Expert therapists use specific clinical reasoning processes in the assessment and management of patients with shoulder pain: a qualitative study

Stephen May¹, Alison Greasley², Sarah Reeve² and Sarah Withers³

¹Sheffield Hallam University, ²Sheffield Teaching Hospitals Foundation NHS Trust, ³Sheffield Primary Care Trust

United Kingdom

Question: What are the key items in the clinical reasoning process which expert clinicians identify as being relevant to the assessment and management of patients with shoulder pain? Design: Qualitative study using a three-round Delphi procedure. Participants: Twenty-six experts in the UK consented to be involved and were contactable, of whom 20 contributed, with 12, 15, and 15 contributing to the different rounds. Results: Clinical reasoning was mostly about diagnostic reasoning, but also involved narrative reasoning. Diagnostic reasoning involved both pattern recognition and hypothetico-deductive reasoning. Diagnostic reasoning emphasised general history items, a constellation of signs and symptoms to identify specific diagnostic categories, and standard physical examination procedures. Narrative reasoning was highlighted by the communication involved in expert history taking, seeing patients in their functional and psychological context, and collaborative reasoning with the patient regarding management. Conclusions: These expert clinicians demonstrated the use of diagnostic pattern recognition, and hypothetico-deductive and narrative clinical reasoning processes. The emphasis was on the history and basic physical examination procedures to make clinical decisions. [May S, Greasley A, Reeve S, Withers S (2008) Expert therapists use specific clinical reasoning processes in the assessment and management of patients with shoulder pain: a qualitative study. Australian Journal of Physiotherapy 54: 261–266]

Key words: Shoulder pain, Decision making, Problem solving, Expert opinion, Delphi technique, Physiotherapy (modality)

Introduction

Patients with shoulder pain are commonly encountered in healthcare settings (van der Windt et al 1995, May 2003) and a number of diagnostic labels have traditionally been used, such as capsulitis, bursitis, and subacromial impingement syndrome (Cyriax 1982). However, the pathophysiology underlying shoulder disorders is still controversial (Lewis et al 2001, Tytherleigh-Strong et al 2001, Carette 2000, Khan et al 2000, Chard et al 1994), tests used to establish a diagnosis have limited reliability (Liesdek et al 1997, de Winter et al 1999, Hanchard et al 2004), and the diagnostic validity of most of the tests is moderate at best (Hegedus et al 2008, Hughes et al 2008, Dessaur and Margarey 2008, Calis et al 2000).

A number of synonyms or overlapping terms are used for similar pathologies without it being clear whether these are the same clinical entities. For instance, subacromial impingement syndrome, impingement syndrome, rotator cuff pathology, supraspinatus tendonitis, and supraspinatus tendinosis are all terms that might be used for similar or the same pathologies, but the implications of each title may vary. Impingement implies some factor external to the tendon, the suffix ‘itis’ implies inflammation of the tendon, and the suffix ‘osis’ implies dysfunction rather than inflammation. These pathologies might display different features on physical examination, warrant different interventions, and have different prognoses. Most of this nomenclature is developed from a pathophysiological model, but even in a relatively recent clinical guideline an operational definition for impingement syndrome was not found (Hanchard et al 2004). Furthermore, not everybody with shoulder pain can be classified with a specific pathoanatomical diagnosis (van der Windt et al 1995). Non-specific classification based on symptom response to repeated movement testing has been proposed (McKenzie and May 2000). The potential value of this approach has been supported by case studies, which provide clinically-based operational definitions (Aina and May 2005, Littlewood and May 2006).

Thus there appear to be a number of problems in linking shoulder pathology with its classification and management. These include the use of a range of diagnostic terms for the same or similar anatomical problems, a lack of standardised operational definitions for diagnostic labels (Schellingerhout et al 2008), the limited validity of most physical examination tests of the shoulder (Hegedus et al 2008, Hughes et al 2008, Dessaur and Margarey 2008), the uncertain reliability of many tests (Ostor et al 2004, Hickey et al 2007) and diagnostic labels (de Winter et al 1999, Walker-Bone et al 2002), and the lack of a link between diagnostic labels and specific interventions (Schellingerhout et al 2008).

