Macefield, R., Jacobs, M., Korfage, I., Nicklin, J., Whistance, R., Brookes, S., ... Blazeby, J. (2014). Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). Trials, 15, [49]. https://doi.org/10.1186/1745-6215-15-49
Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs)

Rhiannon C Macefield1*, Marc Jacobs2, Ida J Korfage3, Joanna Nicklin4, Robert N Whistance1, Sara T Brookes1, Mirjam AG Sprangers2 and Jane M Blazeby1,4

Abstract

Background: Synthesis of patient-reported outcome (PRO) data is hindered by the range of available PRO measures (PROMs) composed of multiple scales and single items with differing terminology and content. The use of core outcome sets, an agreed minimum set of outcomes to be measured and reported in all trials of a specific condition, may improve this issue but methods to select core PRO domains from the many available PROMs are lacking. This study examines existing PROMs and describes methods to identify health domains to inform the development of a core outcome set, illustrated with an example.

Methods: Systematic literature searches identified validated PROMs from studies evaluating radical treatment for oesophageal cancer. PROM scale/single item names were recorded verbatim and the frequency of similar names/scales documented. PROM contents (scale components/single items) were examined for conceptual meaning by an expert clinician and methodologist and categorised into health domains. A patient advocate independently checked this categorisation.

Results: Searches identified 21 generic and disease-specific PROMs containing 116 scales and 32 single items with 94 different verbatim names. Identical names for scales were repeatedly used (for example, ‘physical function’ in six different measures) and others were similar (overlapping face validity) although component items were not always comparable. Based on methodological, clinical and patient expertise, 606 individual items were categorised into 32 health domains.

Conclusion: This study outlines a methodology for identifying candidate PRO domains from existing PROMs to inform a core outcome set to use in clinical trials.

Keywords: Core outcome set, Patient reported outcome (PRO), Patient reported outcome measure (PROM), Randomised controlled trial (RCT), Trial methodology, Health domains, Quality of life, Systematic review

Background

Outcome selection and reporting in randomised controlled trials (RCTs) is often problematic. Heterogeneity in outcomes measured across studies in the same disease or treatment may hamper effective evidence synthesis. A systematic review of oesophageal studies, for example, found 10 different measures for postoperative mortality which were often undefined [1]. In addition, selective reporting of outcomes puts trials at risk of outcome reporting bias and can mean treatment effects are exaggerated [2]. These issues may be further complicated for patient reported outcomes (PROs). PROs are typically assessed using questionnaires (patient reported outcome measures (PROMs)) and many validated questionnaires are available because PROMs have been developed by different groups and disciplines (for example, clinical versus psychological) or for differing purposes (for example, measurement of health in generic populations versus disease-specific patient groups). A single PROM...
can be made up of numerous scales and single items and generic and disease specific PROMs are often combined to assess a range of relevant health domains within an RCT. This means that different (and often ill-defined) outcomes may be reported and the multiplicity of items and scales may also allow selection of statistically significant rather than pre-determined a priori PRO endpoints to be reported, increasing the risk of outcome reporting bias. Problems are further accentuated for PROs because terminology of the scales and items across PROMs is not universally agreed meaning data synthesis across studies is difficult when different questionnaires are used, and while there is overlap in the issues that are measured there is also variation because PROMs have been developed by different methods and for different purposes. Potential solutions to these challenges are to develop and use core outcome sets.

Core outcome sets (COSs) are an agreed minimum set of outcome domains to be measured and reported in all trials of a particular treatment or condition [3]. The routine measurement of COSs has the potential to facilitate data synthesis and reduce outcome reporting bias by standardising the outcomes that are measured across studies and this has been emphasised by the COMET (Core Outcome Measures in Effectiveness Trials) initiative which supports the development and application of COSs for pragmatic (effectiveness) trials [4]. Pragmatic trials are designed to assess whether an intervention is effective for routine clinical practice and outcomes, therefore, need to be relevant and important to patients as well as clinicians and other key decision-makers [5]. In many cases these are the outcomes that are assessed with PROMs, particularly if the questionnaire has been developed with patient input [6] but the availability of so many different PROMs, however, means there are problems with selecting which of the measured health domains are ‘core’. The aim of this study, therefore, was to explore and report methods to identify PRO domains from the wealth of available PROMs and to use this approach to inform the development of a COS to use in pragmatic trials in a specific condition. Consensus on which outcomes to include in the final core set, and the methods to achieve this, are the focus of further research.

