

No room for reflexivity? Critical reflections following a systematic review of qualitative research

Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand and Jon Howard Raphael

J Health Psychol 2012 17: 866 originally published online 6 December 2011

DOI: 10.1177/1359105311427615

The online version of this article can be found at:

<http://hpq.sagepub.com/content/17/6/866>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Journal of Health Psychology* can be found at:

Email Alerts: <http://hpq.sagepub.com/cgi/alerts>

Subscriptions: <http://hpq.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://hpq.sagepub.com/content/17/6/866.refs.html>

>> [Version of Record](#) - Aug 20, 2012

[OnlineFirst Version of Record](#) - Jan 11, 2012

[OnlineFirst Version of Record](#) - Dec 6, 2011

[What is This?](#)

No room for reflexivity? Critical reflections following a systematic review of qualitative research

Journal of Health Psychology
17(6) 866–885
© The Author(s) 2011
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1359105311427615
hpq.sagepub.com


**Benjamin John Newton¹, Zuzana Rothlingova²,
Robin Gutteridge³, Karen LeMarchand⁴ and
Jon Howard Raphael¹**

Abstract

We conducted a systematic search of qualitative research into the individual's experience of chronic low back pain. Two reviewers independently read through 740 unique abstracts. Inter-rater reliability was fair. The final sample comprised 19 articles which we critiqued using the Critical Appraisal Skills Programme checklist. This article focuses on the critical appraisal. Limitations include a lack of an adequate rationale for the theoretical framework, a lack of an account for the decisions made across recruitment and data collection, and a lack of reflexivity. Finally we discuss and offer recommendations for reflexivity and the explication of qualitative methodology in research articles.

Keywords

chronic low back pain, critical appraisal, methodology, reflexivity, systematic review

Introduction

The inception of this review initially came from curiosity. As a team of researchers and clinicians we were interested in the impact of qualitative research on clinical practice; specifically in relation to the work of the Pain Management department. We were interested in whether qualitative research made a difference in the knowledge base of health professionals. In the course of this review however, we became concerned that some aspects (representing the strengths) of qualitative research appeared to be neglected. In this article we first highlight some contemporary issues surrounding the selection

¹Centre for Health & Social Care Research, Birmingham City University, Birmingham, UK

²Trent DClinPsy Programme, University of Lincoln, Lincoln, UK

³School of Health and Wellbeing, University of Wolverhampton, Wolverhampton, UK

⁴Department of Psychology, Black Country Partnership NHS Foundation Trust, Dudley, UK

Corresponding author:

Benjamin J. Newton, Room 311, Centre for Health and Social Care Research, Ravensbury House, Birmingham City University, Westbourne Road, Birmingham, B15 3TN, UK.

Email: b.j.newton@gmx.com

and use of critical appraisal tools. We then describe our systematic approach to article selection and appraisal. We report on the results of the appraisal and subsequently discuss two of our main concerns: (i) that researchers are still explaining and justifying qualitative research; and (ii) that a core part of qualitative research, reflexivity, is not adequately evidenced in any of the articles we reviewed.

In line with other researchers (Miller, 2010; Noyes et al., 2008; Reid et al., 2009) we recognize the issue of quality appraisal in qualitative research is full of tension and ambiguity. Such debate is ongoing, though some argue for the urgent need to achieve some form of conclusion (Dixon-Woods and Fitzpatrick, 2001). There are already a number of articles whose authors offer an in-depth discussion of the key components that comprise good qualitative research. To avoid re-duplicating the literature we direct interested readers to articles such as Horsburgh (2003), Koch (2006) and, more recently, Sin (2010) for further information on quality criteria within qualitative research.

A gold-standard appraisal tool?

The evolutionary nature of the debate has not hindered the development of numerous appraisal tools. Indeed several authors point to the proliferation of qualitative tools that are available to appraise quality (Dixon-Woods et al., 2006; Noyes et al., 2008). Unsurprisingly, some researchers are now moving on from the question of *how* to appraise qualitative research, towards the consideration of *which* appraisal tool is the strongest. For instance Katrak et al. (2004) undertook a systematic review of both quantitative and qualitative appraisal tools in which they sought to identify the most appropriate items across a range of diverse (research) appraisal tools. Their initial search revealed 193 articles which potentially described an appraisal process and through refinement, they arrived at 121 tools, of which seven were qualitative. Katrak and colleagues emphasized that such tools should have an empirical base to support

both their development and application. Furthermore they recommended that researchers should purposively select the appraisal tool according to the needs of the research. The latter suggestion is perhaps recognition that there can be no 'one size fits all' approach to quality appraisal.

In a more recent study, Hannes, Lockwood and Pearson (2010) sought to assess the extent to which three qualitative appraisal tools successfully took account of the validity of qualitative research, arguing that validity is akin to methodological rigor. In selecting tools for evaluation they considered whether the tool was free of cost and available online; its applicability to different qualitative approaches; its use in published syntheses of qualitative research; and the extent of organizational support and involvement (beyond "individual academic interest" (Hannes et al., 2010: 1737)). Of the eight tools they initially selected, Hannes and colleagues reviewed the Critical Appraisal Skills Program (CASP) tool, the Evaluation Tool for Qualitative Studies (ETQS) and the Joanna Briggs Institute (JBI) tool. Of these three the researchers concluded that the CASP was the "least sensitive to validity" (p. 1741). The JBI was the most sensitive.

The issue of context

Of course, one risks over-simplifying the issue by attempting to 'arrive at' a conclusion that a particular appraisal tool is the best one to use. Such a view is open to challenge. Many qualitative researchers hold anti-realist assumptions; the data they generate within their research is contextually bound to a particular time and space. This is no less true in the use of qualitative checklists. For instance Barbour and Barbour (2003) argue that concepts of trustworthiness are not rigid but are subject to change over time. Knowledge is subject to paradigm influence. That which is considered trustworthy in one era might not stand up to scrutiny in a subsequent community of practice. Barbour and Barbour (2003) also note how the

use of qualitative appraisal tools in earlier years of qualitative research may have excluded seminal articles: an absence of techniques such as triangulation does not preclude a research article from offering valuable and groundbreaking insights.

If criteria for quality can change over time, they can also vary in the space of a single moment. Qualitative research is full of diversity. Indeed Dixon-Woods et al. (2004) argue that the quest for a universal checklist potentially undermines the diversity of qualitative methods and therefore the criteria to evaluate one type of approach might not be appropriate or relevant when assessing another. Unsurprisingly for qualitative researchers the notion of a 'true' gold-standard appraisal tool is unlikely to be congruent with an anti-realist approach to knowledge that many qualitative scholars hold. Certainly caution must be exercised when debating and selecting qualitative checklists.

