Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/89298

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) license and may be reused according to the conditions of the license. For more details see: http://creativecommons.org/licenses/by-nc-nd/4.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
Development of a repository of individual participant data from randomized controlled trials of therapists delivered interventions for low back pain

S.W. Hee1, M. Dritsaki2, A. Willis1, M. Underwood1, S. Patel1

1 Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK
2 Oxford Clinical Trials Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK

Abstract

Background: Individual patient data (IPD) meta-analysis of existing randomized controlled trials (RCTs) is a promising approach to achieving sufficient statistical power to identify sub-groups. We created a repository of IPD from multiple low back pain (LBP) RCTs to facilitate a study of treatment moderators. Due to sparse heterogeneous data, the repository needed to be robust and flexible to accommodate millions of data points prior to any subsequent analysis.

Methods: We systematically identified RCTs of therapist delivered intervention for inclusion to the repository. Some were obtained through project publicity. We requested both individual items and aggregate scores of all baseline characteristics and outcomes for all available time points. The repository is made up of a hybrid database: entity-attribute-value and relational database which is capable of storing sparse heterogeneous datasets. We developed a bespoke software program to extract, transform and upload the shared data.

Results: There were 20 datasets with more than 3 million data points from 9328 participants. All trials collected covariates and outcomes data at baseline and follow-ups. The bespoke standardized repository is flexible to accommodate millions of data points without compromising data integrity. Data are easily retrieved for analysis using standard statistical programs.

Conclusions: The bespoke hybrid repository is complex to implement and to query but its flexibility in supporting datasets with varying sets of responses and outcomes with different data types is a worthy trade off. The large standardized LBP dataset is also an important resource useable by other LBP researchers.

Significance: A flexible adaptive database for pain studies that can easily be expanded for future researchers to map, transform and upload their data in a safe and secure environment. The data are standardized and harmonized which will facilitate future requests from other researchers for secondary analyses.

1. Introduction

Globally, low back pain (LBP) is one of the leading causes of years lived with disability, and in developed countries it is the leading contributor to the burden of disability adjusted life years in young adults (Hoy et al., 2014; Global Burden of Disease Study 2013 Collaborators, 2015). Therapist-delivered interventions—non-drug, non-surgical approaches—to the treatment...
of LBP are widely used. There is good evidence that several therapist-delivered treatment approaches are effective. There is also evidence that some of these treatments are cost-effective, e.g. offering a course of manual therapy, including spinal manipulation, comprising up to nine sessions over a period of up to 12 weeks (Savigny et al., 2009). The average effect size of therapists delivered interventions for LBP is typically modest.

One approach for improving outcomes is to identify treatment moderators, baseline characteristics, that predict the greatest benefits or least effectiveness from an intervention for an individual with LBP. To test for a modest interaction between a moderator and a treatment, a randomized controlled trial (RCT) needs at least 503 participants (see, (Gurung et al., 2015) for a full description of this power calculation). Mistry and colleagues reported that most RCTs are too small to reliably identify subgroups (Mistry et al., 2014). Therefore, many of these sub-group analyses were severely underpowered. These data are not suitable for meta-analysis for treatment moderation. Individual participant data (IPD) meta-analysis of RCTs will allow meta-analysis of potential treatment moderators and has the potential to provide adequate statistical power to identify sub-groups who may benefit most from particular treatment options. To achieve this, data from relevant trials need to be assembled and merged into a single useable dataset. Ideally, the dataset structure should also allow additional trial datasets to be added as they become available to facilitate future research.

Clinical trial datasets can be stored in a flat file tabular format such as Microsoft Excel, which typically uses rows to represent a participant record and columns to represent variables captured on case report forms (CRF). Tabular formats are useful for small datasets and have the advantage of being intuitive, relatively simple to create and machine-readable. However, they can be susceptible to excessive growth with each patient record requiring a new row to be inserted and an additional column for every variable. Large numbers of columns can quickly accumulate when clinical and non-clinical items are measured across multiple time points.

Relational databases provide a more robust and efficient solution for larger datasets. This model allows data to be stored and connected in individual tables. Repeating data groups can be separated into their own table and joined back to the main domain using relationships, thus reducing the data redundancy problem associated with flat files. The rules governing a relational database are specified in a schema that can be complex and time consuming to design. The repository relies on data from multiple RCTs and is frequently altered to accommodate new discoveries requiring a more flexible solution. We describe the process of collating data from multiple RCTs. We also describe the process of developing a hybrid database that is flexible and robust for storing multiple datasets to facilitate current analyses and the addition of future datasets.

