**DEVELOPMENT**

# Developing and implementing electronic search strategies to recruit patients with chronic musculoskeletal pain in primary care databases

**Jens Foell, Dawn Carnes, Kate Homer** and **Stephanie Taylor**

NIHR Clinical Lecturer, The Blizard Institute, Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

**Background:** Identifying patients with chronic musculoskeletal pain using database searches is difficult, as chronic pain is not represented with a unique diagnostic code in electronic primary care records. **Aim:** This paper describes the development and implementation of a search strategy to identify patients with chronic musculoskeletal pain in primary care databases to invite them to participate in a randomised controlled trial. **Methods:** We used an exploratory, iterative approach. The first phase involved consultations with IT specialists, practice managers and doctors to gain an understanding about the processes and issues of electronic coding. In the second and third phases, we determined the most appropriate search terms and strategies. In the final phase, we tested, modified and re-tested the search strategy until the quantity and quality of the output appeared good enough to be used in general practices with different IT systems. This strategy was then implemented to recruit participants for a trial. **Findings:** We identified three main search 'domains': prescribing, coding and attendance. We found the most useful identifier for chronic pain was the use of repeat medication. Wide variations in coding terms for chronic pain were seen between practices and individuals. Understanding 'coding cultures' were necessary to inform the electronic searches. In the case of chronic pain, searching on repeat medication for analgesia, low dose antidepressants and carefully selected coding terms captured most relevant patients.

**Key words:** chronic pain; condition coding; health informatics; information infrastructure; read code; search strategy

*Received 22 August 2012; revised 28 March 2013; accepted 28 April 2013;*
*first published online 24 May 2013*

## Introduction

The management of patient data in the UK National Health Service (NHS) is predominantly electronic. In the United Kingdom, primary care clinical software systems are based on 'Read Code' classification and 'SNOMED CT', which includes ICD 10 codes (Bentley *et al*., 1996). These are medical thesauruses of clinical terms, which are organised in hierarchical chapters. They capture patient data such as diagnoses, presenting complaints, occupation, administrative processes, interventions and diagnostic procedures. The systems were developed to be flexible and allow users to interact with the classification

Correspondence to: Dr Jens Foell, NIHR Clinical Lecturer, The Blizard Institute, Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Yvonne Carter Building, 58 Turner Street, London E1 2AB, UK. Email: j.foell@qmul.ac.uk

system in 'natural language' (Read *et al.*, 1995; NHS, 2000). An extensive thesaurus allows the user to choose the most appropriate term to describe each patient. Each term is linked to an equivalent Read code. These codes allow computerised searches. Some health conditions, such as diabetes have standardised uniform coding as primary care practices are reimbursed financially for delivering these services under the Quality and Outcomes Framework (QOF; Department of Health, 2004). Conditions outside these clear 'QOF' categories can be classified without generally agreed coding rules, they are difficult to capture for secondary work like research or service planning. Mental health and physical symptoms in combination are 'hard to code', practitioners can choose multiple variations on themes, these choices allow for flexibility, but introduce diversity in classification (Henningsen and Creed, 2009). Chronic pain is another example of a condition that can be classified using a variety of codes (Nordin *et al.*, 2006). Chronic pain is defined by the International Association for the Study of Pain as 'pain beyond three months duration or the expected time for tissue repair' (Merskey and Bogduk, 1994). It is not represented in an equivalent Read code in primary care but may be recognised at secondary care level by referral to pain specialists, therefore patient identification can be made by service delivery coding (Armstrong, 2011). Our primary aim was to develop a search strategy to identify potential participants in primary care with chronic pain for a randomised controlled trial (RCT; Carnes *et al.*, 2013).

## Methods

We took a very pragmatic approach to this study due to the constantly changing 'live' databases we were using in a real-time clinical setting. We did not develop stringent reliability and validity tests as you would for a diagnostic test as these were not appropriate for this study for two reasons:

 i) patient details in the practice databases were continually being updated and changing, so, regardless of the search parameters, results were liable to change with each search attempted;
 ii) we had no way of knowing exactly how many patients with chronic musculoskeletal pain were registered in the practices at any

one time. Therefore, the validity of our searches could not be tested using sensitivity or specificity calculations as there were no fixed denominators against which to test the accuracy of the searches.