Methods

This study was undertaken within one disease site and treatment - radical treatment for oesophageal cancer, selected because the research team have clinical and PROM expertise in this area and have previously tried to summarise PRO evidence [7-9]. There were three phases of work: (1) a systematic literature review to identify validated PROMs used in oesophageal cancer studies and the scope of these instruments; (2) a detailed content analysis to explore PROM diversity; and (3) categorisation of PROM content into health domains (Figure 1).

Identification of PROMs used in oesophageal cancer studies

A systematic review was performed to identify and present the scope of existing validated PROMs in order to provide knowledge of the current state of PRO measurement in this field.

Search strategy

Electronic searches in MEDLINE, Embase, PsycINFO and CINAHL databases between January 2006 and May 2011 were performed. The search strategy included terms for patient-reported outcomes, oesophageal cancer, surgery and chemotherapy, radiotherapy or combined therapy (see Additional file 1). Searches were limited to studies published in English language. Relevant studies published prior to 2006 were identified from a previous systematic review [8]. Abstracts of identified records were screened for inclusion and full text articles were assessed for eligibility by one of three reviewers (RW, MJ, RCM) with reasons for exclusion documented. No studies were excluded based on a risk of bias assessment or judgement of methodological quality because the purpose of the current study was to identify PROs rather than examine the quality of the data or treatment effect.

Selection criteria

Included were studies that used at least one validated PROM to evaluate health-related quality of life (HRQL) after radical treatment of oesophageal cancer, including surgical, chemotherapy and/or radiotherapy interventions. Valid PROMs were defined as those that had been tested for psychometric validity and reliability in appropriate patient populations with methodology verified from published papers. No restrictions on study design or sample size were applied. Studies of palliative treatment, comparisons of clinician- or hospital-related factors, and those limited to investigating satisfaction with care or health utilities were excluded.

Data extraction

Data were extracted using a pre-designed form, piloted before full data extraction with a sample of included studies. Study publication date, design and treatment intervention, the name of the PROM(s), the reported PRO scales and single items, and details of any additional non-validated questions were extracted. These were recorded by one reviewer (RM) and checked by additional members of the study team (MJ, MAGS). The validated PROMs were obtained, including other validated disease-specific PROMs known to authors. Verbatim names for the PRO scales and single items as termed by the PROM
developers were extracted and all PROM items (scale components and any single items) were recorded. Data were stored in an electronic database.

Examination of PROM content
A detailed content analysis of the identified instruments was performed to explore the diversity of PROs in this field. Verbatim names for scales and single items were listed. Scales with identical names and others that were similar (defined as having at least one identical word) were documented, counted and compared for consistency and overlap of the component items.

Categorisation into health domains
To synthesise the existing content of instruments and provide a framework for future core set development, all PROM items (scale components and any single items) were examined and systematically categorised into conceptual health domains according to the issue they addressed. This was performed by expert methodologists (an oesophageal cancer surgeon and a psychologist) with experience of questionnaire development in health-related quality of life research and cancer (JMB and MAGS) based on their knowledge, familiarity and practiced skill of grouping questionnaire items in this field. Health domains were defined as generic aspects of quality of life affected by health or disease-specific issues and symptoms [10]. Further domains were defined until saturation, that is, all individual PROM items had been mapped onto a domain. Issues addressed in non-validated questions were additionally mapped to domains to verify that the conceptual health domains encompassed all outcomes measured in the included studies. Mapping of items to domains was checked for completeness and consistency by two authors (IK and RCM) and a patient advocate working within oncology research to maximise validity and reliability of the method. Variances were resolved by discussion within the study team and with the senior author (JMB). Data were recorded electronically.