Clinical reasoning is the decision-making process used to determine the diagnosis and management of patients’ problems (Terry and Higgs 1993, Jones et al 1994); it has been defined as the ‘thinking and decision making associated
with clinical practice that enables therapists to take the best-judged action for individual patients’ (Jones and Rivett 2004, p. 3). There are several models of clinical reasoning including pattern recognition, hypothetico-deductive or diagnostic reasoning, and narrative reasoning (Jones 1992, Terry and Higgs 1993, Jones and Rivett 2004, Edwards and Jones 2007). In pattern recognition, the clinician associates problems of the current patient with previously-seen clinical problems and adopts a previously-successful management strategy. In hypothetico-deductive reasoning, the clinician generates a hypothesis based on data from the patient, which is then tested, and further hypotheses are generated until a management pathway is defined clearly. Narrative reasoning is a ‘process of enquiry, examination and reflective management’ (Jones and Rivett 2004, p. 5) by which the clinician understands the patient’s problem, the patient’s perspective, and the context of that problem. It demands collaborative reasoning between the patient and the clinician, effective communication by the clinician, and on-going reasoning until a plan of management is agreed upon. Diagnostic and narrative reasoning are said to be fundamentally different forms of reasoning, but intrinsically linked in a dialectical model (Edwards and Jones 2007).

Clinical reasoning by health professionals has generally been examined using interpretive paradigm studies, which seek to generate practical knowledge through the description and interpretation of phenomena explored as a whole, in context, and to include the meanings and significant aspects of the situation from the perspective of the people being studied. This has been done in nursing (Benner 1984), in occupational therapy (Fleming 1991), and in physiotherapy (Jensen et al 1992) although clinical reasoning strategies within physiotherapy remain largely under-researched (Edwards et al 2004a). In physiotherapy, Jensen et al (1992) examined the complex processes that occur during therapeutic interaction and attempted to examine the differences between novice and expert physiotherapists in order to define the characteristics of the latter (Jensen et al 1990, 1992, 2000). In an exploration of management approaches, when ‘expert’ was defined by outcome rather than by years of clinical experience, experts were deemed to have a patient-centred approach to care characterised by collaborative reasoning and encouragement of patient empowerment (Resnik and Jensen 2003). Most of the work on clinical reasoning has used an observational study design, and has been of a theoretical nature rather than exploring specific problems.

This study explored the clinical reasoning of experts related to a complex musculoskeletal problem using a series of questionnaires known as a Delphi process, which is useful where clinical judgements need to be made but empirical evidence on the topic is limited (Powell 2003). It has been used to explore musculoskeletal topics, namely the classification of low back pain (Binkley et al 1993), and signs and symptoms of spinal clinical instability (Cook et al 2005). The aim of this study was to explore the clinical reasoning process of expert occupational and physiotherapists in assessing and managing patients with shoulder pain. The research question was:

What are the key items in the clinical reasoning process which expert clinicians identify as being relevant to the assessment and management of patients with shoulder pain?

### Method

#### Design

We conducted a Delphi survey that allowed respondents to propose and then refine and weight items until consensus was achieved (Powell 2003, Rowe and Wright 1999). Round 1 was unstructured and allowed multiple open responses. Clinicians were asked to share the clinical reasoning process by which they assessed patients with shoulder pain and arrived at a management strategy. The responses were consolidated using qualitative thematic analysis, in which items were grouped together through shared themes relating to the narrative reasoning process of shoulder assessment, classification, and management. Similar words or phrases were reduced into a single item. In Round 2, respondents were asked to rank the relative importance of all the items using the scale:

1 = Essential; this item would be used in all shoulder examinations
2 = Very important; this item would be used in most shoulder examinations
3 = Important; this item would be used in a lot of shoulder examinations
4 = Less important; this item would be used in a few shoulder examinations
5 = Unimportant; this item would be used rarely in shoulder examinations.