A comparison of three appraisal tools

With this in mind we now discuss how we came to select the CASP for use in the present study. We began planning the protocol for this review in March 2009. As a useful starting point we consulted the *Cochrane Handbook for Systematic Reviews of Interventions* (Noyes et al., 2008) because this contained a useful chapter on qualitative research. We also consulted the website of the Joanna Briggs Institute, given its role in evidence appraisal. These led us to consider three appraisal tools which were free of cost and easily accessible; we considered the CASP (Public Health Resource Unit, 2006); the QARI Critical Appraisal Instrument (Joanna Briggs Institute, 2007) and a report entitled *Quality in Qualitative Evaluation* (QQE) (Spencer et al., 2003).

a) Critical Appraisal Skills Program. The CASP comprises ten questions addressing clarity of aims; appropriateness of qualitative methodology, research design, recruitment strategy and data collection method; consideration of

reflexivity and ethical issues; rigor of analysis; clarity of findings; and the value of the research. Following each of the ten questions there follows a (number of) prompt(s) for the reviewer to use when addressing the question. The first two questions (which concern the statement of aims and the appropriateness of a qualitative methodology) act as screening questions. The reviewer can then choose whether to continue with the remaining questions. The priority of these two questions suggests they are important and integral to qualitative research, though we would question whether omitting an article on the basis of unclear research aims would be appropriate. In terms of completing the CASP, the first two screening questions invite a yes or no response. The remaining eight questions leave room for the reviewer's comments.

b) QARI Critical Appraisal Instrument. The QARI tool, like the CASP, has ten items. These are presented as statements which the reviewer must decide are either present (yes), absent (no) or unclear. These items cover congruity between the philosophical perspective and the research methodology; congruity between the research methodology and: research aims, data collection methods, data analysis and presentation, and data interpretation; whether the researcher offers a cultural or theoretical reflexive statement; whether the researcher addresses outside influences; the extent to which participant voices are adequately represented; whether the research was ethically conducted; and, the grounding of the conclusions in the data. There is room for the reviewers to sum and score the responses to these ten items and then make a decision to include the article. There is also room for the reviewer's comments.

c) Quality in Qualitative Evaluation. Finally, we also considered the QQE. This report (and tool) begins with a seven page discussion about the context of the framework and the appraisal questions. Following this the authors offer 18 questions. Five of these questions consider the

research findings in relation to their credibility, their contribution to knowledge, how they address the research aims, their ‘generalizability’, and, the application of critical appraisal. One question concerns the design rationale. Two questions consider the sample rationale. Another question assesses the conduct of data collection. A further four questions assess data analysis with respect to the clarity of the analytic approach, the retention of data context, the use of multiple perspectives, and, the communication of data complexity. Two questions consider the way the results are reported, including the grounding of interpretation in the data. The remaining three questions consider reflexivity, ethics and audibility. Following each of these 18 items there are a number of “quality indicators” which expand on the question. Finally, there is room for the reviewer to write any comments they have. Like the CASP, this appraisal tool does not score the article.

Rationale for the CASP

The selection of an appraisal tool is a subjective decision. Following a comparison of the three appraisal tools, we selected the CASP. Although any of the three we considered could have offered a useful framework to appraising quality, we felt the CASP occupied some middle ground. We acknowledge the work of Hannes et al. (2010) who suggest that the JBI (QARI) tool is stronger than the CASP in assessing validity. However their article was published over a year after the design of our protocol. Moreover our decision not to employ the QARI was influenced by the lack of guidance for each question: The availability of anchors can facilitate the critical appraisal, particularly for novice reviewers. The anchors help to reduce ambiguity surrounding the question so that a similar interpretation of the question can be achieved between two raters. As we, like others, (Dixon-Woods et al., 2004) felt the QQE was somewhat unwieldy, we settled on selecting the CASP.

Literature search methods

We undertook this review as part of assessing whether qualitative research made a difference to the knowledge of frontline clinicians. As chronic low back pain (CLBP) is a relatively common illness (NICE, 2009) with high costs at personal, social and economic levels (Miller and Timson, 2004; NICE, 2007; Reid, 2004) we settled on this topic. Specifically, our review sought to explore recently published qualitative research on the patient experience of living with CLBP.

Search strategy

Our search strategy was developed by BN in conjunction with a Consultant Anaesthetist (JR) and two Chartered Psychologists (RG & KL). In addition, we consulted a university information specialist for advice. We used a combination of keyword and thesaurus search terms across a broad range of health-related databases: AMED, BNI, CINAHL Plus with Full Text, EMBASE, IBSS, MEDLINE, PsychINFO and PsychARTICLES. We combined the search terms for PsychINFO and PsychARTICLES databases because they use the same thesaurus terms. Appendix 1 shows the search strategy for the CINAHL database as a typical example. The keywords used across databases remained the same, whereas the subject headings chosen differed according to the unique indexing of each database.

Selection criteria

We used the following selection criteria when assessing an abstract’s suitability for inclusion:

1. We defined recent research as articles published within the previous five years. Given the common nature of CLBP we anticipated a large amount of articles. By narrowing our search to recent research we were able to focus on contemporary and therefore more relevant

articles. For ease of searching this period was considered to be from the start of 2004 until the search day, 12 October 2009.

2. We only looked at articles that focussed on the patient experience of CLBP. We decided to narrow our criteria to the patient experience because of the large number of qualitative studies that have been published in the field of CLBP. We believe this enabled the synthesis component to be more manageable. In defining the 'patient experience' we asked whether the article considers:
 - a) what it is like to live with CLBP;
 - b) how CLBP has impacted the life of the patient;
 - c) whether CLBP has changed the patient's life story;
 - d) how CLBP affects different aspects of their "person" (e.g. their cognition, emotions, behaviour).
3. We defined chronic pain as pain lasting more than three months (Bond et al., 2006) in the lower back. Where there were several conditions alongside low back pain, we included the article if it was clear the majority (50% or more) of participants experienced CLBP.
4. We sought a definition of qualitative research and settled on Noblit and Hare's (1988) broad and encompassing definition: the article was selected if it was primarily concerned with generating explanation and understanding that was derived in part from individuals' reported experience and in line with an interpretive framework.
5. We excluded articles that were case-studies, systematic reviews or from dissertation abstracts international. Case-studies often do not provide sufficient information, while systematic reviews typically contain their own syntheses. We excluded articles from dissertation abstracts international as

it was not feasible to request original dissertations.

Abstract review

The abstracts were reviewed by BN and ZR, both psychology graduates. BN has significant experience in undertaking qualitative research in chronic pain; ZR has clinical experience through working in a physical health psychology service. The reviewers independently reviewed the abstracts and decided whether they were suitable, not suitable or uncertain for inclusion according to the selection criteria we described above. Where there were discrepancies the reviewers resolved these through discussion. In cases that could not be resolved, the reviewers consulted a third party (RG/JR/KL). When there was a lack of clarity concerning whether the article met the selection criteria, one of the reviewers (BN) made contact with the author(s) of the article. Typical queries the reviewer made include finding out the number of participants in the article who experienced low back pain, and, the duration of pain chronicity. A good response was received from many of the authors he contacted. We removed duplicate abstracts prior to calculating the inter-rater reliability for the two reviewers.