Due to the practical issues of working with live active databases we decided to assess 'feasibility' and 'representativeness' of the quality of our searches. Feasibility in this context was the ease with which the search could be performed, this included: the amount of data transformations or computer commands required to perform the search, the time it took to perform the search, the skills of the person needed to do the work and the difficulties of data processing. Representativeness in this case was how well the search identified an index list of known people with chronic musculoskeletal pain and the estimated prevalence in the practice population with chronic pain (5% of adults). We used a set of 'indexed' or 'reference' patients known to have chronic pain and assessed whether the searches identified them. The study was conducted in primary care practices. We used an exploratory iterative approach to investigate the issues and processes necessary to identify chronic pain patients. We provide a descriptive narrative of the issues that we encountered at each stage of the research.

The method we used involved four exploratory stages (Figure 1). Stage 1 involved a consultative exercise, in stage 2 we generated appropriate search terms and strategies, during stage 3 we tested the search strategies and in stage 4 we evaluated the results of our recommended approach. We assumed that most primary care electronic patient records systems operated using information about patient characteristics/demographics, consultation records, prescriptions, referrals and tests and types of conditions and included a search facility.

## Stage 1: Consultation (search strategy requirements)

Our target population to identify was adults (aged 18 and above) with chronic musculoskeletal pain (pain that persists for more than three months).

This was a scoping exercise to help identify effective search approaches and search terms (labels and domains used to classify patients). We interviewed practice managers, specialists in healthcare
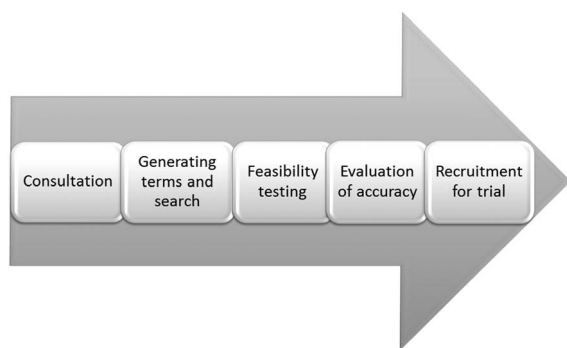
**Figure 1**  The exploratory process

informatics, GPs, pain specialists and administrative general practice staff about their experience of using electronic databases, technical issues, limitations and benefits of using the software, issues around clinical coding preferences and the characteristics of 'pain patients'. We approached practices with different population demographics and different IT systems. We consulted professionals working in primary care based on our own peer networks. J.F. led the information gathering process and in conjunction with D.C. organised the findings into commonly occurring themes and issues.

   We also examined the electronic profiles of known chronic musculoskeletal pain patients to inform our search strategy. We reviewed coding terms, test requests, prescription data, attendance, referral letters and other correspondence held on file.

### Clinical software programmes

   In the United Kingdom, there are two commonly used electronic patient record systems: VISION and EMIS (Egton Medical Information Systems Limited, 1987–2013; In Practice Systems Ltd, 1994–2013). These systems have different interfaces for clinical coding and appointment management. EMIS can be operated at a practice level as EMIS LV, or accessed through an internet browser as EMIS Web. We took these differences into account throughout the study.

### Stage 2: Generating the search approach and terms

From the data gathered in stage 1, we identified the most commonly used codes and terms used to

categorise chronic pain patients. We used the information from the consultations to design the searches that allowed us identify our target group.

### Stage 3: Testing the feasibility of the search strategy

We devised different ways of structuring the search strategies and we tested the feasibility of the search terms and domains. We tested and re-tested variations of the searches until we had a search strategy that suited the capability/capacity of the search engines used in the software and that was feasible for practitioners and practice staff to administer. Stages 3 and 4 were closely associated. Each search was tested for its feasibility and also evaluated for its representativeness.

### Stage 4: Evaluating the accuracy of the searches

For this stage, we tested the search strategies to find the most appropriate search. We collaborated with two GPs and two practice managers from different practices to test several versions of the searches.