2. Methods

2.1 Collation of trials

We systematically identified all RCTs of therapist delivered treatments for LBP up until September 2011. We have described this process in detail elsewhere (Mistry et al., 2014; Gurung et al., 2015). We searched this dataset to identify unique trials with >179 participants. We started with an original lower limit of 200 for the sample size. Allowing for some loss to follow-up, a trial of 200 participants will have 90% statistical power to identify a standardized mean difference of 0.5 between two treatment groups. Any individual trials smaller than this are likely to be seriously underpowered for their primary outcome. Upon screening the trials there were many that obtained a final sample size of just under 200; typically these were studies aiming for around 200 participants that fell short of the final target. We therefore revised our inclusion criteria to >179 participants.

From a practical perspective of approaching trial investigators, our inclusion criteria yielded a manageable number of trials to approach; large trials (those of thousands of participants) and small trials (<100 participants) each create a similar amount of work to collate. We also obtained data from trials that were not on our original list, as investigators became aware of our project. Although these trials had smaller sample sizes than our target studies, we decided to include them to add power to our analysis. As the primary purpose of this project was to identify sub-groups rather than main treatment effects, the omission of some small trials from our dataset is unlikely to have materially affected our conclusions.

Between 2011 and 2012, each investigator was invited to participate and share their data with us to the standardized LBP repository via email. If a response was not received within 6–8 weeks, a reminder email was sent. For those interested a
personalized data sharing agreement was created and sent to the investigator to review and sign. Once the signed document was received by the University, the investigator was provided with details on how to securely send their data to us. We used the University of Warwick secure file transfer service. Investigators were advised that any datasets sent to us needed to be anonymized and encrypted.

On receipt of the data, the statistician (SWH) and health economist (MD) queried them before mapping and transforming the original data to the standard for the repository. Details of the mapping and transformation procedures are described below (Section 2.2).

Data integrity is vital to the repository. To check that the mapping and transformation procedures were done correctly, the repository data were routinely checked against the original datasets. To achieve this, at each time point (baseline and all follow-ups), a random sample of data was extracted and manually cross checked against the source data. Any inconsistencies were flagged and, if required, the mapping and transformation instructions were amended. This process was repeated until the data were deemed to have been transformed correctly, i.e. zero error.

### 2.2 Bespoke database

Our bespoke database is a hybrid of an entity-attribute-value (EAV) model and a fixed schema relational model. This design is commonly used in clinical trial database management systems (Brandt et al., 2002) as it provides the flexibility of storing sparse heterogeneous data while enforcing high data integrity. A detailed technical description of the system architecture and our approach to data transfer can be found in the Supporting Information (Method S1).

The basic components of the database can be broken down into four main tables. Fig. 1 shows a simplification of three of these tables. The ‘Subject’ table stores the participant’s original identifier (ID) and a unique identifier generated by the system (Fig. 1C). The ‘Object’ table stores a reference to the ‘Subject’ table, a unique object identifier, and a key value representing either a single CRF or a group of repeating questions that have been separated into a child object and then re-joined to a parent using a parent/child relationship (Fig. 1D). This approach to storing repeating data groups is an interpretation of the EAV with classes and relationships (EAV/CR) (Nadkarni et al., 1999).

The ‘Attribute’ table is used to simply store a list of all of the repository’s variables. The ‘EAV’ table (Entity, Attribute and Value) stores references to the related ‘Object’ and ‘Attribute’ tables, as well as the actual value (Fig. 1E). Thus, a complete record can typically be recreated by selecting a collection of rows from the EAV table that have the same object identifier.

A simplification of how two original tabular source datasets are stored in the EAV model is shown in Fig. 1. In Fig. 1A, Trial A collected gender and the EuroQol five dimensions questionnaire (EQ5D) (EuroQol Group, 1990) and EuroQol visual analogue scale (EQ-VAS) of health state at baseline (labelled, EQ1, EQ2, EQ3, EQ4, EQ5 and EQ6, respectively) and at the first follow-up (labelled, EQ1_1, EQ2_1, EQ3_1, EQ4_1, EQ5_1 and EQ6_1), where the first follow-up was 4 weeks post-randomization. The coded values 1 and 2 for the variable ‘sex’ represent male and female, respectively. In Fig. 1B, Trial B collected gender, age, and the Roland Morris Disability Questionnaire (RMDQ) (Roland and Morris, 1983) at baseline (labelled, RDQ_0) and 3-month follow-up (labelled, RDQ_3mo).

