapists was performed to revise the questionnaire.

Chapter 8 discusses the results and implications of this thesis for the current management against guidelines, the use of diagnostic ultrasound and prognostic factors that can help guide patients recovery and/or treatment.

SAMENVATTING

In Nederland bezoeken patiënten met schouderpijn veelvuldig de fysiotherapeut. Het diagnostisch proces bij deze patiëntengroep is complex en er is veel discussie over de diagnostische labels. **Hoofdstuk 1** is de algemene introductie waarin het fysiotherapeutisch handelen bij patiënten met schouderpijn uit eerdere studies wordt beschreven. Recente trends laten een stijging in het gebruik van echografie zien door fysiotherapeuten, om zo tot een patho-anatomische diagnose te komen. De pathologieën die op deze beeldvormende technieken worden gevonden verklaren echter niet altijd de symptomen van de patiënt. Het kan leiden tot asymptomatische bevindingen of onnodige verwijzingen. Daarentegen zou het wel tot een gerichter behandelplan en specifiekere prognose voor de fysiotherapeut kunnen leiden. Op het moment is de prognose voor patiënten met schouderpijn niet optimistisch. Daarbij zouden fysiotherapeuten rekening moeten houden met prognostische factoren die kunnen helpen bij de besluitvorming. Deze studie beoogde het fysiotherapeutisch handelen in kaart te brengen en meerdere prognostische factoren die het herstel beïnvloeden te vinden.

Hoofdstuk 2 presenteert het protocol van de proscpectieve cohort studie (ShoCoDiP), met een interbeoordelaarsbetrouwbaarheidsstudie, in de eerste lijn fysiotherapie voor patiënten met schouderpijn. De observationele studie was primair opgezet om het fysiotherapeutisch handelen in kaart te brengen en te onderzoeken wat prognostische factoren waren voor herstel. Naast bekende factoren werden daarbij de werkalliantie en het gebruik van echografie ook als mogelijke prognostische factoren onderzocht.

Hoofstuk 3 beschrijft het huidige fysiotherapeutisch handelen bij patiënten met schouderpijn. Fysiotherapeuten rapporteerden 1) de hypothese na anamnese en lichamelijk onderzoek, 2) het gebruik van specifieke lichamelijke testen, 3) of echografie was gebruikt en de gevonden pathologie(ën) en 4) het behandelplan op basis van de klinische diagnose. Patiënten met schouderklachten in de fysiotherapiepraktijk vertoonden in de meeste gevallen een subacromiaal inklemmings-/pijn syndroom. Echografie werd gebruikt bij 31% van de gevallen en bij 29% van de patiënten veranderde de klinische diagnose door de echo. De gekozen interventies kwamen overeen met de voorschriften uit het evidence statement. Bij een kleine groep patiënten werd echter nog massage en tape/brace technieken gebruikt. Veel patiënten waren, ondanks de afwezigheid van herstel, nog steeds onder behandeling na 12 weken.

Hoofdstuk 4 presenteert het effect van een systematische literatuurstudie over het gebruik van routinematige beeldvormende technieken op het herstel bij patiënten met musculoskeletale klachten. Deze literatuurstudie verkent of het gebruik van beeldvormende technieken, als diagnostisch instrument, tot betere patiënt-gerelateerde uitkomsten leidt. Elf studies waren gevonden die alleen patiënten met kniepijn en lage rugklachten includeerden. Geen studies includeerden patiënten met schouder pijn. Alles tezamen

was er geen additioneel effect van beeldvormende diagnostiek bij alle uitkomstmaten in vergelijking met controlegroepen met een matige bewijskracht. Een significant, maar klinisch irrelevant effect was gevonden in het voordeel van de controlegroepen bij patienten met lage rugklachten op de korte termijn [SMD 0.17 (0.04-0.31)] en lange termijn [SMA 0.13 (0.02-0.24)] voor de uitkomst pijn en voor herstel [RR1.15(1.03-1.28)].

Hoofstuk 5 rapporteert de invloed van het gebruik van echografie op de diagnose, gekozen interventies en herstel van de patiënt. Patiënten die een echo hadden ontvangen werden vaker met oefentherapie behandeld en patiënten die geen echo hadden ontvangen met massage, triggerpoint therapie of mobilisatie technieken. In de groep mensen die geen echo hadden ontvangen was een grotere proportie hersteld in vergelijking met de groep die wel een echo hadden ontvangen. Een logistische regressie analyse kon echter geen statisch significant verschil hierin aantonen.