   They were familiar with their practice population and able to estimate the accuracy of the search by visually scanning the generated list of patients. We used this to assess feasibility, but we also tested accuracy in more detail by reviewing case notes. We used the following methods to test the quality of the searches:

i) Between 5% and 21% of adults registered with GPs consult with chronic pain over a 12-month period (Jordan *et al.*, 2007). Around 75–80% of the population (UK Census, 2001) are adults, that is, over 18 years old. Our test general practice had 5500 registered patients, we assumed that around 4125 (75%) would be adults of whom between 5% and 21% will have chronic pain. If we take the lower estimate of 5%, we would expect a minimum of 206 patients to be identified by our search.

ii) A reference list of 20 known patients ( ~10% of the above sample estimate) with chronic pain was used to test the inclusiveness of the search. For each search conducted, we checked whether these 20 reference/index patients were selected.

iii) A random sample of 10 further patients identified by the search output list were assessed to see whether they had chronic pain. A clinical member of the staff reviewed the patient record to determine their health state.

Unfortunately due to resource constraints and ethical issues, it was not possible to assess each patient file identified in the searches and classify whether they had chronic pain or not and work out the inter- and intra-rater reliability of the evaluators or the true false positive figure.

Once the search satisfied the above three criteria, we tested it in other GP practices and with other personnel applying the same approach. The final search strategy had to fulfil the above three quality criteria in four pilot practices to be regarded as suitable for use in our RCT.

As the study required auditing data only it did not require ethical approval.

## Results

### Stages 1 and 2: Consultation and generating search terms and the strategy

#### Consultation

We went to four primary care centres to gain and seek information about databases, software and coding. Three were inner London practices and one was a semi-rural practice in the east of England. Eight people were consulted: three GPs, two practice managers, one specialist in health informatics, one pain psychologist and one consultant in pain medicine. Overall, three areas emerged that were considered relevant for identifying chronic pain patients: prescribing, clinical coding and attendance.

#### Prescribing

Medication was generally regarded as the most consistent and reliable way of identifying patients with persistent pain. Commonly prescribed analgesic pharmacological regimes included psychotropic and peripherally acting drugs. When these drugs are coded as 'repeat' or 'automatic' medications (the terms differ in the software programmes), they were more likely to be associated with persistent conditions. This delivery option enables patients to obtain their medication without prior

healthcare contact. There were two ways of accessing information about drug prescribing using the software search functions. Drug names can be entered either directly or searched by the action the drugs have on the body, for example drugs for the central nervous system (CNS) or 'musculoskeletal and joint diseases'. The difficulty with searching via these categories is that they include other non-relevant drugs for musculoskeletal pain. Terms such as 'analgesics' were too broad for our purposes – they included compound drugs like Co-Codamol, weak opiates like Tramadol and simple drugs like Paracetamol alongside drugs used in substance misuse (Methadone) and palliative care (Fentanyl lozenges). Each drug needed to be selected individually so only those relevant to musculoskeletal pain were used in the search. The final selection of relevant drugs is shown in Table 1.

#### Clinical coding

We noted that practices and individual practitioners had different 'coding cultures' for ill-defined conditions like chronic pain. Clinicians reported that they familiarise themselves with the patient's history by looking at patterns of medication consumption, disability assessments, presentations and referrals. The majority of coding described symptoms in body regions via the symptoms chapter in the Read code system. These clinical symptom entries (regardless of software package) are often graded as 'low priority' (category 3) and do not usually get extracted into note summaries. Most relevant information for identifying chronic pain patients was commonly entered as free text or appeared as a combination of presenting complaints such as a trail of multi-problem consultations. This type of input is not searchable using the search-function of the clinical software systems. Table 1 shows the most common terms used to define and classify chronic pain patients in primary care patient records.

#### Attendance

The third relevant descriptor we identified as important was frequent attendance with multi-problem consultations. This was a commonly reported feature of the population presenting with chronic pain.