For each participant, the repository generated a unique ID as seen in Fig. 1C. For each domain occurrence, a row is created in the ‘Object’ table (Fig. 1D). For example, for subject #011 (equivalently, unique ID #1000001), one row was created for Demographics and two rows for EQ5D. The demographic data were only recorded once, at baseline per participant, hence, only one row was required. As the EQ5D and EQ-VAS were collected at both baseline and at 4-week follow-up, a row was required for each time point.

A row is then created for each populated cell in the ‘EAV’ table (Fig. 1E). One row is also created for each time point that the item was collected at. Rows are only created for populated cells from the original data source. For example, five EQ5D rows and five rows for their corresponding follow-up were created for subject #011 to capture the values recorded at baseline. Note that the labels for all of the variables (attributes) in the ‘EAV’ table follow the repository standard. In addition, original values are transformed, if necessary, to the repository standard. In the standardized repository, male and female values are represented numerically by codes 1 and 2, respectively. In our example, the values ‘M’ and ‘F’ for the variable sex from trial B were transformed to 1 and 2, respectively, in the ‘EAV’ table. For mapping and transformation details, please refer to Supporting Information Method S1 Section 2.2.2.
Figure 1  Sample of tabular clinical data in an EAV table. (A) and (B) Examples of original clinical data from two trials in a tabular format, (C) the ‘Subject’ table with a new unique ID for each participant, (D) the ‘Object’ table with an instance of a domain per participant for every derived tabular record, and (E) the ‘EAV’ table with a row for each populated cell and a row for the follow-up time point where applicable.
2.3 Healthcare resource-use dataset

The mapping and transformation of healthcare resource-use data were more challenging because the different types of resources used across RCTs did not conform to any standard. However, each question and answer in a typical healthcare resource-use questionnaire could be broken into eight parts: the recall period (e.g. 3-month follow-up), the type of resource (e.g. visit to physiotherapist), the reason for using the resource (e.g. LBP), the location of the resource (e.g. community healthcare centre), the unit of measurement (e.g. home visit), the quantity, the cost or expenses incurred, and the payer (e.g. national health care system).

Fig. 2 shows a simplified version of a typical healthcare resource-use questionnaire (Patel et al., 2016). In this example, participants were asked to record all of the healthcare resources they had used at the 3-month follow-up time point (Fig. 2A). The answers provided were stored in a tabular format that used 12 columns to capture all responses to the five questions (Fig. 2B). By using this format, the number of columns required to accommodate the data would grow in line with the maximum number of responses provided by any individual. Fig. 2C shows a view of the repository healthcare resources data generated and pivoted from the EAV/CR model. This view displays the eight standard repository healthcare resource-use attributes and an additional attribute called ‘Text’ that is used to capture any comments that were written on the CRFs.

The process for creating the transformed healthcare resource-use data involves splitting the original questions into a number of derived parts that will map to the standard attributes. In our example, question 1 asked how many times the participant had consulted their primary care doctor for any reason in the last 3 months. Thus, the recall period was set to ‘13’ (because the repository standard stores the time point in unit weeks), the type of resource was ‘GP’, the reason for using the resource was ‘Any condition’, the location of the resource was ‘Primary Care Clinic’, the unit of measurement was ‘Visit’, and the payer was ‘Public Health Service’. All of these values were derived solely from the information contained in the original question, as opposed to the value of the variable. Only the attribute ‘quantity’ was directly mapped to the original variable’s value. For question 3, both quantity and cost were directly mapped to the original variables’ values. Mapping and transformation details for healthcare resource-use data are presented in Supporting Information Method S1 Section 2.2.3.

2.4 Using the repository data

Data in the EAV/CR format are not suitable for analysis because of their fragmented structure. Thus, data from the same domain are pieced together and pivoted so that the dataset for each domain resembles a long format tabular structure. Technical details of the extraction, transformation, loading and outputting the repository data are in Supporting Information Method S1 Sections 2.2.4 and 2.2.5. For example, suppose we would like to view and analyse demographics data from our simplistic example in Fig. 1. We can extract data from Object ID #1, #4, #7, #10 and #12 (Fig. 1D) and their corresponding data from tables ‘Subject’ (Fig. 1C) and ‘EAV’ (Fig. 1E), joined by their unique subject ID and object ID. The extracted data are presented in a tabular format for analysis, as seen in Fig. 3A.