Hoofdstuk 6 beschrijft de ontwikkeling van een prognostisch model voor patiënten met schouderpijn. Potentiele prognostische factoren waren uit de biomedische literatuur geselecteerd en met behulp van multivariabele regressie analyse geanalyseerd. De werk alliantie en het gebruik van echografie werden daarbij als nieuwe prognostische factoren meegenomen. Imputatie technieken werden gebruikt voor de missende gegevens en een bootstrap werd toegepast voor een interne validatie van het model. Kort durende klachten, lagere beperkingsscores, het hebben van een baan, een betere werk alliantie en de afwezigheid van gevoelens van depressie of angst waren factoren die geassocieerd waren met een beter herstel na 6 maanden. Bij de werkende populatie waren de duur van de klacht en de beperkingsscore geassocieerd met herstel. De oppervlakte onder de curve (area under the curve) was 0.67 voor het uiteindelijke model van de totale populatie en 0.63 voor de werkende populatie. Na de interne validatie daalde de oppervlakte onder de curve maar met 0.01 punt.

Hoofstuk 7 presenteert de validiteit en aanpassingen voor de Vlaamse werk alliantie vragenlijst (WAV-12). In totaal hadden 274 patiënten een of meerdere items van deze de WAV-12 ingevuld. Een RASCH analyse toonde goede discriminitieve mogelijkheden aan van alle items en alle 12 droegen ze bij aan een uni-dimensionaal construct. Hoewel de resultaten van de analyse goed waren, geloofden we dat er aanpassingen nodig waren bij de verwoording van een aantal vragen door wat leek op selectieve missende gegevens bij sommige items. Een Delphi studie was gedaan voor de herformulering van de vragen van de WAV-12. Het panel bestond uit onderzoekers, patiënten en fysiotherapeuten.

Hoofdstuk 8 discussieert de resultaten en implicaties van deze studie voor het huidige fysiotherapeutisch handelen met betrekking tot de aanbevelingen uit richtlijnen, het gebruik van echografie en prognostische factoren die leidend kunnen zijn bij het herstel of behandelkeuzes.

DANKWOORD

Promoveren….Die roze olifant op je schouders.

Van zwoegen, tobben, wroeten, pluizen, bijeengaren, aanreiken, offeren, voordragen en ruisen; tot het proefschrift dat voor u ligt. Vijf jaren van onder andere utopisch denken, doordringende dataverzameling, complexe statistische analyses, robuuste harde schijven, diepgaande discussies en ontelbare manuscriptversies. Het moment is aangebroken en hiervoor wil ik een aantal mensen bedanken.

Allereerst mijn promotor Bart Koes en een van de grondleggers van het ShoCoDiP onderzoek. Er was geen manuscript dat jij niet hebt gelezen en van waardevolle feedback hebt voorzien. Bedankt voor de waardevolle begeleiding tijdens mijn promotietraject.

De copromotoren en bedenkers van dit onderzoek waarbij ik altijd terecht kon met vragen. Ik voelde me altijd op mijn gemak bij jullie en er was geen vraag die ik niet aan jullie kon stellen. Wendy, bedankt voor je energieke begeleiding. Mijn ontzag voor je harde werken en de manier waarop jij alle werkzaamheden naast een druk moederbestaan kunt realiseren.

Arianne. Ik heb ontzettend veel geluk gehad om jou als mijn directe begeleider te hebben. Ik bewonder je messcherpe blik en passie voor de wetenschap. Daarnaast nam je niet alleen de inhoudelijke begeleiding op je schouders, maar was je zo toegewijd dat we ook nog eens wat gezellige andere activiteiten konden ondernemen. Nooit was er een dag dat ik geen zin had om naar Avans te komen. Het was altijd gezellig op de werkvloer. Je was altijd bereikbaar en hopelijk zal je dit altijd blijven. Zelfs in Australië! Bedankt voor de kansen die jij me hebt gegeven.

Mede onderzoekers Edwin en Marloes. We mogen ontzettend trots zijn met de gegevens die we verzameld hebben. Zonder dit team waren we nooit tot deze mooie resultaten gekomen. Marloes, samen hebben wij ontzettend hard gewerkt en vooral veel plezier gehad bij het uitvoeren van ons onderzoek. Van het organiseren van een symposium tot het bellen van fysiotherapeuten en patiënten om maar ervoor te zorgen dat we genoeg gegevens konden verzamelen. Wat een heerlijk moment om hier samen ook een einde aan te maken. Toch zal ik onze gezellige werkmomenten missen! Ik ben blij je een vriendin te mogen noemen na onze gezamenlijke toewijding aan dit project. Edwin, dankjewel voor alle gezellige avonturen, congressen en volle koffiebekers die wij samen hebben gehad.