VISION and EMIS LV have facilities to search for attendance data, but it requires several command

**Table 1** Selection of commonly used repeat prescription drugs and classification codes for chronic pain

| Prescription drugs | Commonly used |
|---|---|
| Controlled drugs | Buprenorphine and Fentanyl patches, Oxycodone |
| Opiates, analgesics | Tramadol, Gabapentin, Dihydrocodeine (DHC), Pregabalin, Duloxetin Codeine, Co-codamol, Codrydramol, NSAIDs (Ibuprofen, Naproxen, Diclofenac), low-dose tricyclic antidepressants like Amitriptyline |
| Over the counter preparations | Paracetamol, topical NSAIDs, rubefacients |
| | |
| Coding terms | Commonly used |
| Inner city General Practice A | Body region plus pain plus free text, 'Fibromyalgia', 'Chronic low back pain' |
| Inner city General Practice B | 'Osteoarthritis', 'Multiple symptoms', 'Arthritis of multiple joints', 'Arthralgia', 'pain in multiple joints' |
| Semi-urban General Practice C | 'Pain' plus free text, 'Back pain', 'OA Knees', 'Somatoform pain disorder' |
| Inner city General Practice D | Pain location plus free text, 'Aches and pains', 'Fibromyalgia', Mood-affective labels, e.g. 'somatoform pain disorder' |

steps until detailed information is available. Searching for patients by attendance was time-consuming and required more sophisticated data-handling skills. We tested this procedure once: it provided information about the attendance frequency and identified the subgroup of patients with increased healthcare utilisation, but it was not feasible within the time and labour constraints of a typical primary care worker (see Figure 2 illustrating the search strategy testing process).

EMIS Web allows basic information about whether or not contacts with healthcare practitioners have happened in a given time interval. This generated useful additional information in our EMIS searches and was incorporated in our final search.

## Stages 3 and 4: Testing the feasibility and evaluating the representativeness of the searches

Four people (two GPs, one practice manager, one IT-consultant) tested a variety of search approaches using 'trial and error' methodology, that is, devising, testing and refining. Each of the domains was tested in various combinations (see flow chart, Figure 2). After around 20 search modifications an agreement was reached, that no further procedural amendments would improve the quality of the search. Reliability and validity testing was hampered by the live database altering the search results on a continuous basis and the lack of a definitive list of chronic pain patients.

The search enhancement modifications we tested included searching for repeat medications of the drugs and using different classification codes as shown in Table 1. We tested the search with three, six or twelve months as observation window for repeat prescriptions. Six months seemed to be most useful time bracket.

After multiple testing of different search strategies, the optimal search was a successive combination of information from repeat prescribing, clinical coding and attendance. The EMIS search allowed additional filtering by attendance, which reduced the output significantly and made the search more appropriate. Figure 2 shows the process and the strengths and limitations of the testing process at each stage for each software system, EMIS and VISION.