Fig. 3B and C show the long format tabular structure for data captured at various time points for EQ5D and RMDQ, respectively. End users would need to combine data from different domains, e.g. demographics and RMDQ for analysis to investigate potential variables that moderate RMDQ outcomes. Note that the LBP repository database is more detailed than the one we present here for illustration. In the repository information such as the name of the trial is stored in a fixed table (see, Supporting information Fig. S1) and is pieced together with the domain data into a long tabular structure as seen in Fig. 3.

3. Results

3.1 Identification of trials

Our initial search yielded 658 hits. After exclusions, we identified 42 unique trials datasets that met our entry criteria. These authors were contacted for their data. We obtained no response from 15 corresponding authors, six datasets were no longer available and for seven datasets we were unable to conclude negotiations prior to freezing our database. We were able to satisfactorily import data from 14 of these trials. We also included data from five smaller trials offered to us by researchers who were aware of this project (Fig. 4). The final dataset included data from 20 datasets (one trial had a feasibility study prior to the main trial) and a total of 9328 participants (Patel et al., 2016).
### A Sample of questions and answers at 3-month follow-up.

**Public healthcare professionals**

<table>
<thead>
<tr>
<th>No. of visits</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pri1</td>
<td>1. In the last 3 months, how many times have you consulted your doctor or another doctor for any reason?</td>
</tr>
<tr>
<td>Pri2</td>
<td>2. In the last 3 months, how many times have you consulted the NHS physiotherapist for low back pain?</td>
</tr>
</tbody>
</table>

**Private healthcare professionals**

<table>
<thead>
<tr>
<th>No. of visits</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>nPriv1</td>
<td>3. In the last 3 months, how many times have you consulted any private physiotherapist for low back pain?</td>
</tr>
<tr>
<td>cPriv1</td>
<td></td>
</tr>
</tbody>
</table>

**Medication**

4. In the last 3 months, list the medicines prescribed by your doctor for your low back pain.

<table>
<thead>
<tr>
<th>Medicine prescribed</th>
<th>No. of prescriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. pmel1</td>
<td>nmed1</td>
</tr>
<tr>
<td>ii. pmel2</td>
<td>nmed2</td>
</tr>
</tbody>
</table>

5. In the last 3 months, list the medicines or treatments that you personally bought over-the-counter to treat your low back pain.

<table>
<thead>
<tr>
<th>Aid/medicine bought</th>
<th>Cost (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. bmed1</td>
<td>cmed1</td>
</tr>
<tr>
<td>ii. bmed2</td>
<td>cmed2</td>
</tr>
</tbody>
</table>

### B Sample of data in a tabular database.

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>Pri1</th>
<th>Pri2</th>
<th>nPriv1</th>
<th>cPriv1</th>
<th>pmel1</th>
<th>nmed1</th>
<th>pmel2</th>
<th>nmed2</th>
<th>bmed1</th>
<th>cmed1</th>
<th>bmed2</th>
<th>cmed2</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>40</td>
<td>Ibuprofen</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Mattress</td>
<td>325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>3</td>
<td>120</td>
<td>120</td>
<td></td>
<td>Mattress</td>
<td>299</td>
<td>Walking stick</td>
<td>5.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### C The view of the healthcare resource-use data.