Critical appraisal

The reviewers obtained full-text articles in cases where the abstract was initially judged as meeting the selection criteria. They subsequently excluded articles that did not meet the selection criteria. As per the QUOROM flow-chart (Figure 1) the reviewers excluded a total of 19 full-text articles. The main reason for excluding articles was because the focus was not on the patient experience of CLBP. Some articles were excluded by the reviewers for several of the reasons listed in Figure 1 and so the corresponding numbers do not add up to 19 articles. The reviewers independently read the articles in conjunction with the CASP (Public Health Resource Unit, 2006) so that a

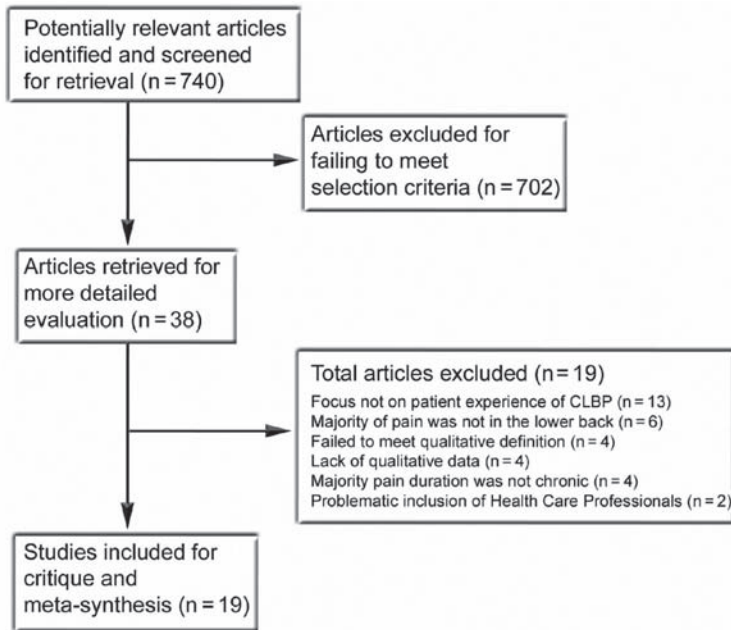


Figure 1. QUOROM statement flowchart.

critique of the article could be made. In line with others (Atkins et al., 2008) we adapted the CASP, replacing the CASP’s definition of qualitative research in question two with Noblit and Hare’s (1988) definition which we described earlier. This offered a more comprehensive understanding of qualitative research. Aside from this change we made no other alterations to the CASP tool. We did not exclude studies based on the result of the critical appraisal unless the authors inappropriately used a qualitative approach (CASP question two). We did not appraise excluded articles.

The reviewers met to discuss their individual appraisal of each article. Together they went through their appraisal for each of the CASP’s ten questions and jointly, produced an appraisal summary for each article. Finally, using the summary of each individual appraisal, one of the reviewers (BN) produced a further summary for each section of the CASP’s ten questions. He re-read a selection of the articles, identifying their strengths and weaknesses. There were a number of reasons for the production of the additional

summary. Primarily we sought to honour the work of the authors whose articles we reviewed; we wanted to produce a fair review and one that we had confidence in. We felt the joint appraisal alone, did not achieve this. The reviewers experienced the common research phenomenon of *practice effect* whereby the later reviews were stronger than the earlier articles they reviewed. Re-reviewing the articles allowed for greater consistency across all 19 articles. Given the fluid nature of the CASP prompts under each question, we found some details were not initially covered. The re-review allowed for such details to be obtained. Overall we believe this additional process enabled us to produce a thorough review, subsequently strengthening the credibility of the critique.

Reflections on the critical appraisal process

As first author and one of the main reviewers (BN), I wanted to offer some reflections on the process of undertaking the critical appraisal.

The second reviewer (ZR) and I met up a number of times over a 12 month period. The first stage of the research was to decide which articles to request. This involved independently reviewing the abstracts, comparing our selection and then discussing our disagreements. The second stage was similar in many respects in that we independently read the 19 full-text articles, made notes on them using the CASP and then met to develop a synthesized critique.

As a doctoral student undertaking qualitative research, I had more experience in working with qualitative research than ZR; this certainly affected the dynamic between us in the early months of the research. I was aware of this and exercised caution in expressing my own voice, so that I facilitated my colleague to express her dissent where necessary. The earlier phase of the research was characterized by developing a joint understanding of how we interpreted the protocol. As time progressed we both developed confidence and ZR developed greater freedom in expressing her opinion. I believe my colleague helped bring a healthy balance to the process of discussion; as I naturally lean towards the critical side, it was a great strength to have a colleague bring another voice into the conversation and decision making process, particularly during the joint synthesis of the CASP critique.

The supervisory team (RG/JR/KL) also provided invaluable input. In the early days we drew on their depth of clinical and academic experience in resolving disagreements that we could not resolve ourselves. This occurred at various stages of the research process; for instance in understanding how CLBP was defined and in making key methodological decisions (such as inter-reliability calculations). Towards the end of the research process I encountered a further dilemma which I resolved with the support of my academic supervisor (RG). As a doctoral student, two of the articles we reviewed were authored by my future external examiner. Through liaising with my supervisor and the journal editor we were able to assess the potential impact on the critique that

ensured it was fair and balanced. We are confident this impact is minimal and perhaps more importantly, that our disclosure allows us to maintain an integrity of transparency. Research can be replete with such dilemmas and we believe that reflexivity and audibility go a long way to successfully holding and managing these tensions. We encourage others to take similar reflexive steps.

Results

Results of abstract selection

The search yielded a total of 1482 abstracts (see Table 1). There were a total of 740 unique abstracts and 742 duplicate abstracts. One abstract was unintentionally treated as a duplicate and subsequently found to be a unique abstract. This was because there were two other updated abstracts as part of a Cochrane Systematic Review, all sharing similar titles and authors. In any case all three abstracts failed to meet the selection criteria. The first author carried out a re-checking process of the abstracts and found this to be an isolated incident. The inter-rater reliability was calculated without this abstract. The authors are confident this had no effect on the review. The kappa value for the two reviewers was 0.573. Orwin (1994, cited by Higgins and Deeks, 2008) suggests kappa values between 0.40 and 0.59 reflect fair agreement whereas values between 0.60 and 0.74 suggest good agreement. The reviewers' agreement can therefore be considered fair, approaching good.

Characteristics of final sample

Of the 740 unique abstracts, 38 full-text articles were requested. The reviewers independently read these and following discussion agreed that 19 of these met the selection criteria. Table 2 provides detailed information about these articles. For one of the articles (May, 2007), the reviewers considered an earlier article (May, 2001) referenced by the author as this contained

Table 1. Breakdown of Resultant Abstracts & Articles Selected Across Databases

Database	Unique Abstracts	Duplicate Abstracts ^a	Total Abstracts	Full-Text Articles Requested ^a	Critiqued Articles
AMED	65	–	65	8	5
BNI	9	2	11	4	1
CINAHL	148	66	214	13	8
EMBASE	404	111	515	6	1
IBSS	29	7	36	2	1
MEDLINE	67	452	519	2	1
PsycINFO & PsycARTICLES	18	104	122	3	2
TOTAL	740	742	1482	38	19

Note: ^aDuplicate abstracts were removed in a cumulative manner. Hence the BNI database contained 2 abstracts that were duplicated in the AMED database. In the same manner, there were full-text articles that were requested in earlier databases (i.e. CINAHL), that were also present in later databases (i.e. MEDLINE) but which were not counted for the sake of calculating inter-rater reliability.

more detailed information surrounding the method. A further article (Rahman et al. 2004, referenced by Harding et al., 2005) was also considered by BN during the summary process.

The majority of studies draw on participants from the United Kingdom. Three studies emanate from Sweden and two studies from Australia. The remaining two studies originate from the United States and from Holland. The mode publication year is 2007. The approach, data collection method and analysis method are presented separately given these are far from singular, unitary concepts. The most common theoretical frameworks used are narrative (5) and phenomenology (5) approaches. All but two studies used interviews to collect data, either individually or as part of a focus group: one used an open-ended questionnaire and the other used a cumulative group discussion. Some of the studies supplemented interview data with a questionnaire. Finally for data analysis, six studies used a form of phenomenology, six studies used a form of thematic analysis, four studies used a form of Grounded-Theory approach, four studies used a framework approach and one study used “common ideas in qualitative analysis (Dey, 1993)” (Hansson et al., 2006: 2186). These approaches were sometimes mixed.