The final search structure is shown in Figure 3. The searcher first sets the inclusion parameters for the search, that is, 'active' registered patients 18 years or over. After a conversation with staff familiar with the local 'coding culture', the searcher identifies the most commonly prescribed drugs used for patients with chronic pain that are on repeat prescriptions, followed by the most commonly used classification codes used in the practice. In our test practice, the final search approach generated 209 patients from a list size of 5500 (we estimated the search would generate 206 patients). We tested the approach on four other practices, each search generated patient list sizes around the size of our estimates and included the quality control of 20 index patients and 10 randomly selected patients who could be classified as having chronic pain.
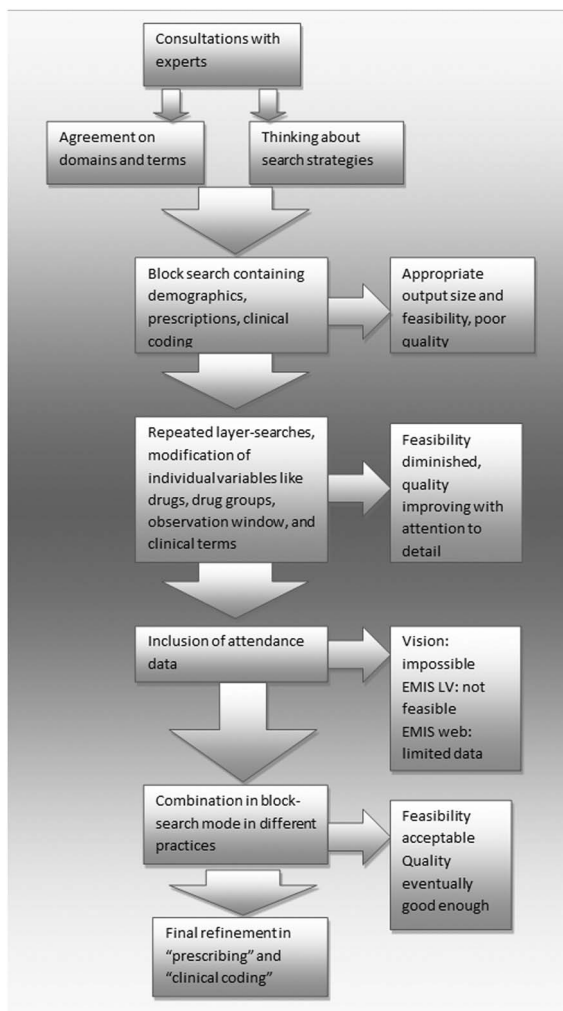
**Figure 2** The testing process



**Figure 3** The final search strategy

terms were tailored to the 'coding cultures' of the practices. We did not use a universal uniform search but applied the search protocol to reflect the practice culture. The searches worked well but there were some outliers (Table 2, Practices 'SS', 'OM', 'CB'): note the number of patients excluded by the GPs. Several issues accounted for the unusually high number of patients screened out by the GPs. Exclusion criteria for the trial were: being housebound, not fluent in English and not having any more serious other co-morbidity than their pain that would make the person unsuitable to attend a pain management course. Reasons for the high rejection rate were large numbers of the patient register did not have English as a first language (SS), a proportion of the GPs in the practice did not screen the lists (OM, so we had to reject all their patients) and practice 'CB' had a predominantly elderly patient population in nursing homes. Screening the output list of patient names by somebody familiar with the practice population was useful to understand the particular circumstances of some practice populations.

Patients identified from the electronic searches were invited by letter to participate in the study, if they were interested they could contact the trial administration team via an expression-of-interest-form or directly via telephone. During the trial recruitment phase we kept records of conversations from patients ringing up to find out more

**Stage 5: Implementing the search strategy**

At this point, we felt that the search was good enough to be employed for the trial recruitment. From a population of 229 143 people, 8494 were sent invitations via their GPs to participate in a trial, this equates to 3.7%: we had originally estimated 5%. We eventually enrolled 703 people with chronic pain into the trial. Table 2 shows the results from 20 general practices in inner city London and rural Warwickshire. In line with the findings from the consultation phase, a flexible/ bespoke search strategy was employed. The search
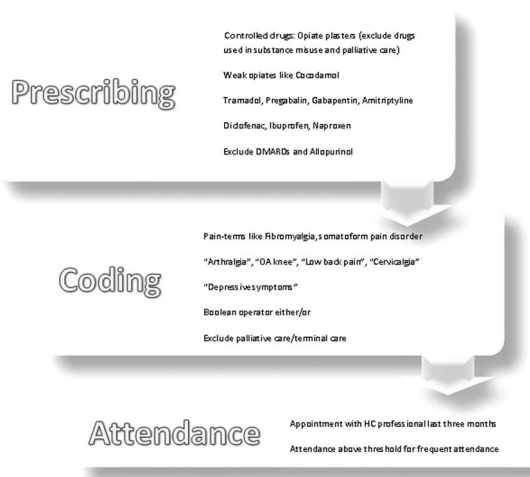
about the trial, this gave us some information about the appropriateness of our searches and the type of people we invited to the study. In London, we covered 10 practices, we had 640 phone calls of which 20 (3%) were from patients informing us that they either: no longer had pain, they had never had pain or that they had only just started to have pain. Patients who sent the study team an 'expression of interest' form were telephoned to assess their suitability to be in the trial. We had 673 expression of interest forms, of these 25 (3.7%) were from people who: no longer had pain, never had pain or that they had only just started to have pain. We recruited patients from the remaining interested patients who rang or sent in expression of interest forms who were suitable and eligible for the study.