<table>
<thead>
<tr>
<th>ID</th>
<th>FU</th>
<th>RP</th>
<th>Type</th>
<th>Text</th>
<th>Reason</th>
<th>Location</th>
<th>Unit</th>
<th>Quantity</th>
<th>Cost</th>
<th>Payer</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>13</td>
<td>13</td>
<td>GP</td>
<td>Any</td>
<td>PRI</td>
<td>Visit</td>
<td>1</td>
<td>PHS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>13</td>
<td>13</td>
<td>Physio</td>
<td>LBP</td>
<td>COMM</td>
<td>Visit</td>
<td>1</td>
<td>40</td>
<td>IND</td>
<td>PHS</td>
</tr>
<tr>
<td>100</td>
<td>13</td>
<td>13</td>
<td>NSAIID</td>
<td>Ibuprofen</td>
<td>LBP</td>
<td>N/A Px</td>
<td>1</td>
<td>40</td>
<td>IND</td>
<td>PHS</td>
</tr>
<tr>
<td>101</td>
<td>13</td>
<td>13</td>
<td>GP</td>
<td>Any</td>
<td>PRI</td>
<td>Visit</td>
<td>0</td>
<td>PHS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>13</td>
<td>13</td>
<td>Physio</td>
<td>LBP</td>
<td>COMM</td>
<td>Visit</td>
<td>5</td>
<td>PHS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>13</td>
<td>13</td>
<td>Physio</td>
<td>LBP</td>
<td>PTE</td>
<td>Visit</td>
<td>0</td>
<td>IND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>13</td>
<td>13</td>
<td>Aid</td>
<td>Mattress</td>
<td>LBP</td>
<td>Item</td>
<td>325</td>
<td>IND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>13</td>
<td>13</td>
<td>Physio</td>
<td>LBP</td>
<td>PTE</td>
<td>Visit</td>
<td>120</td>
<td>IND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>13</td>
<td>13</td>
<td>Aid</td>
<td>Mattress</td>
<td>LBP</td>
<td>Item</td>
<td>299</td>
<td>IND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>13</td>
<td>13</td>
<td>Aid</td>
<td>Walking stick</td>
<td>LBP</td>
<td>Item</td>
<td>5.65</td>
<td>IND</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2** Sample healthcare resource-use data. (A) A simplified healthcare resource-use questionnaire. (B) Sample of healthcare resource-use data in a tabular format. (C) The view of the original source data generated and pivoted from the EAV/CR table. FU, follow-up; RP, recall period; GP, primary care doctor; Any, any reason; PRI, primary care clinic; PHS, public health service; Physio, physiotherapist; LBP, low back pain; COMM, community clinic; PTE, private clinic; IND, individual; NSAIID, non-steroidal anti-inflammatory drugs; Px, prescription; Aid, aids and adaptations.
The characteristics of the included studies are summarized in Supporting information Table S1. Initial examination of the data showed that no two trials studied identical interventions. Even the usual care arms of the included studies are likely to differ according to jurisdiction, site of recruitment and age of the study. To make meaningful comparisons we needed to broadly pool the interventions into groups for our analyses.

Considering the potential mechanisms through which the participant characteristics might affect the outcome, we decided to pool interventions that might under other circumstances appear rather heterogeneous. In particular, the decision to include several superficially different interventions as ‘passive physiotherapy’ may surprise some readers. Our view, however, is that these are very distinctly different from active exercise based interventions, or those working primarily through a psychological approach. Essentially they all consist of an assessment, whatever reassurance and education is provided as part of the treatment session, plus whatever modality is being offered; be it massage/mobilization/manipulation or needling. We consider these to be conceptually sufficiently close in their mode of action that it is unlikely there will be distinctions in how the potential moderators included in our analyses might affect the outcomes. They are, however, distinctly different from active physical and psychological interventions in how treatment moderation might operate. The American Pain Society/American College of Physicians guidelines use a broadly similar approach to group non-pharmacological interventions (Chou and Huffman, 2007).

We first identified the control interventions and classified them as either usual care or sham control. We then split our active interventions into three broad categories; active physical (exercise and graded activity), passive physical (individual physiotherapy, manipulation and acupuncture) and psychological (advice/education and psychological therapy) (Patel et al., 2016). Fig 5 shows a network of the treatment subtypes from trials in the repository.

### 3.2 Repository database

Data in the repository were captured in the same granularity as the original data provided to us.
Where available, we included both the individual items and the summary score for each questionnaire (domain) at all possible time points when they were collected. This proves to be helpful when only certain items of a questionnaire may be of interest to future researchers. For example, the visual analogue scale (VAS) is usually a standalone instrument that asks the participant to describe their average or worst pain at the present time. The participant marks their pain level on the VAS line, which is visually presented as either a horizontal or vertical line with ‘no pain’ at one end and ‘worst possible pain’ at the other end. Similar questions are found in the chronic pain grade scale (CPG) which has two dimensions and one of them is the pain intensity scores (von Korff et al., 1992). Three items contribute to the pain intensity score (von Korff et al., 1992). Three items contribute to the pain intensity score. Each RCT is unique and no single questionnaire was used by all trials. Thus, future researchers will still be able to pool data from the VAS and the equivalent items from CPG pain intensity items if they were interested to analyse the effect of treatment on pain.