The items were ranked according to the responses to form a new consolidated list with items weighted by their relative importance, the percentage agreement for that rank, and the composite score. In Round 3, respondents voted on their agreement or disagreement with the ranking that emerged from Round 2. Both the ranking of the item and the percentage consensus for that ranking was available to the respondents. The final list of items was generated from the responses. All correspondence was via e-mail; two reminders were used at each round to maximise the response rate.

#### Participants

Participants were recruited by telephone from Extended Scope or Clinical Specialist Physical Therapy or Occupational Therapy practitioners who specialised in shoulder or upper limb problems (as defined by their job title) in the UK. The sample was located through personal contacts, delegate lists from conferences, ICSP (an interactive website run by the Chartered Society of Physiotherapy), and through the Extended Scope Practitioners Clinical Interest Group. Demographic details and the clinical characteristics of their patient population were collected.

#### Data analysis

After Round 2, items ranked as essential or very important by ≥ 75% of respondents were ranked as primary items. Items ranked as important by ≥ 75% of respondents were ranked as secondary items. Items ranked as less important or unimportant by ≥ 75% of respondents were classified as tertiary items. Previous studies have used ≥ 75% as indicating consensus (Binkley et al 1993, Cook et al 2005, 2006, McCarthy et al 2006). When this measure of consensus was not reached in this study, but there was a majority decision (50% to 74%), this majority decision was used to rank items as above. Items where a majority decision

was not reached (< 50%) were discarded from the study, but retained in the final list for information. Therefore, in Round 2 the respondents produced a list of ranked items where there was a majority decision about the importance of their inclusion in a shoulder assessment.

Composite scores to determine a numerical ranking for the different items were determined with the following formula (Cook et al 2005):

\[ \text{Composite score} = (n_1 \times 5) + (n_2 \times 4) + (n_3 \times 3) + (n_4 \times 2) + (n_5 \times 1). \]

That is where the number of respondents voting an item as: Essential (1) is multiplied by 5, Very important (2) is multiplied by 4, Important (3) is multiplied by 3, Less important (4) is multiplied by 2, and Unimportant (5) is multiplied by 1; and then each separate total is added together to gain the composite score.

In Round 3 of the Delphi study, the respondents confirmed or denied the primary, secondary or tertiary ranking of the item in the clinical reasoning process, and the authors produced a percentage composite score that represented the level of consensus about that ranking from the respondents’ responses.

**Results**

**Flow of participants through the study**

Thirty-six upper limb and shoulder experts were located of whom 30 consented to participate in the study, but four could not be contacted at the email addresses provided. Of the 26 who were contactable, 20 contributed to at least one round of the Delphi process. Twelve contributed to Round 1, 15 to Round 2, and 15 to Round 3, with 8 contributing to all rounds (Figure 1). Demographic details and the clinical characteristics of the patient population are provided in Table 1.

**Key items relevant to the assessment of patients with shoulder pain**

After Round 1, 887 separate items were consolidated into 238 items comprising seven themes:

1. General history
2. Diagnostic classifications
3. Signs and symptoms linked to specific classifications
4. General physical examination
5. Management decisions
6. Investigations
7. Intervention options.

Themes were established by the authors on a consensual basis with complete agreement about which items belonged with which themes. There was limited change by the respondents in the ranking of items between Rounds 2 and 3. Composite scores from Round 2 are not reported as they correlated closely with rank and percentage agreement.

After Round 3, items fell into one of four categories based on participants’ responses: majority agreement about primary importance, majority agreement about secondary importance, majority agreement about tertiary importance, or a wide range of responses so lack of consensus. Specific rankings and percentage agreement with rank are presented in Tables 2 to 5 (see eAddenda for Tables 2 to 5). The overall rankings reported here are for groups of items rather than for individual items.