## Discussion

The aim of this study was to develop a search strategy to identify patients with chronic musculoskeletal pain in different primary care practice electronic patient record systems. Despite considerable differences in data organisation and software systems, we generated a search strategy protocol, which identified the population of interest, but allowed flexibility in its application to cater for the different coding cultures that exist between general practices. We cannot say that this is the most accurate or inclusive way of identifying chronic pain patients, as we have no definitive data to support such a claim, but we can say that this search strategy yielded the type of patients that we were interested in.

We found that the quickest way of identifying chronic pain patients was to search for repeat prescriptions for specific drugs commonly prescribed in the practice for chronic pain. Local knowledge of the population and commonly prescribed drugs were the key to selecting the index drugs on repeat prescriptions. A search strategy in an inner city practice with high fluctuation of staff and registered patients may not allow this simplicity.

There were some disadvantages of using repeat prescriptions alone to identify chronic pain patients. Some drugs generated a bias towards the older population whose pain was chronic but may not necessarily be persistently bothersome

and/or intrusive. Drug-orientated searches may miss people with pain who do not have analgesic drugs on repeat and those who prefer not to take medication for their pain. There are also clinicians who avoid putting medication on the repeat prescription system. It is therefore important to understand the prescribing regimens of the practices in which the searches are being conducted. Using repeat prescriptions of analgesics has been used in other studies and was also thought to yield the highest likelihood of reaching the appropriate patient group (Smith *et al.*, 2005; McCracken and Velleman, 2010).

Our brief survey of commonly used terms for musculoskeletal pain among clinicians confirmed the heterogeneity in coding practice. Our testing processes showed that effective searching in the clinical coding domain is dependent on local coding practices or 'coding cultures' and software capabilities. Basing a search for chronic pain, patients on morbidity coding alone was not feasible due to the variety of index codes and that the condition coded is often regarded as 'non-specific' (Khan *et al.*, 2010).

Increased healthcare seeking behaviour by those with chronic pain is not unusual (Boersma and Linton, 2005). We therefore chose the frequency of attendance as proxy-marker for the need-state of patients (Smith *et al.*, 2001; Ciechanowski *et al.*, 2003; Smith *et al.*, 2004). Extracting and sorting attendance frequencies per patient was difficult and these data alone were not accurate enough to reflect the presumed need-state of patients. Attendance data were confounded by the presence of QOF – financially incentivised care for specific conditions (Department of Health, 2004). These conditions or diseases may co-exist with chronic pain and consultations may therefore be initiated by the provider and do not necessarily represent persisting pain symptoms (Love and Burton, 2005). Repeated consultations in primary care can also reflect an organised approach to accommodate complex diagnoses in coordinated structured consultations (Dunn and Croft, 2006). We did not find attendance data easy to retrieve and concluded that using this as a primary search strategy was not always practical but where possible useful.

The strength of the study was that it looked at a search strategy, which could identify patients

**Table 2** Search results from primary care centres

| Practice | Location | Practice list size | Estimated chronic pain patients[a] | Number identified by search | Excluded by GPs |
|---|---|---|---|---|---|
| WL | London | 7059 | 265 | 143 | 36 |
| T | London | 3496 | 131 | 143 | 33 |
| CSt | London | 12 181 | 457 | 312 | 74 |
| SG | London | 8107 | 304 | 281 | 41 |
| N | London | 8300 | 311 | 350 | 10 |
| IH | London | 12 051 | 452 | 309 | 87 |
| EA | London | 5700 | 214 | 300 | 37 |
| EB | London | 2900 | 109 | 175 | 25 |
| SS | London | 12 190 | 457 | 602 | 439 |
| SL | London | 14 000 | 525 | 546 | 34 |
| MC | Warwick | 9200 | 345 | 404 | 37 |
| OM | Warwick | 10 500 | 394 | 428 | 254 |
| CA | Warwick | 5500 | 206 | 291 | 11 |
| S | Warwick | 10 000 | 375 | 375 | 44 |
| CB | Warwick | 12 600 | 473 | 568 | 285 |
| CC | Warwick | 10 878 | 408 | 372 | 104 |
| TF | Coventry | 14 147 | 531 | 390 | 30 |