Table 1 presents the list of patient reported outcome measurements stored in the repository alongside demographics and medical history collected at baseline, and healthcare resource-use data at subsequent follow-ups.

All of the trials were able to provide information on sex and age. Other demographics information, such as ethnicity, smoking status, employment status and body mass index (BMI), were given by some trials but were not routinely collected by others (see, Table S2 in the Supporting information for a comprehensive summary). Seven RCTs across three countries provided healthcare resource-use data (Patel et al., 2016).

Other data fields in the repository included a variety of established patient reported outcomes on pain related physical disability (CPG disability, FFBHR, ODI, PDI, PSFS, RMDQ and Troublesome), pain (CPG pain intensity and VAS), and health related quality of life (SF-12/36 and EQ5D3L) (see Table 1 for list of abbreviations). Of note was there was no common instrument that was used by all of the trials (see Table 1).

Supporting information Table S2 shows the response rates for various instruments of interest by treatment group at each possible time point. Most of the RCTs collected outcomes at short- (within 2/3 months post-randomization) and mid-term (6 months post-randomization). Two trials collected outcomes beyond 12 months post-randomization. The follow-up time for trials that collected data 3 months post-randomization or entry to the trial was stored in the repository as 13 weeks. However, one RCT had specifically mentioned in their protocol to collect data at 12 weeks and thus this was stored in the repository, as per protocol (Patel et al., 2016).

4. Discussion

We created a purpose-built repository to store IPD from multiple RCTs and developed the software and procedures to morph heterogeneous datasets. This infrastructure was necessary for our primary objective, identifying treatment moderators for LBP.

The repository and ETL software was intended to provide a solution for standardizing and storing heterogeneous and sparse datasets and output into friendly formatted datasets. Creating the hybrid system based on a relational and EAV model was a relatively simple process. However, creating the bespoke software and standardizing processing solutions was more challenging as it was not possible to predefine classes and attributes until data sharing agreements

Figure 4 Identification of trials.
Figure 5 Summary of all included trials.

Individual physiotherapy (n = 333)
- Cecchi (n = 70)
- Dufour (n = 144)
- HullExProB (n = 119)

Exercise (n = 770)
- UK BEAM (n = 303)
- Dufour (n = 142)
- HullExProB (n = 118)
  *Pengel (n = 65)
  Smeets (n = 53)
  YorkBP (n = 89)
  *(+ sham advice/education)

Advice/Education (n = 219)
- Cecchi (n = 70)
- Macedo (n = 86)
  *Pengel (n = 63)
  *(+ sham electrotherapy)

Manipulation (n = 736)
- UK BEAM (n = 346)
  - Cecchi (n = 70)
  - Hancock (n = 119)
  - Hay (n = 201)

Control (n = 3573)
- Usual/GP Care (n = 2885)
  - UK BEAM (n = 332)
  - BeST (n = 233)
  - Brinkhaus (n = 79)
  - Haake (n = 388)
  - Smeets (n = 51)
  - BIA (n = 121)
  - SC2 (n = 113)
  - Witt (n = 1390)
  - YACBAC (n = 80)
  - YorkBP (n = 98)

Psychological (n = 901)
- BeST (n = 468)
  - Hay (n = 201)
  - BIA (n = 119)
  - SC2 (n = 113)

Sham (n = 688)
- Brinkhaus (acupuncture; n = 73)
- Carlsson (mock TENS; n = 16)
- Haake (acupuncture; n = 387)
- Hancock (electrotherapy; n = 120)
- Kennedy (mock TENS; n = 24)
- Pengel (electrotherapy + education; n = 68)

Combination (n = 451)
- UK BEAM (exercise + manipulation; n = 327)
- Pengel (exercise + education; n = 63)
- Smeets (education + graded activity; n = 61)

Graded Activity (n = 144)
- Macedo (n = 86)
- Smeets (n = 58)

Acupuncture (n = 2201)
- Brinkhaus (n = 146)
- Carlsson (n = 34)
- Haake (n = 387)
- Kennedy (n = 24)
- Witt (n = 1451)
- YACBAC (n = 159)
were in place. Our approach avoided the need to pre-specify all variables of interest and follow-up time points before constructing the database. It also allows us to easily add new datasets in a relatively straightforward manner as they become available.