Critique results

We wish to emphasize that our critique is based upon the words authors have used in their article. These words may or may not effectively represent the quality underpinning their research. Our discussion therefore reflects the tensions of the context in which qualitative research is published. Our intention is to stimulate discussion about how most effectively researchers can present their work to a research community that does not wholly accept qualitative research as a credible endeavour. We present the results of the critique as per the CASP headings. For brevity's sake we reference individual articles using a study code, which is detailed in Table 2.

1. Was there a clear statement of the aims of the research? We judged that 16 of the 19 articles presented the aims in a very clear manner. Concerning the remaining three articles we thought article S19 concisely explained what the study consisted of but felt the authors could have presented the aims clearer; where articles reported on a specific theme (S6 and S15) rather than on the study findings as a whole, we believe the statement of aims was weaker.

Table 2. Full-Text Articles Selected for Critique & Synthesis

Article Code	Authors	Year	Country of Sample	Title	Approach	Method	Analysis
S1	Ong et al.	2004	United Kingdom	Establishing self and meaning in low back pain narratives	Narrative approach	Questionnaire and interview	Open coding, in vivo coding and category formation
S2	Satink et al.	2004	Holland	Daily occupations with or without pain: Dilemmas in occupational performance	Narrative approach	Interview	Constant comparison
S3	Busch	2005	Sweden	Appraisal and coping processes among chronic low back pain patients	Glaser's grounded theory	Interview	Glaser's grounded theory
S4	Harding et al.	2005	United Kingdom	"It struck me that they didn't understand pain": The specialist pain clinic experience of patients with chronic musculoskeletal pain	Inductive qualitative approach	Questionnaire and interview	Framework approach
S5	Hansson et al.	2006	Sweden	Sickness absence and sickness attendance - what people with neck or back pain think	Exploratory longitudinal study	Interview	Qualitative analysis
S6	Osborn and Smith	2006	United Kingdom	Living with a body separate from the self: The experience of the body in chronic benign low back pain: An interpretative phenomenological analysis	Interpretative phenomenological analysis (IPA)	Interview	IPA
S7	Walker et al.	2006	United Kingdom	The experience of chronic back pain: Accounts of loss in those seeking help from pain clinics	IPA; narrative approach to data collection	Interview	Thematic analysis (and IPA?)
S8	Afrell et al.	2007	Sweden	Living with a body in pain - between acceptance and denial	Bodily existential perspective	Interview	Phenomenological - hermeneutic
S9	Campbell and Guy	2007	United Kingdom	'Why can't they do anything for a simple back problem?': A qualitative examination of expectations for low back pain treatment and outcome	Narrative approach	Cumulative group discussion	Thematic analysis
S10	Corbett et al.	2007	United Kingdom	Living with low back pain - stories of hope and despair	Mixed methods	Questionnaire and interview	Thematic analysis
S11	Holloway et al.	2007	United Kingdom	The stigmatisation of people with chronic back pain	IPA; narrative approach to data collection	Interview	IPA and thematic approach

(Continued)

Table 2. (Continued)

Article Code	Authors	Year	Country of Sample	Title	Approach	Method	Analysis
S12	Liddle et al.	2007	United Kingdom	Chronic low back pain: Patients' experiences, opinions and expectations for clinical management	No approach explicitly stated	Focus group interview and background questionnaire	Thematic analysis
S13	May	2007	United Kingdom	Patients' attitudes and beliefs about back pain and its management after physiotherapy for low back pain	Explorative qualitative approach	Interview	Framework analysis
S14	Smith and Osborn	2007	United Kingdom	Pain as an assault on the self: An interpretative phenomenological analysis of the psychological impact of chronic benign low back pain	Idiographic, phenomenological perspective	Interview	IPA
S15	Carnes and Underwood	2008	United Kingdom	The importance of monitoring patient's ability to achieve functional tasks in those with musculoskeletal pain	Phenomenological / ethnomethodological approach	Questionnaire and interview	Framework approach
S16	Hush et al.	2009	Australia	Recovery: What does this mean to patients with low back pain?	No approach explicitly stated	Questionnaire and focus group interview	Framework analysis
S17	Slade et al.	2009	Australia	Stigma experienced by people with nonspecific chronic low back pain: A qualitative study.	No approach explicitly stated	Focus group interview and demographic questionnaire	Grounded theory
S18	Snelgrove and Lioffi	2009	United Kingdom	An interpretative phenomenological analysis of living with chronic low back pain	Longitudinal qualitative approach	Interview	IPA
S19	Vroman et al.	2009	United States	Now let me tell you in my own words: Narratives of acute and chronic low back pain	Narrative approach within a Personal Project Analysis study	Open-ended question in a questionnaire	Thematic content analysis

Occasionally we found articles did not put forward a strong rationale for why the topic was worth considering and furthermore some articles (e.g. S10) had unwieldy introductions: a more concise introduction would suffice and might better engage the reader. We believe article S7 offers a strong example of an introduction: The research aims are situated in the context of background literature and the literature review adopts at times, a critical stance. This critique helps form the rationale for the study.

2. Is a qualitative methodology appropriate? The reviewers excluded studies that inappropriately used a qualitative approach therefore by default, all the articles in the final sample appropriately used a qualitative approach. We did hold some concerns surrounding the stated methodology within this sample. As previously mentioned, the theoretical approach, method of data collection and the analysis method represent different concepts. Some articles drew on approaches that guide the entire study, namely Interpretative Phenomenological Analysis (IPA) or Grounded Theory. However the majority of articles highlighted multiple theoretical influences. A shortcoming in seven of the articles we reviewed (S5, S10, S12-13, S16-18) was the lack of an adequate justification or explanation of the theoretical framework that guided the research. We found three articles (S12 and S16-17) did not offer an indication of the type of qualitative approach used. Given the presence of multiple theoretical influences (in those articles that highlight them) it is imperative for researchers not only to make explicit the theoretical approach(es) to their research (e.g. drawing on narrative theory), but also to offer some background and explanation on such theory and further, to defend the value of drawing on their chosen framework(s).

In contrast we found on a number of occasions authors sought to explain the rationale or the philosophy of a qualitative approach per se. Such explication is not necessarily wrong or a shortcoming, however we believe this indicates qualitative research might continue to be less

privileged by the research community. The 'need to explain' has the potential to consume words. Such words might be better directed towards the discussion of reflexivity. For instance in two of the articles the authors offer an outline of what qualitative research is (S12 and S17) and in three articles the authors describe the purpose of qualitative approaches (S7, S12 and S17). Others highlight the appropriate use of qualitative approaches for the study of human experience (S2 and S16). Authors of one article (S18) suggest the qualitative methodology they used (IPA) should not be seen as seeking to replace quantitative approaches. Finally one article (S11) describes how qualitative research attains credible data. We will discuss some of the concerns surrounding these findings later on.

3. Was the research design appropriate to address the aims of the research? One of the shortcomings of the CASP is the ambiguity surrounding what research design is. The sub-statement asks "have they discussed how they decided which methods to use?" (Public Health Resource Unit, 2006). This question appears to focus on the method. We believe that design encompasses both methods and methodology, because the theoretical influence will have an impact on the data collection procedure selected. The reviewers subsequently assessed the appropriateness of both the theoretical framework chosen as well as the specific procedure followed.