**Table 1. Characteristics of respondents.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n = 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr), mean (SD)</td>
<td>43 (5.3)</td>
</tr>
<tr>
<td>Time since qualifying (yr), mean (SD)</td>
<td>20 (2.2)</td>
</tr>
<tr>
<td>Years in present post, mean (SD)</td>
<td>7 (1.9)</td>
</tr>
<tr>
<td>Health care sector, n (%)</td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>2 (14)</td>
</tr>
<tr>
<td>Secondary care</td>
<td>12 (86)</td>
</tr>
<tr>
<td>Proportion of shoulder patients seen (%), mean (SD)</td>
<td></td>
</tr>
<tr>
<td>Impingement/rotator cuff disease (n = 12)</td>
<td>44 (3.6)</td>
</tr>
<tr>
<td>Atraumatic instability (n = 12)</td>
<td>28 (4.1)</td>
</tr>
<tr>
<td>Adhesive capsulitis (n = 9)</td>
<td>12 (2.4)</td>
</tr>
<tr>
<td>Osteoarthritis (n = 5)</td>
<td>11 (2.3)</td>
</tr>
</tbody>
</table>

Figure 1. Flow of participants through the study.
Majority agreement about primary importance occurred with the following issues:
- General history items
- Two most frequently considered diagnostic categories (rotator cuff disease, neck pain)
- Use of a constellation of signs and symptoms in clinical recognition of diagnostic categories (frozen shoulder, instability, impingement, rotator cuff problems, glenohumeral osteoarthritis, and acromioclavicular joint problems)
- General physical examination items, mostly of a ‘non-specialist’ nature (such as active and passive range of movement and resisted tests)
- Management decisions based on a range of considerations (such as stage of disorder, diagnosis, and response to intervention)
- Use of investigations
- Certain intervention options (mostly advice and exercise based).

Majority agreement about secondary importance occurred with the following items:
- Next most frequently considered diagnostic categories (frozen shoulder, acromioclavicular joint)
- Some specific physical examination tests.

Majority agreement about tertiary importance occurred with the following items:
- Specific diagnostic categories (non-specific shoulder pain, avascular necrosis, long-head of biceps dysfunction, Ehlers Danlos syndrome, Marfan’s disease, serious pathology, nerve palsy, Pancoast tumour, visceral disease)
- Some specific physical examination tests
- Therapist-conducted intervention options (such as TENS, acupuncture or upper limb tension test mobilisations).

There was a wide range of responses about the relative importance of a number of items; therefore, there was lack of consensus in these areas:
- Specific diagnostic categories (including instability and labral tears)
- Physical examination items, especially specific diagnostic tests (such as Burkhead’s, Hawkins-Kennedy, and Neer tests)
- Intervention options (such as manual therapy, injection, and taping).

Discussion

Both diagnostic and narrative clinical reasoning processes were used by this group of expert clinicians, however much weight was given to diagnostic reasoning (Jones et al 2008). Hypothetico-deductive reasoning was used in the history-taking to inform the use of specific tests, differentiation and management decisions; and diagnostic reasoning was underpinned by pattern recognition, which is commonly used by more experienced clinicians (Jones 1992, Terry and Higgs 1993). This was done through a constellation of signs and symptoms, was heavily dependent on the history taking, and relied not simply on diagnostic classifications, but also on patterns of impairment. This reflects the lack of standardised operational definitions for diagnostic labels (Schellingerhout et al 2008), the limited validity and reliability of most physical examination tests of the shoulder (Hegedus et al 2007, Hughes et al 2008, Dessaur and Margaret 2008, Ostor et al 2004, Hickey et al 2007), and the fact that not everyone can be given a structural diagnosis.

That impairment as well as pathology featured strongly in the reasoning process is confirmed by the importance of the standard parts of the physical examination compared to specialist orthopaedic tests. Several of the experts indeed identified responses to movement as a key feature of the diagnostic reasoning process (Edwards et al 2006). This is consistent with low back pain impairments, in which pattern recognition based on clinical features is a key feature of a number of classification systems (McKenzie and May 2003, Van Dillen et al 2003, O’Sullivan 2006, Brennan et al 2006). In a similar way management decisions were not based on recipes related to pathology but also involved hypothetico-deductive reasoning and took into account diagnostic classification, stage of the disorder, functional requirements, and response to intervention.