Seven RCTs across three countries provided healthcare resource-use data. The lack of standardization in the recording of healthcare resource-use items for health economic analyses increased the complexity in harmonizing data from multiple sources. Nevertheless, by splitting each healthcare resource-use item into eight distinct parts, we were able to standardize and harmonize the data from different centres and countries.

Two major strengths of our database are the rigour with which the data were checked and cleaned, and its arrangement, which will allow more trials and variables to be added easily. By including mainly trials with >179 participants, we have included higher quality trials. We have not, however, sought to include all trials on an individual intervention. The database is not currently suitable for comparing the main effects of different treatments. Additional work is needed to identify and upload smaller trials that would contribute to any such analysis and an updated search for relevant studies is also needed.

We were limited by the nature of the available trial data. The heterogeneous nature of the interventions, which were often poorly characterized, the heterogeneous populations studied, and the variable choices in outcome selection will limit the comparisons that can be made using the current database. This in marked contrast to an IPD database of a drug

Table 1 List of patient reported outcome measurements and the number of datasets that collected such information.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>List of measurements</th>
<th>No. of datasets (m = 20)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABPS</td>
<td>Aberdeen Back Pain Scale</td>
<td>2</td>
<td>Ruta et al., (1994)</td>
</tr>
<tr>
<td>ALBP SQ</td>
<td>Acute Low back Pain Screening Questionnaire</td>
<td>3</td>
<td>Linton and Haldén, (1998)</td>
</tr>
<tr>
<td>BBQ</td>
<td>Back Beliefs Questionnaire</td>
<td>2</td>
<td>Symonds et al., (1996)</td>
</tr>
<tr>
<td>BDI</td>
<td>Beck Depression Inventory</td>
<td>1</td>
<td>Beck et al., (1961, 1979)</td>
</tr>
<tr>
<td>CES-D</td>
<td>The Centre for Epidemiologic Studies Depression Scale</td>
<td>1</td>
<td>Radloff, (1977)</td>
</tr>
<tr>
<td>CPG</td>
<td>Chronic Pain Grade</td>
<td>6</td>
<td>von Korff et al., (1992)</td>
</tr>
<tr>
<td>CSQ</td>
<td>Coping Strategies Questionnaire</td>
<td>2</td>
<td>Rosenstiel and Keefe, (1983)</td>
</tr>
<tr>
<td>DASS-21</td>
<td>Depression Anxiety Stress Scales</td>
<td>1</td>
<td>Henry and Crawford, (2005)</td>
</tr>
<tr>
<td>DRAM</td>
<td>Distress and Risk Assessment Method</td>
<td>3</td>
<td>Main et al., (1992)</td>
</tr>
<tr>
<td>EQ-5D-3L</td>
<td>EQ-5D-3L</td>
<td>8</td>
<td>EuroQol Group, (1990)</td>
</tr>
<tr>
<td>FABQ</td>
<td>Fear-Avoidance Beliefs Questionnaire</td>
<td>6</td>
<td>Waddell et al., (1993)</td>
</tr>
<tr>
<td>FFbHR</td>
<td>Hannover Functional Ability Questionnaire for Measuring Back Pain-Related Functional Limitations (Funktionsbeeinträchtigung durch Rückenschmerzen)</td>
<td>3</td>
<td>Kohlmann and Raspe, (1996)</td>
</tr>
<tr>
<td>HADS</td>
<td>Hospital Anxiety and Depression Scale</td>
<td>1</td>
<td>Snaith, (2003)</td>
</tr>
<tr>
<td>IPAQ</td>
<td>International Physical Activity Questionnaire</td>
<td>1</td>
<td>Craig et al., (2003)</td>
</tr>
<tr>
<td>LSI</td>
<td>Lumbar Spine Instability Questionnaire</td>
<td>1</td>
<td>Cook et al., (2006)</td>
</tr>
<tr>
<td>MPOQ-SF</td>
<td>McGill Pain Questionnaire (short form)</td>
<td>2</td>
<td>Melzack, (1987)</td>
</tr>
<tr>
<td>MSPQ</td>
<td>Modified Somatic Perception Questionnaire</td>
<td>4</td>
<td>Main, (1983)</td>
</tr>
<tr>
<td>MZDI</td>
<td>Modified Zung Depression Index</td>
<td>4</td>
<td>Main et al., (1992)</td>
</tr>
<tr>
<td>ODI</td>
<td>Oswestry Disability Index</td>
<td>1</td>
<td>Fairbank et al., (1980)</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Analogue Scale (average/worst)</td>
<td>9</td>
<td>McCracken and Dhingra, (2002)</td>
</tr>
<tr>
<td>PASS-20</td>
<td>Pain Anxiety Symptoms Scale</td>
<td>1</td>
<td>Tait et al., (1990)</td>
</tr>
<tr>
<td>PCI</td>
<td>Pain Disability Index</td>
<td>1</td>
<td>Flor et al., (1993)</td>
</tr>
<tr>
<td>RMDQ</td>
<td>Roland-Morris Disability Questionnaire</td>
<td>15</td>
<td>Roland and Morris, (1983)</td>
</tr>
<tr>
<td>SES</td>
<td>Pain Experience Scale (Schmerzempfindungsskala)</td>
<td>1</td>
<td>Geissner, (1995)</td>
</tr>
<tr>
<td>SF-12</td>
<td>12-item Short Form Health Survey</td>
<td>3</td>
<td>Ware et al., (2002)</td>
</tr>
<tr>
<td>SF-36</td>
<td>36-item Short Form Health Survey</td>
<td>9</td>
<td>Ware et al., (2000)</td>
</tr>
<tr>
<td>TSK</td>
<td>Tampa Scale for Kinesiophobia</td>
<td>5</td>
<td>Vlaeyen et al., (1995)</td>
</tr>
<tr>
<td>Troublesome</td>
<td>Troublesomeness</td>
<td>5</td>
<td>Parsons et al., (2006)</td>
</tr>
<tr>
<td>WHODAS 2.0</td>
<td>World Health Organization Disability Assessment Schedule 2.0</td>
<td>1</td>
<td>Üstün et al., (2010)</td>
</tr>
</tbody>
</table>
intervention used on a well-defined population with a clear outcome, e.g. death. In the long-term, researchers should use intervention taxonomies when reporting their trials to allow for better replication and pooling of trials (Schulz et al., 2010). Notwithstanding these challenges this database is an important step forward in back pain research.