Only four articles (S6, S7 and S18-19) offered a strong justification for both the underpinning methodology and the method used. Authors frequently offered a rationale for either method or methodology, but rarely for both. In some instances, authors detailed a theoretical framework but then used a different method of analysis. This is evident from the articles listed in Table 2. We believe authors need to discuss and justify how the method of analysis relates to their stated theoretical perspective of choice.

Articles such as S6 represent studies with a strong research design. Authors of such articles describe the aims of their philosophical approach

with some historical context. Furthermore they justify their choice of approach specifically in relation to the research aims. They also justify their use of semi-structured interviews.

4. Was the recruitment strategy appropriate to the aims of the research? Seven of the articles explicitly described inclusion criteria (S3, S8, S12 and S17), exclusion criteria (S2) or both (S5 and S16). The remaining 12 did not explicitly describe the study selection criteria, however the sample was often broadly described in terms of similarities. Four of the articles (S3, S5, S7 and S17) provided information on non-responders, or those who did not participate, though for two articles (S8 and S14) this was not relevant since all those invited took part. Information on non-responders offers the reader greater insight into the study.

We felt one of the main shortcomings in this section was the lack of justification for the decisions made in the recruitment process. Five articles (S1, S4, S6, S9 and S11) could have provided greater detail of the recruitment process. In two articles the authors employed sampling processes which might not be considered appropriate for qualitative research. For instance S5 used a random sampling process and S13 used a systematic sampling strategy. In both cases it was unclear why these strategies were used. In contrast we did find that in eight articles (S2, S6, S7, S10-11, S13, S15 and S18) the authors described using purposive sampling processes, in one article (S3) the author described using theoretical sampling and in a further article (S4) the authors described the sampling strategy in terms of both of these. All of the 19 articles offered data on participant characteristics. This ranged from minimal information to detailed data presented in tables.

5. Were the data collected in a way that addressed the research issue? Within this question the CASP offers a number of areas for the reviewer to consider including setting, clarity of data collection method and the medium of data collection, and, the application of saturation

principles. We note a number of findings: With regard to offering a rationale, we found authors often did not justify the choices made in the data collection method. With respect to the research context we believe the setting was unclear in three of the articles (S6, S14 and S15). With regard to the questions asked within interviews we found seven articles (S6 and S12-17) offered a topic guide and a further six (S1-3, S7, S9 and S11) offered a brief outline of the types of questions asked. This was not relevant for article S19, in which the authors analysed the responses to a specific question detailed in the article.

One of our main concerns arising from section five of the CASP was the lack of discussion surrounding data saturation. Only five of the 19 articles explicitly discussed saturation (S3-4, S9, S13 and S15). The principle of saturation is not necessarily pertinent to all research designs however we believe reference should be made to the choice of sample size (as S14 achieves) and how the decision to stop recruitment came about. In the case of an unplanned analysis (for instance, in the case of open-ended questionnaire items) we suggest that authors should confirm their belief (or not) in the saturation of the themes.

6. Has the relationship between researcher and participants been adequately considered? We did not find a single article which provided evidence of good reflexive practice. At best, some articles offered a few lines which suggested reflexivity had been considered. In some articles there was simply no evidence of reflexivity. Despite the lack of an exemplar we suggest that reflexivity was evidenced by authors in a number of ways. In article S2 one of the researchers maintained a field journal to become aware of bias, however the authors do not describe how this impacted on the trustworthiness of their study. In another article (S9) the researchers fed back emerging themes to participants. Articles S7 and S11 (undertaken by the same authors) described reflexive techniques such as seeking a diverse sample, avoiding an interview schedule

and deferring additional questions until the end of the study, to avoid bias. And finally, the authors in article S18 evidenced reflexivity in that they sought to avoid imposing topics onto the individual. We note the presence of positivist overtones in two of the articles. For instance the authors of article S12 were apologetic about the influence of bias on recruitment procedures and in S7 the authors described the desire to avoid socially desirable responses.

7. Have ethical issues been taken into consideration? Although we found the level of detail about ethical procedures varied across the 19 articles, we had no concerns surrounding the conduct of the research. The only limitation we noted was concerning the detail in which ethical procedures were described. For instance in two of the articles (S3 and S10) there was no written statement stating that ethical approval had been obtained though we would reiterate we had no concerns about ethical misconduct in these (or any) articles. Similarly in ten of the articles (S1-4, S8, S10-11, S13, S16 and S19) the authors did not explicitly state whether informed consent was gained. This might reflect an assumption that consent is a pre-requisite to conducting research.

In contrast however we report that many of the articles discussed anonymity and confidentiality issues. Furthermore ten of the articles specified the name of the ethics committee approving the study (S1-2, S4-5, S8, S12, S15-17 and S19). One article (S6) put in place psychological procedures to protect participants who were distressed and another sought to emphasize the disconnected nature of the research from clinical activities (S2).

8. Was the data analysis sufficiently rigorous? The majority of articles had an adequate description of the analysis process. Article S8 offered a good description of the transcript reading and familiarization, the process of producing interview synopses and the subsequent comparison process between synopses: It is clear from the account how the researchers arrived at the

results subsequently giving the reader confidence in the data interpretation.

We found that bias was addressed in all of the studies either through the use of an independent researcher (S3-4, S6-7, S11, S14-15 and S17-18) through the use of multiple researchers undertaking an analysis (S1-2, S4-5 S8-10, S12 and S19), or through discussion with peers (S13 and S16).

The quotes chosen for inclusion were rarely given a rationale behind their selection. We found that three of the articles contained quotes which did not always appear to support the themes (S1, S4 and S15). In two articles (S13 and S16) we felt some of the themes were too broad or simplistic. In the majority of cases however we considered the quotes successfully supported the interpretation or themes.

Finally, we could only find direct evidence of contrasting data in three articles (S3-4 and S10) though a number of articles did offer evidence of data variation. We would caution placing too much confidence in this aspect of the critique however, given that the reviewers found it difficult to assess contradictory data and data variation.

9. Is there a clear statement of findings? Although many of the authors presented their findings in a clear manner, we judged a lack of clarity in six of the articles (S1, S3, S6, S10, S13 and S18). The use of more headings might have offered the reader greater clarity. We found it rare for authors to consider alternative explanations in their discussion. We note that in 11 articles (S2-3, S5, S7-8, S11-15 and S17) the authors made some reference to the credibility of their findings. The issue of credibility was often described in the analysis section and it might therefore be cumbersome for this to be repeated again in the discussion. Finally, authors typically discussed the research in line with the study aims, though it was not typical for them to explicitly restate the study aims and discuss the results with reference to these. This became further complicated where the focus of the article was on the discussion of an entire theme that emerged

rather than on all of the study's results. In such circumstances we felt the link between the theme and the study aims could have been clearer in some of the articles.

10. How valuable is the research? All of the articles offered clinical implications in discussing their findings, though in two of the studies (S8 and S10) we felt these were limited or described in an abstract manner. We believe all 19 articles offer a valuable contribution to a wider understanding of how CLBP is experienced.