[a] We estimate that 5% of adults (~75% of list size) will have chronic pain.

with chronic pain in the absence of generally agreed and ubiquitously applied classification term. Unlike conditions and treatments incentivised by payments, there were no coding policies for chronic pain and data entry was not standardised or influenced by political pressure to implement coding policies. We used the information organised and coded by GPs in their clinical notes. The themes derived were used to inform the search, leading to a practice-orientated search strategy.

Reliability and validity as quality standards for diagnostic tests were not applicable to this study due to the constantly changing database and the absence of a fixed external gold standard as comparator. It is unknown how many patients are there with chronic pain in a practice population. In an ideal world, we would have conducted the searches in a controlled environment and compared the information in the clinical records with those that were selected and those that were not. This was a pragmatic study in a clinical setting; we worked with clinicians and practice managers with limited time and information technology skills. We decided that feasibility, practicality and reasonable representativeness were sufficient for our purposes to identify patients for a large RCT (ISRCTN: 24426731). We applied the search strategy in 20 general practices and it enabled the trial team to successfully identify and recruit 703 participants.

This study demonstrated the challenges of identifying a population of people with poorly defined conditions in a database. Smith *et al.* described a similar task: they targeted people with 'medically unexplained symptoms' who were repeatedly referred into secondary care. Their strategy also included a combination of database searches and manual case note reviews. The quality of their search outcome was described as 'reasonably accurate' (Smith *et al.*, 2009). Chronic musculoskeletal pain was our index condition, but our approach could be adapted for other conditions such as 'medically unexplained symptoms' (Barsky *et al.*, 2005; Bensing and Verhaak, 2006; Fink and Rosendal, 2008).

Our key findings were: recognising the existence of local coding and prescribing cultures and the tension between the primary purpose of the data we were searching (managing patient information for clinical use) and its secondary use for things like audit, research or billing (Greenhalgh *et al.*, 2009). Standardisation of data entry makes searches using shared databases possible however data must be regarded with due caution and the context within which it is collected.

We recognise that more work is needed to test overall validity and reliability of our search approach in different settings and with different populations to see whether the principles of searching identified in this study can be generalised to other research projects.

## Conclusion

Overall, in the absence of a definitive list of chronic pain patients, a reasonably representative search conducted electronically can only ever be an estimate of the pain population. Our work shows that chronic pain is coded in various ways in primary care records. Knowledge of 'coding cultures' and the practices of clinical staff including how they prescribe and document these processes is essential for successful and representative search data.

## Acknowledgements

## References

**Armstrong, D.** 2011: Diagnosis and nosology in primary care. *Social Science & Medicine* 73, 801–7.

**Barsky, A.J., Orav, E.J.** and **Bates, D.W.** 2005: Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Archives of General Psychiatry* 62, 903–10.

**Bensing, J.M.** and **Verhaak, P.F.M.** 2006: Somatisation: a joint responsibility of doctor and patient. *Lancet* 367, 452–54.

**Bentley, T., Price, C.** and **Brown, P.** 1996: Structural and lexical features of successive versions of the Read Codes. In Teasdale, S.E., editor, *Annual Conference of the Primary Healthcare Specialist Group of the British Computer Society.* Cambridge, Worcester: PHCSG, 91–103.

**Boersma, K.** and **Linton, S.J.** 2005: Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *The Clinical Journal of Pain* 21, 38–43.

**Carnes, D., Taylor, S.J., Homer, K., Eldridge, S., Bremner, S., Pincus, T. Rahman, A.** and **Underwood, M.** 2013: Effectiveness and cost-effectiveness of a novel, group self-management course for adults with chronic musculoskeletal pain: study protocol for a multicentre, randomised controlled trial (COPERS). *BMJ Open* 3.