Het hele ShoCoDiP team waaronder Ramon, Maaike, Geert-Jan, Eric, Annechien, Joost en Marcel. Bedankt voor jullie waardevolle input in het project. Tevens alle panelleden van de Delphi studie.

Collega's van Avans. Bedankt voor jullie begrip voor mijn promotietraject. Met name Inge waarmee ik veel samenwerk. Jij begreep wat het inhoudt om een promotie traject af te ronden. Een bijzonder bedankje aan de directie, het college van bestuur en het Lectoraat Leven Lang in Beweging voor de financiële middelen en de mentale ondersteuning voor deze promotie.

Dank aan alle fysiotherapeuten en radiologen in de regio Noord-Brabant, Zeeland en Zuid-Holland die hebben meegeholpen met het verzamelen van gegevens. Ook alle co-auteurs ben ik uiteraard dankbaar voor de bijdrage aan de artikelen van dit proefschrift.

Mijn vrouw Luisa die er altijd voor me was. De moeilijke beslissing om naar Nederland te komen. Bedankt voor alle leuke momenten en lachbuien die we hadden. Jij steunde mij door dik en dun. Je hebt me van mijn goede en minder mooie kanten gezien. Er was maar één plek waar ik alles kon vergeten en dat was in je armen.

Nooit had ik gedacht zoveel van een hond te kunnen houden. Manuela je bent de beste viervoeter in de wereld!

Zonder vrienden zou ik niet zoveel plezier hebben in het leven.

Yvonique, altijd een lach op je gezicht. Je bent een ontzettend lieve en excentrieke vrouw. Ik zal je nooit als vanzelfsprekend beschouwen, maar altijd waarderen! Dankjewel dat je mijn vriendin bent.

Vriendin Ineke. Wij delen veel van onze herinneringen. Misschien kunnen we ook nu weer naar huis rijden en een paar extra rondjes om de rotonde rijden!? Ik houdt van de diversiteit van onze vriendschap. Met jou doe ik dingen die ik met niemand anders doe. Tegenwoordig zien we elkaar niet zo veel, maar elke ontmoeting voelt weer als oudsher. Ook bij dit belangrijke moment in mijn leven ben jij een belangrijk onderdeel geweest. De kleuren van de kaft zijn de kleuren die jij aan mijn leven geeft. Bedankt voor de mooie kaft die jij hebt gemaakt! Ellen, het begon allemaal met die befaamde avocado in de bus. Nu zijn we collega's op Avans, collega's bij Enjoy en ben je paranimf tijdens mijn promotie, maar eigenlijk zijn wij vooral vriendinnen! Ik apprecieer je nuchtere commentaar en optimistische levenshouding. Zonder jou was het werk op Avans een stuk ongezelliger!

Katrien en Emmanuel, Ik ben blij dat ik jullie heb leren kennen en dankbaar dat jij, Katrien, mijn paranimf wil zijn.

Alyona, Frank, Lei, Baukje, Dana, Nikki, Robert, Fenna, Niels & Maranna, Marjolein, Mariëtte, Constance en Wolfgang bedankt voor al jullie gezelligheid en blijdschap die jullie in mijn leven brengen. Maria & Jos, jullie maakten de integratie in België ontzettend gezellig! Ook mijn dank aan de capoeira vrienden (Cyriel, Mestre Vulcao, Willemijn, Veronica,Katinka, Sophia, Diane, Roeland and Miguel) van Eindhoven en de spinning-groep van de zondagochtend voor de wekelijkse ontspanning.

To all my friends from the European Society for Shoulder and Elbow Rehabilitation. Thank you, Suzanne, Anju, Beate, Marco, Edwin, Andre for all the great discussions regarding shoulder pain! Ingrid, thank you for welcoming me to your beautiful home and for becoming one of my dearest friends. I admire how skilled you are and your ability to motivate people in the society.