Whereas authors often made reference to previous research in their discussion, many of these references were typically in support of their research findings, demonstrating how their study extends previous research. Only a minority of references within articles represented a challenge to the study's findings. A more critical stance might be warranted in which authors both compare and contrast their findings with previous research. Finally we note in eight of the articles the authors did not appear to suggest areas for further research (S1, S3-4, S9-10, S15 and S18-19).

Discussion

Summary

Following an appraisal of 19 articles the reviewers identified a number of strengths within this sample. We believe most articles had clear aims, described in detail the questions researchers asked participants, contained research conducted to a good ethical standard, offered a rigorous analysis, sought to address bias, described their findings in a clear manner and proffered clinical implications. We note a number of key limitations however. Researchers typically failed to offer an adequate explanation and rationale for the guiding theoretical framework used, they did not give an adequate account for the decisions made in the recruitment and data collection process, they often did not describe either saturation or how the decision to stop collecting data was made, and finally, they

failed to demonstrate adequate reflexivity throughout the research process.

Are we thinking qualitatively?

This review has led us to question why researchers feel the need to explain qualitative methods. The following points highlight some of our concerns arising from the review:

- (a) The 'need to explain' qualitative methodology and the use of inappropriate sampling strategies left us with the impression that some researchers are still influenced by positivist ideas. Detailed explication of qualitative methodology is perhaps unnecessary; qualitative research is after all an established discipline. For instance some (Miller, 2010) note within nursing practice, qualitative approaches have become the "standard way in which researchers generate knowledge" (p. 193). Similarly others (Pope and Mays, 2009) suggest qualitative research has become well-accepted within research conducted in the health services, with the focus moving towards concerns surrounding rigor and quality. We therefore believe researchers need to move away from justifying qualitative approaches and instead explain why their methodology is the appropriate approach for the research question.
- (b) We occasionally found that researchers did not follow the methodology through into all aspects of the research process. This is evident to the reader from Table 2. There are many studies here which draw on influences (in their approach to the research) which are not followed through in to the data analysis. A rigorous qualitative study cannot employ a "cook-book" approach to study design without thoughtful justification. We believe researchers need to avoid the mindset that "anything goes"

in the design process of qualitative research and employ strong justification in outlining their approach.

- (c) Only a minority of studies described how saturation was achieved. While accepting that not all qualitative methods call for data saturation, we suggest researchers do need to articulate clearly how they reached a decision to stop data collection, or in the case of retrospective data, how confident the researchers are that the themes are well-saturated. Of course unsaturated themes do not make for a worthless study. The richness of qualitative data allows for valuable insights to be gained without attaining saturation. However researchers must be transparent in discussing their recruitment process and how they decided to stop collecting data.
- (d) Finally, the reviewers were concerned about the use of independent checks to prevent bias; either in the form of multiple coders, using an independent coder or seeking external advice. Although efforts to privilege the participant's voice should be applauded, it was not evident to us that independent checks always had this purpose. The notion of multiple coders has echoes of positivism in which there is a single reality to be uncovered. Our concerns lie in the apparent unacceptability of a single researcher's interpretation. Done with sensitivity, adequate reflexivity and within a constructivist framework, the interpretation of a single researcher should be considered an acceptable qualitative approach. For instance Daly (1997, cited by Horsburgh, 2003) believes that "all facts are interpreted facts" (p. 308). He also argues for the importance of "preserving" the voice and meaning of the participant. We therefore believe it is more important to deliver a single interpretation that

contains the meaning and nuances (as close to those) intended by the participant, than to have multiple interpretations that potentially fail in this.

Reflexivity: are we practicing what we preach?

Reflexivity is an integral part of conducting good qualitative research. Furthermore there are different approaches to reflexivity. For some it is about minimizing the impact of the researcher on the data (Sin, 2010), whereas for others, reflexivity is about making use of subjectivity (Gough, 2003). This stands in contrast to research located within a positivist framework, where subjectivity is seen as a threat to validity. Gough suggests reflexivity "facilitates a critical attitude towards locating the impact of research(er) context and subjectivity on project design, data collection, data analysis, and presentation of findings" (p. 22). In the current review we were surprised at the lack of reflexivity demonstrated across the 19 articles. In the few articles that were reflexive we occasionally felt these had positivist overtones, suggesting different perceptions about the function of reflexivity. There was not a single article that we considered an exemplar of reflexive practice.

In the course of her clinical work, one of us (RG) came across a recently published article discussing the experiences of visually impaired individuals. This article might be considered such an exemplar. Thurston (2010) offers a detailed reflexive account following her results. She became aware of her own assumptions and prejudices through the process of journaling and bracketing. She acknowledges the potential effect of such bias on her data. Thurston then describes the impact of the research on herself, noting how it has challenged the way she sees herself as a blind individual.

A consequence of failing to be reflexive is to risk obscuring the subjective. Researchers might subsequently present their approach, intentionally or otherwise, as objective. Even

within this review the authors made subjective decisions: The selection criteria consist of decisions jointly made based on clinical and academic experience; nevertheless they are subjective calls. Horsburgh (2003) might agree, noting "reflexivity refers to active acknowledgement by the researcher that her/his own actions and decisions will inevitably impact on the meaning and context of the experience under investigation." (p. 308). We believe that researchers must acknowledge and embrace the subjectivity within their work. This will help form the basis of rigorous qualitative research.

Through presenting preliminary results at the 16th Annual Qualitative Health Research Conference, we gained insight into some of the possible explanations for the lack of reflexivity. Delegates highlighted the restrictions that journal editors place on manuscript length. Reflexive accounts therefore might be cut short or omitted because of word limit restrictions. In other words the researchers may have taken reflexive steps in the research process but may not have reported doing so in their article. There is then the difference between the *conduct* and the *reporting* of the study. A limitation of evaluating the conduct of the study based on the *report* of the article is that a critique can only be made based on the article contents.

It is not just qualitative work that could benefit from evidence of reflexivity. The notion of reflexivity appears to be alien in quantitative research. If reflexivity benefits research through facilitating a critical examination of how the researcher interacts and influences data, then surely quantitative researchers would also do well to embrace a reflexive approach? Quite how this could be incorporated into a very different paradigm remains to be seen, but it is a debate that is yet to be seriously entertained.

Although we have advocated in this article for greater evidence of reflexivity, we do not pretend that reflexive practice is without its difficulties. One of the dangers of incorporating reflexive narratives into our work is to risk disengaging other parts of the research community for whom personal narrative is synonymous

with stream of consciousness. This risks the perception that such work is inadmissible or indeed irrelevant. Such a risk is perhaps acutely felt in auto-ethnographical work (e.g. Holt, 2003) where narratives of the self are (for some, questionably) used as data.

A further problem of open reflexive accounts is to risk unnecessarily undermining the quality of the research. In a culture where reflexivity is not yet widely embraced, to expose all of one's subjectivities, internal tensions and weaknesses is to create an unlevel playing field among the many articles that do not undertake this practice. One solution might be to make available reflexive notes online, in a forum detached from the journal article yet referenced within it. This would allow reflexivity to be evidenced while removing such exposition from the peer-review process. There are certainly challenges for the reflexive researcher to manage and we acknowledge that incorporating reflexivity is no easy feat.