Where diagnostic reasoning was used by these clinicians there were also elements of narrative or patient-centred reasoning. This was highlighted in the importance of the history in the clinical reasoning process, awareness of patients’ psychosocial context, taking the functional requirements of the patients into account, the emphasis on collaborative reasoning with the patient, and the provision of advice and exercises as part of the management plan. The primacy given to history items emphasised the importance of clinician-patient communication and probably highlights the advanced skill levels of this group in their history taking and interpretation.

Diagnostic and narrative reasoning are meant to be underpinned by different research paradigms: diagnostic reasoning reflecting a positive or quantitative approach, and narrative reasoning an interpretative or qualitative approach (Jones et al 2008). Narrative reasoning is context-dependent and socially constructed and leads to what has been termed ‘communicative’ decision-making and management (Edwards et al 2004b). The use of both types of clinical reasoning by these clinicians confirms the dialectical model of the two (Edward and Jones 2007).

Some elements appear to be contradictory, especially regarding specific diagnostic tests. For instance, Neer and Hawkins–Kennedy tests were considered items of primary importance in the diagnostic constellation of signs and symptoms for establishing the diagnostic category of impingement. But the same tests were excluded from the general physical examination items as there was lack of consensus about their importance. We interpreted this as follows: a single test used in isolation has limited diagnostic accuracy, but combined with other tests and information from the history it can be useful.

A number of limitations need to be recognised with this study. Not everybody who consented to participate in the study responded, but it is impossible for us to know why this occurred. About 77% of consenting and contactable respondents contributed to data collection overall, but at each round response rates ranged from 46% to 58%. However, in the first round of data collection, the last two respondents contributed few new items to those already raised, suggesting that we had achieved saturation point of new items. In addition, we cannot know if respondents were actually able to make the diagnostic classifications from history and physical examinations items that they suggested they were capable of doing. However, the combination of an uncertain scientific background and plenty of diverse opinion makes this an excellent topic for expert opinion consensus.
The study has a number of clinical, research, and educational implications. It has emphasised movement impairments, based on the patient’s history and standard examination procedures, rather than specific structural diagnoses based on specific tests. This has both clinical and educational implications regarding diagnostic reasoning. It is in accord with a recent review, which concluded that, because of a lack of standardised operational definitions for diagnostic labels and the lack of a link between diagnostic labels and specific interventions, it is time for a different approach to the management of shoulder problems based on other than specific structural pathology (Schellingerhout et al 2008). Whether the constellation of signs and symptoms that our respondents suggested is actually of useful diagnostic accuracy should be validated by further research. Furthermore, it would be of interest to compare the clinical reasoning process of this group of experts with less experienced therapists to evaluate whether there are differences and, if so, where these lie.

These expert clinicians employed diagnostic clinical reasoning that used both a hypothetico-deductive and a pattern recognition model. There was consensus about the primacy of items from the history and standard physical examination procedures in this process, and the use of a constellation of signs and symptoms for pattern recognition. Although specific pathological diagnoses were in the diagnostic model, the process also highlighted non-specific movement impairments. And although diagnostic reasoning dominated, there were also examples of narrative and patient-centred reasoning and recognition of the context-dependent nature of the patient information.

eAddenda: Appendix 1 and Tables 2–5 available at AJP. physiotherapy.asn.au

Ethics: The Leeds (East) Research Ethics Committee approved this study (REC ref number: 07/Q1206/12). Informed consent was gained from all participants before data collection began.

Acknowledgements: The authors thank all the participating therapists who contributed their time, effort and knowledge so generously to this study. The analysis was ours, but the source of information was all theirs.

Corresponding author: Stephen May, Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK, S10 2BP. Email: s.may@shu.ac.uk

References


Research