The research community is in favour of having a central repository of data collected for IPD meta-analyses (Tudur Smith et al., 2014). They alluded to the many advantages of standardizing, safeguarding and storing such data centrally. There are other researchers who are attempting to create secure central databases of previously collected IPD but their work is ongoing (Tudur Smith et al., 2011). Our pooled dataset, with standardized and harmonized data, will provide an excellent resource for back pain researchers. Although primarily developed for our own work on the identification of sub-groups, there are many potential uses to which such a database could be used, e.g. secondary analyses. However, any use is dependent on the agreement of the data custodians who donated the original datasets.

Acknowledgements
We acknowledge the Low Back Pain Repository team for their earlier contributions to this work and Dr Jennifer de Beyer for proofreading the manuscript.

Collaborators
Dr Christer Carlsson, Dr Francesca Cecchi, Dr Nina Dufour, Dr Heinz Endres, Dr Mark Hancock, Professor Elaine Hay, Dr Michael Von Korff, Professor Sarah Lamb, Dr Luciana Macedo, Dr Hugh MacPherson, Professor Chris Maher, Professor Suzanne McDonough, Professor Rob Smeehts, Professor David Torgerson and Professor Claudia Witt, who all contributed anonymized individual participant data from randomized controlled trials.

Author contributions
Both SWH and MD made a substantial contribution to the conception of the repository database structure and the XML. AW made a substantial contribution to the conception, design and architecture of the repository database structure, XML and XSD. MU made a substantial contribution to the conception and design of the study and the data acquisition. All authors were involved in the drafting and revision of this manuscript, and gave final approval for this version to be published.

References


Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Summary of the trials in the low back pain standard repository.

Table S2. Comprehensive summary of the number of trials (m) and participants (n) for each outcome by follow-up time point and treatment arm.

Method S1. Technical details of the bespoke hybrid database.

Figure S1. Entity-relationship diagram, depicting the fixed schema with the sub-scheme entity-attribute-value (EAV) tables.

Figure S2. XML instructions to map the original clinical data to the equivalent repository attributes and transform the original values to the repository standard.

Figure S3. XML instructions to map the original health-care resource-use data to the equivalent repository attributes and to transform the original values to the repository standard.

Figure S4. Sample of XML instructions to map more than one time point of the original healthcare resource-use data to the equivalent repository attributes and standard.