Limitations

The authors sought to design and implement a rigorous and trustworthy review process. The audit trail, maintained through field notes by one of the reviewers (BN), helped towards assuring this. However the strength of the review was limited in a couple of respects. Firstly, upon reflection we feel the appraisal process could have benefited from the involvement of another experienced qualitative researcher. Although the main reviewer is a doctoral researcher engaged in qualitative work, the second reviewer (ZR) lacked experience in qualitative research. Despite this, we consider her clinical experience within physical health psychology particularly helped with identifying the clinical utility of the research studies. Furthermore we believe she brought a healthy challenge to the joint appraisal discussions, enabling a more balanced approach to be achieved. Secondly, we were disappointed not to have achieved a higher inter-rater reliability. The fluid definition of "patient experience" might have proved

too loosely defined and was therefore partly responsible for a lower rate of agreement than we had hoped for; it was often difficult to determine from the abstract alone whether an article was appropriate.

Recommendations

Given qualitative research is an established field, researchers might do well to consider before describing what qualitative research is, whether their scientific counter-parts would use valuable manuscript space describing more established methodologies such as the RCT. We suggest there is scope for researchers to grow in using qualitative approaches in a confident and rigorous manner. The research community must resist the urge to justify qualitative methods. When editors ask researchers to describe qualitative approaches, authors should consider offering a rebuttal, drawing on the literature that demonstrates the established nature of qualitative methods.

Moving on to reflexivity, more needs to be done on a number of levels. We suggest that reflexivity needs to be more clearly evidenced in articles. However this forces difficult decisions about what to leave out, in order to meet the word length requirements that journals require. At the 16th Annual Qualitative Health Research Conference delegates suggested that researchers need to “lobby” journal editors. Indeed editors of both *qualitative* and *quantitative* journals need to develop an expectation of reflexivity within research articles. Reviewers similarly must welcome open and honest accounts of reflexivity, even if this appears to acknowledge what might be seen as “biased interpretations”. Finally, we believe authors must ensure their research is both conducted reflexively and reported in a reflexive manner. Granted not all editors will be open to devoting journal space to accounts of reflexivity. However a short two-line reflexive statement is enough to demonstrate authors have considered the impact they, as researchers, have had on the research process. Perhaps the extra space will come from removing the

statement, “it’s only a qualitative study” (Morse, 2008).

Conclusion

There is a substantial body of qualitative research focusing on CLBP; the majority of which, in this review, stems from the United Kingdom. The bulk of the articles we reviewed describe research conducted to high standards with findings that can be transferred into the clinical context. However the lack (in substance or in reporting) of some core qualitative principles is of concern. For qualitative research to be considered credible, those using its principles need to offer evidence to demonstrate their work is rigorous and trustworthy.

Acknowledgements

We acknowledge and thank Stephen Gough, Liaison Librarian at Birmingham City University, for his guidance in the protocol design; Dr Sayeed Haque, Consultant Statistician, for his support with calculating inter-rater reliability; we thank the two anonymous reviewers and the journal editor for all their invaluable input; we also acknowledge that parts of the critique contained within this article were presented in rudimentary form at the 16th Annual Qualitative Health Research Conference, 3rd-5th October 2010 in Vancouver, and thank the delegates for sharing their thoughts.

Declaration of Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

Ben Newton is a PhD student at Birmingham City University and has been funded by a bursary.

References

- Afrell M, Gabriele B, and Rudebeck CE (2007). Living with a body in pain—between acceptance and denial. *Scandinavian Journal of Caring Sciences*, 21(3): 291–296.
- Atkins S, Lewin S, Smith H, et al. (2008). Conducting a meta-ethnography of qualitative literature: Lessons learnt. *BMC Medical Research Methodology* 8: 21. Available at: www.biomedcentral.com/1471-2288/8/21.

- Barbour RS, and Barbour M (2003). Evaluating and synthesizing qualitative research: the need to develop a distinctive approach. *Journal of Evaluation in Clinical Practice*, 9(2): 179–186.
- Bond M, Breivik H, Jensen T, et al. (2006). Pain associated with neurological disorders. In: *Neurological disorders: Public health challenges*. Retrieved from: www.iasp-pain.org. Geneva: WHO.
- Busch H (2005). Appraisal and coping processes among chronic low back pain patients. *Scandinavian Journal of Caring Sciences*, 19(4): 396–402.
- Campbell C and Guy A (2007). ‘Why can’t they do anything for a simple back problem?’: A qualitative examination of expectations for low back pain treatment and outcome. *Journal of Health Psychology*, 12(4): 641–652.
- Carnes D and Underwood M (2008). The importance of monitoring patient’s ability to achieve functional tasks in those with musculoskeletal pain. *International Journal of Osteopathic Medicine*, 11(1): 26–32.
- Corbett M, Foster NE, and Ong BN (2007). Living with low back pain—Stories of hope and despair. *Social Science & Medicine*, 65(8): 1584–1594.
- Daly K (1997). Re-placing theory in ethnography: A postmodern view. *Qualitative Inquiry*, 3(3): 343–365. Cited by Horsburgh D (2003) Evaluation of qualitative research. *Journal of Clinical Nursing*, 12(2): 307–312.
- Dey I (1993). Qualitative data analysis. *A user-friendly guide for social scientist*. London: Routledge.
- Dixon-Woods M and Fitzpatrick R (2001). Qualitative research in systematic reviews. *BMJ* 323(7316): 765–766.
- Dixon-Woods M, Shaw RL, Agarwal S, et al. (2004). The problem of appraising qualitative research. *Quality and Safety in Health Care*, 13(3): 223–225.
- Dixon-Woods M, Bonas S, Booth A, et al. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6(1): 27–44.
- Gough B (2003) Deconstructing reflexivity. In: Finlay L and Gough B (eds) *Reflexivity: A practical guide for researchers in health and social sciences*. Oxford: Blackwell Science Ltd, 21–35.
- Hannes K, Lockwood C, and Pearson A (2010). A comparative analysis of three online appraisal instruments’ ability to assess validity in qualitative research. *Qualitative Health Research*, 20(12): 1736–1743.
- Hansson M, Boström C, and Harms-Ringdahl K (2006). Sickness absence and sickness attendance—What people with neck or back pain think. *Social Science & Medicine*, 62(9): 2183–2195.
- Harding G, Parsons S, Rahman A, et al. (2005). ‘It struck me that they didn’t understand pain’’: The specialist pain clinic experience of patients with chronic musculoskeletal pain. *Arthritis and Rheumatism*, 53(5): 691–696.
- Holloway I, Sofaer-Bennett B, and Walker J (2007). The stigmatisation of people with chronic back pain. *Disability and Rehabilitation*, 29(18): 1456–1464.
- Holt NL (2003). Representation, legitimation, and autoethnography: An autoethnographic writing story. *International Journal of Qualitative Methods*, 2(1). Retrieved from: www.ualberta.ca/~iiqm/backissues/2_1final/html/holt.htm
- Horsburgh D (2003). Evaluation of qualitative research. *Journal of Clinical Nursing*, 12(2): 307–312.
- Hush JM, Refshauge K, Sullivan G, et al. (2009). Recovery: What does this mean to patients with low back pain? *Arthritis and Rheumatism*, 61(1): 124–131.
- Joanna Briggs Institute (2007) *SUMARI: The Joanna Briggs Institute system for the unified management, assessment and review of information*. Adelaide, Australia: Joanna Briggs Institute. Retrieved from: www.joannabriggs.edu.au
- Katrak P, Bialocerkowski AE, Massy-Westropp N, et al. (2004). A systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology*, 4, 22. Retrieved from: www.biomedcentral.com/1471-2288/4/22
- Koch T (2006). Establishing rigour in qualitative research: the decision trail. *Journal of Advanced Nursing*, 53(1): 91–103.
- Liddle SD, Baxter GD, and Gracey JH (2007). Chronic low back pain: Patients’ experiences, opinions and expectations for clinical management. *Disability and Rehabilitation*, 29(24): 1899–1909.
- May S (2001). Patient satisfaction with management of back pain. Part 2: An explorative, qualitative study into patients’ satisfaction with physiotherapy. *Physiotherapy*, 87(1): 10–20.
- May S (2007). Patients’ attitudes and beliefs about back pain and its management after physiotherapy for low back pain. *Physiotherapy Research International*, 12(3): 126–135.
- Miller J and Timson D (2004). Exploring the experiences of partners who live with a chronic low