Zonder mijn liefdevolle familie was ik nu niet waar ik ben. De rust en stilte van oma Zus was een plek waar ik me kon verstoppen. De lach van oma Laura die me meerdere malen tot een melige lach heeft doen uitbarsten. De overtuigingskracht die ik van opa Jacques heb geleerd en de wijsheid van opa Phillip mis ik nog elke dag. Alle gekke bedoelingen van Maudy, Paul, Thécla, Peter, Daan, Judith en Helga. Jullie zijn de steunpilaar van onze familie. Maudy, mijn generatie in de familie mag blij zijn met een tante zoals jij! Je accepteert iedereen zoals ze zijn. Ik denk vaak terug aan de momenten dat ik als kind samen met jou in het verpleeghuis activiteiten mocht doen met de ouderen. Wellicht is dit oorsprong geweest voor mijn keuze voor een beroep in de zorg. Paul, ik hoop zo actief en positief te mogen blijven als jij! Ik bewonder je doorzettingsvermogen. Thécla, het is ontzettend moeilijk om te beseffen dat jij niet meer aan de eettafel met een koffie zit. Je was altijd ontzettend geïnteresseerd in mijn werkzaamheden. Ik ga je lach en rust nooit vergeten. Peter, jij leert me dat je alles zelf kan bouwen als je het maar slim en simpel aanpakt. Daan, ik bewonder je blik op het leven en de jazz progressies die jij zo uit je mouw schudt. Judith, dankjewel voor al je gezelligheid. Helga, jij hebt me geleerd dat veerkracht een van de belangrijkste intrinsieke vaardigheden van een persoon is. Lothar en Irma, zonder jullie had ik nooit met zo'n ontspannen gevoel mijn masterstudie kunnen doen. Bedankt dat ik in deze tijd in jullie prachtige appartement in Maastricht mocht verblijven. Alle lieve neefjes en nichtjes, ik ben trots jullie nichtje te mogen zijn!

Thank you to my beautiful family in New York; Rosa, Brighan, Fabi and all the lovely family from Colombia to teach me that nothing is what it seems.

De grootse dank aan mijn ouders, Reginald en Jacqueline, voor de onvoorwaardelijke liefde die ik mag ontvangen en dat jullie altijd voor me klaar stonden. Ik maakte vaak gebruik van de mogelijkheid om bij jullie even te ontsnappen aan de dagelijkse sleur. Ik heb geleerd dat een warm thuis heel waardevol is. Dat het gebrek aan geldzorgen fijn

is, maar dat geld niet gelukkig maakt. Hopelijk zijn jullie trots, want ik ben trots op jullie. Op de creativiteit van mijn moeder en de stabiliteit van mijn vader. Ik had geen betere ouders kunnen wensen.

ABOUT THE AUTHOR

Yasmaine Karel was born on May 14th 1988 in Eijsden, Limburg, the Netherlands. After completing secondary school in 2004, she moved to Breda to study Physiotherapy at the Avans University of Applied Sciences. In her fourth year she took the opportunity to combine her bachelor with a premaster at the University of Maastricht. In 2009 she obtained her bachelor degree in Physiotherapy. She started working as a physiotherapist and at the same time studying for a master's degree. In 2010 she got a master's degree in Public Health Science.

After obtaining her master's degree she kept working as a primary care physiotherapist and in 2012 was offered a position as a teacher and researcher at the Avans University of Applied Sciences. For two years she collected data for the ShoCoDiP study (Shoulder Complaints and Diagnostic ultrasound in Physiotherapy), while working as a physiotherapist for two years as well. When in 2014 a promotional voucher was acquired, she fully committed to doing a PhD project about shoulder complaints in primary care physiotherapy. Since 2014 she has been an active member of the European Society for Shoulder and Elbow Rehabilitation (EUSSER) and has been chairing the Scientific Committee since 2015. Currently, she is a researcher in public health for the Research group "life in motion" and is a teacher and coordinator for the graduation program in the bachelor of Physiotherapy in Breda.

PHD PORTFOLIO

Name PhD student: drs. Yasmaine Helga Jacques Marie Karel	Promotor : Prof. Dr. Bart Koes
Erasmus MC Department: Department of Family Medicine	Supervisors: Arianne P Verhagen
PhD period: 2013-2018	Wendy Scholten-Peeters