**Ciechanowski, P., Sullivan, M., Jensen, M., Romano, J.** and **Summers, H.** 2003: The relationship of attachment style to depression, catastrophizing and health care utilization in patients with chronic pain. *Pain* 104, 627–37.

**Department of Health.** 2004: QOF Quality and Outcomes Framework. London: Department of Health.

**Dunn, K.M.** and **Croft, P.R.** 2006: Repeat assessment improves the prediction of prognosis in patients with low back pain in primary care. *Pain* 126, 10–15.

**Egton Medical Information Systems Limited.** 1987–2013: *EMIS.* Leeds: Egton Medical Information Systems Limited.

**Fink, P.** and **Rosendal, M.** 2008: Recent developments in the understanding and management of functional somatic symptoms in primary care. *Current Opinion Psychiatry* 21, 182–88.

**Greenhalgh, T., Potts, H.W.W., Wong, G., Bark, P.** and **Swinglehurst, D.** 2009: Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method. *Milbank Quarterly* 87, 729–88.

**Henningsen, P.** and **Creed, F.** 2009: Medically unexplained symptoms and somatisation. In EACLPP Medically Unexplained Symptoms Study Group, editor, Noordwijkerhout: The European Association for Consultation Liaison Psychiatry and Psychosomatics.

**In Practice Systems Ltd.** 1994–2013: *Vision,* third edition. London: In Practice Systems Ltd.

**Jordan, K., Clarke, A.M., Symmons, D.P.M., Fleming, D., Porcheret, M., Kadam, U.T.** and **Croft, P.** 2007: Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. *British Journal of General Practice* 57, 7–14.

**Khan, N.F., Harrison, S.E.** and **Rose, P.W.** 2010: Validity of diagnostic coding within the General Practice Research Database: a systematic review. *British Journal of General Practice* 60, e128–136.

**Love, T.** and **Burton, C.** 2005: General practice as a complex system: a novel analysis of consultation data. *Family Practice* 22, 347–52.

**McCracken, L.M.** and **Velleman, S.C.** 2010: Psychological flexibility in adults with chronic pain: a study of acceptance, mindfulness, and values-based action in primary care. *Pain* 148, 141–47.

**Merskey, H.** and **Bogduk, N.** (editors). 1994: *Classifications of chronic pain – descriptions of chronic pain syndromes and definition of pain terms.* Seattle: IASP Press.

**NHS.** 2000: The Clinical Terms Project Version 3 (The Read Codes). In NHS Information Authority, editor, Birmingham: NHS Executive.

**Nordin, H., Eisemann, M.** and **Richter, J.** 2006: The accuracy of the DSM-IV pain disorder and the ICD-10 persistent somatoform pain disorder in chronic pain patients. *Journal of Clinical Psychology in Medical Settings* 13, 303–10.

**Read, J., Sanderson, H.** and **Drennan, Y.** 1995: Terming, encoding, and grouping. *Medinfo* 8, 56–59.

**Smith, B.H., Elliott, A.M.** and **Hannaford, P.C.** 2004: Is chronic pain a distinct diagnosis in primary care? Evidence arising from the Royal College of General Practitioners' Oral Contraception study. *Family Practice* 21, 66–74.

**Smith, B.H., Penny, K.I., Elliott, A.M., Chambers, W.A.** and **Smith, W.C.** 2001: The level of expressed need – a measure of help-seeking behaviour for chronic pain in the community. *European Journal of Pain* 5, 257–66.

**Smith, B.H., Hannaford, P.C., Elliott, A.M., Smith, W.C.** and **Chambers, W.A.** 2005: The 'number needed to sample' in primary care research. Comparison of two primary care sampling frames for chronic back pain. *Family Practice* 22, 205–14.

**Smith, B.J., Mcgorm, K.J., Weller, D., Burton, C.** and **Sharpe, M.** 2009: The identification in primary care of patients who have been repeatedly referred to hospital for medically unexplained symptoms: a pilot study. *Journal of Psychosomatic Research* 67, 207–11.