- back pain sufferer. *Health and Social Care in the Community*, 12(1): 34–42.
- Miller WR (2010). Qualitative research findings as evidence: Utility in nursing practice. *Clinical Nurse Specialist*, 24(4): 191–193.
- Morse JM (2008). "It's only a qualitative study!" Considering the qualitative foundations of social sciences. *Qualitative Health Research*, 18(2): 147–148.
- NICE (2007). *Low back pain: Final scope*. Dated 9 May 2007. Retrieved from: www.nice.org.uk/nicemedia/pdf/LowBackPain_FinalScope.pdf
- NICE (2009). *Low back pain: Early management of persistent non-specific low back pain*. London: National Institute for Health and Clinical Excellence. Retrieved from: www.nice.org.uk/nicemedia/pdf/CG88NICEGuideline.pdf.
- Noblit GW and Hare RD (1988). *Meta-Ethnography: Synthesizing qualitative studies*. California, CA: SAGE.
- Noyes J, Popay J, Pearson A, et al. (2008). Chapter 20: Qualitative research and Cochrane reviews. In: Higgins JPT and Green S (eds) *Cochrane handbook for systematic reviews of interventions version 5.0.1* [updated September 2008]. The Cochrane Collaboration. Retrieved from: www.cochrane-handbook.org
- Ong BN, Hooper H, Dunn K, et al. (2004). Establishing self and meaning in low back pain narratives. *The Sociological Review*, 52(4): 532–549.
- Orwin RG (1994). Evaluating coding decisions. In: Cooper H and Hedges LV (eds) *The handbook of research synthesis*. New York (NY): Russell Sage Foundation. Cited by Higgins JPT and Deeks JJ (2008) Chapter 7: Selecting studies and collecting data. In: Higgins JPT and Green S (eds) *Cochrane handbook for systematic reviews of interventions version 5.0.1* [updated September 2008]. The Cochrane Collaboration. Retrieved from: www.cochrane-handbook.org.
- Osborn M and Smith JA (2006). Living with a body separate from the self. The experience of the body in chronic benign low back pain: An interpretative phenomenological analysis. *Scandinavian Journal of Caring Sciences*, 20(2): 216–222.
- Pope C and Mays N (2009). Critical reflections on the rise of qualitative research. *BMJ*, 339, b3425.
- Public Health Resource Unit (2006). Critical Appraisal Skills Programme (CASP): 10 questions to help you make sense of qualitative research. England: Public Health Resource Unit. Retrieved from: www.sph.nhs.uk/what-we-do/public-health-workforce/resources/critical-appraisals-skills-programme
- Rahman A, Ambler G, Underwood MR, et al. (2004). Important determinants of self-efficacy in patients with chronic musculoskeletal pain. *The Journal of Rheumatology*, 31(6): 1187–1192.
- Reid M (2004). An assessment of health needs of chronic low back pain patients from General Practice. *Journal of Health Psychology*, 9(3): 451–462.
- Reid B, Sinclair M, Barr O, et al. (2009). Reflections on the methodological challenges of undertaking a meta-synthesis: A response to Heyman. *Social Science & Medicine*, 69(11): 1577–1579.
- Satink T, Winding K, and Jonsson H (2004). Daily occupations with or without pain: Dilemmas in occupational performance. *OTJR: Occupation, Participation and Health*, 24(4): 144–150.
- Sin S (2010). Considerations of quality in phenomenographic research. *International Journal of Qualitative Methods*, 9(4): 305–319.
- Slade SC, Molloy E, and Keating JL (2009). Stigma experienced by people with nonspecific chronic low back pain: A qualitative study. *Pain Medicine*, 10(1): 143–154.
- Smith JA and Osborn M (2007). Pain as an assault on the self: An interpretative phenomenological analysis of the psychological impact of chronic benign low back pain. *Psychology & Health*, 22(5): 517–534.
- Snelgrove S and Liossi C (2009). An interpretative phenomenological analysis of living with chronic low back pain. *British Journal of Health Psychology*, 14(4): 735–749.
- Spencer L, Ritchie J, Lewis J, et al. (2003). *Quality in Qualitative Evaluation: A framework for assessing research evidence: A Quality Framework*. London: Strategy Unit, Government Chief Social Researcher's Office.
- Thurston M (2010). An inquiry into the emotional impact of sight loss and the counselling experiences and needs of blind and partially sighted people. *Counselling and Psychotherapy Research*, 10(1): 3–12.
- Vroman K, Warner R, and Chamberlain K (2009). Now let me tell you in my own words: Narratives of acute and chronic low back pain. *Disability and Rehabilitation*, 31(12): 976–987.
- Walker J, Sofaer B, and Holloway I (2006). The experience of chronic back pain: Accounts of loss in those seeking help from pain clinics. *European Journal of Pain*, 10(3): 199–207.

Appendix I

Search Strategy Used With CINAHL Plus With Full Text

1. low W2 back (title) or low W2 back (abstract)
2. lower W2 back (title) or lower W2 back (abstract)
3. 1 OR 2
4. degenerative (title) or degenerative (abstract)
5. non-specific OR nonspecific (title) or non-specific OR nonspecific (abstract)
6. unexplained (title) or unexplained (abstract)
7. idiopathic (title) or idiopathic (abstract)
8. 4 OR 5 OR 6 OR 7
9. back (title) or back (abstract)
10. 8 AND 9
11. spine OR spinal (title) or spine OR spinal (abstract)
12. musculoskeletal (title) or musculoskeletal (abstract)
13. 11 OR 12
14. 3 OR 10 OR 13
15. pain* OR ache* (title) or pain* OR ache* (abstract)
16. 14 AND 15
17. lumbago (title) or lumbago (abstract)
18. (MH "Low Back Pain")
19. 16 OR 17 OR 18
20. qualitative (title) or qualitative (abstract)
21. constructivism OR constructivist (title) or constructivism OR constructivist (abstract)
22. feminis* AND theory (title) or feminis* AND theory (abstract)
23. narrative* (title) or narrative* (abstract)
24. phenomenology OR phenomenological (title) or phenomenology OR phenomenological (abstract)
25. inductive OR induction* (title) or inductive OR induction* (abstract)
26. grounded W2 theory (title) or grounded W2 theory (abstract)
27. (MH "Qualitative Studies+") (explode)
28. patient* OR doctor* (title) or patient* OR doctor* (abstract)
29. story OR stories (title) or story OR stories (abstract)
30. 28 AND 29
31. 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 30
32. 19 AND 31
33. 32: Limit to Publication Date Current – 2004