1. Phd Training		Workload	
		Hours	ECTs
Courses			
Medipoint Hands-on training Musculoskeletal Ultrasound	2012	16	1
SECCEC-Shoulder and elbow unit - Advance Shoulder Course	2013	12	1
Reading advanced shoulder course	2013	16	1
Seminars and Workshops			
Athens seminar: scapular dyskinesis – Athens GR (workshop)	2017	12	1
Int course upper extremity surgery – Poznan PL (workshop)	2018	12	1
Int. course upper extremity surgery – Poznan PL (mini-battle)	2018	12	1
Presentations			
KNGF congres (poster presentation)	2012	8	1
Schoudercongres Flexibility Matters (poster presentation)	2017	8	1
(International) Conferences			
Nursing and Healthcare – Texas USA (oral presentation)	2016	55	1
Nice Shoulder course rehab – Nice FR (2 oral presentations)	2016	55	2
NHG Wetenschapsdag – Amersfoort NL (oral presentation)	2017	30	1
SECEC-ESSSE congress – Berlin GER (organizer/moderator)	2017	75	1
Int course upper extremity surgery – Poznan PL (oral presentation + moderator)	2018	50	1
Other			
Writing Factsheet KNGF	2018	20	1
Chair of Scientific Committee EUSSER	2015-2018	150	3

2. Teaching	Year		
		Hours	ECTs
Supervising master's thesis (2 students)	2016-2017	12	10
Supervising Bachelor's thesis (35 students)	2015-2018	280	10
Teaching Bachelor physiotherapy programme (3 days per week)	2012-2018	>1000	10
Total			

LIST OF PUBLICATIONS

Karel, Y., Thoomes-de Graaf, M., Scholten-Peeters, G., Ferreira, P. H., Rizopoulos, D., Koes, B. W., & Verhagen, A. P. (2018). Validity of the Flemish working alliance inventory in a Dutch physiotherapy setting in patients with shoulder pain. *Physiotherapy, Theory and Practice*, 1-9.

Thoomes-deGraaf, M., Scholten-Peeters, W., Karel, Y., Verwoerd, A., Koes, B., Verhagen, A. (2018). One question might be capable of replacing the Shoulder Pain and Disability Index (SPADI) when measuring disability: a prospective cohort study. *Qual Life Res,* 27(2):401-410

Karel, Y.H.J.M., Scholten-Peeters, G.G.M., Thoomes-de Graaf, M., Duijn, E., van Broekhoven, J., Koes, B., Verhagen, A.P. (2017). Physiotherapy for patients with shoulder pain in primary care: a descriptive study of diagnostic- and therapeutic management. *Physiotherapy*, 103(4):369-378

Thoomes-de Graaf, M., Scholten-Peeters, W., Duijn, E., Karel, Y., de Vet, H.C., Koes,B., Verhagen, A. (2017). The responsiveness and Interpretability of the Shoulder Pain and Disability Index. *J Orthop Sports Phys Ther*, 47(4):278-286

Karel, Y.H.J.M., Verhagen, A.P., Thoomes-de Graaf, M., Duijn, E., van den Borne, M.P.J., Beumer, A., Ottenheijm, R.P.G., Dinant, G.J., Koes, B.W., Scholten-Peeters, G.G.M. (2017). Development of a Prognostic model for Patients With Shoulder Complaints in Physical Therapist Practice. *Phys Ther*, 97(1):72-80

Karel, Y.H.J.M., Verkerk, K., Endenburg, S., Metselaar, S., Verhagen, A.P. (2015). Effect of routine diagnostic imaging for patients with musculoskeletal disorders: A meta-analysis. *Eur J Intern Med*, 26(8):585-95

Thoomes-de Graaf, M., Scholten-Peeters, G.G.M., Duijn, E., Karel, Y.H.J.M, Koes, B.W., Verhagen, A.P. (2015). The Dutch Shoulder Pain and Disability Index (SPADI): a reliability and validation study. *Qual Life Res*, 24(6): 1515-9

Thoomes-de Graaf, M., Scholten-Peeters, G.G.M., Duijn, E., Karel, Y.H.J.M., van den Borne, M.P., Beumer, A., Ottenheijm, R.P., Dinant, G.J., Tetteroo, E., Lucas, C., Koes, B.W., Verhagen, A.P. (2014). Interprofessional agreement of ultrasound-based diagnoses in patients with shoulder pain between physical therapists and radiologists in the Netherlands. *Man Ther*, 19(5):478-83

Karel, Y.H.J.M., Scholten-Peeters, G.G.M., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R.P., van den Borne, M.P., Koes, B.W., Verhagen, A.P.; ShoCoDiP (Shoulder Complaints and using Diagnostic ultrasound in Physiotherapy practice) study group, Dinant, G.J., Tetteroo, E., Beumer, A., van Broekhoven, J.B., Heijmans, M. (2013). Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. BMC Musculoskelet Disord, 14:62