Results
Identification of PROMs used in oesophageal cancer studies
A total of 1,351 records were screened for inclusion and 111 full-text articles were assessed for eligibility. Of these, 56 were excluded because they did not meet the criteria for eligibility, including seven studies that used PROMs without sufficient psychometric validation. Some 55 relevant articles reporting 56 studies were identified (Table 1) [11-65]. Almost all studies (n = 54, 96%) included data on PROs after surgery, either alone or with neoadjuvant chemo/radiotherapy. Nineteen validated PROMs were used (Table 1) [56,66-83]: nine for gastrointestinal diseases, five cancer-specific instruments and five generic instruments. One oesophageal specific PROM was adapted from a cancer instrument (adapted Rotterdam Symptom Checklist). Three were earlier versions of an updated PROM (EORTC QLQ-C36, QLQ-OES24 and MOS SF20). The most frequently used PROMs were the EORTC...
Table 1 Oesophageal cancer studies (n = 56) using validated PROMs (n = 21)

| Study design     | n  | (%) |
|------------------|----|-----|
| RCT              | 4  | (7) |
| Longitudinal     | 25 | (45)|
| Cross-sectional  | 27 | (48)|

| Publication year | n  | (%) |
|------------------|----|-----|
| <1995            | 1  | (2) |
| 1995-1999        | 6  | (11)|
| 2000-2004        | 12 | (21)|
| 2005-2009        | 26 | (46)|
| 2010-2011        | 11 | (20)|

| Primary outcome   | n  | (%) |
|-------------------|----|-----|
| Clinical measure  | 2  | (4) |
| PRO               | 40 | (7) |
| Undefined/both    | 14 | (25)|

| PROM               | n  | (%) |
|--------------------|----|-----|
| QLQ-C30            | 34 | 15  |
| QLQ-oes18          | 19 | 10  |
| SF36               | 12 | 8   |
| QLQ-oes24          | 8  | 11  |
| GIQL               | 3  | 5   |
| FACT-E             | 3  | 5   |
| Adapted RSLC       | 4  | 4/5a|
| QLQ-C36            | 2  | 18  |
| GERD-HRQL          | 1  | 2   |
| DAUGS32            | 1  | 7   |
| LAGS               | 1  | 3   |
| RSLC               | 1  | 3   |
| Adapted RSLC       | 1  | 5   |
| PNPC               | 1  | 9   |
| WOCS               | 1  | 1   |
| MOS SF20           | 1  | 6   |
| PAIS               | 1  | 7   |
| POMS               | 1  | 6   |
| HADS               | 1  | 2   |
| EQOLb              | 0  | 5   |
| QLQ-OES23b         | 0  | 16  |

*One study used an adapted version with an additional scale [56].
*One study used an adapted version with an additional scale [56].
Sourced from authors’ knowledge; PROM yet to be used in a published study at the time of the conducted search.

**Table 1** Oesophageal cancer studies (n = 56) using validated PROMs (n = 21)

QLQ-C30 (n = 34, 61%), and the disease-specific modules EORTC QLQ-oes18, or earlier version QLQ-oes24 (n = 27, 48%). PROMs were not always used in their entirety, with evidence of selective outcome reporting of scales and single items in 33 (59%), although there was variation across studies in the outcomes that were selected (data not shown). Twenty-one (37%) studies added an additional 74 non-validated items. A further two validated disease specific PROMs; the EORTC QLQ-OG25 [84] and EQOL [85], were sourced from authors’ knowledge, neither of which had been used in a published study since development and validation at the time of the conducted search (May 2011).

### Examination of PROM content

There were 116 scales (composed from 574 individual items) and 32 single items in total, with 94 different verbative scale/item names (Table 2). ‘Pain’ and ‘physical function’ were the most common verbative name for a scale, used in six different PROMs, but other PROM scale names were also very similar (for example, physical wellbeing, physical problems, physical distress, physical activity, role physical) (Table 3). Some scales with identical names, however, had different component items. For example, ‘physical function’ in one PROM consisted of seven items relating to tiredness/fatigue, feeling unwell, waking up at night, changes in appearance, physical strength, endurance and feeling unfit [72], compared to ‘physical function’ in another PROM consisting of five items that referred to strenuous activity, ability to walk certain distances, time spent in bed or a chair, and need for help with self-care [74]. Similar heterogeneity was found for PROs assessed with single items, for example ‘cough’ in one PROM assessed waking at night because of coughing [67], whereas in another it was an assessment of coughing following eating [69]. While the two items assessed slightly different aspects of coughing they had the same name (‘cough’) and thus reporting would only refer to cough and not the actual issue being assessed within the item.

### Categorisation into health domains

All PROM individual items (n = 606) were categorised into 32 conceptual generic or symptom specific domains by the study authors (Table 4). Illustrative examples of this categorisation process are provided for some of the generic health domains (Table 5). The most common assessed health domain (concept), that is, the health domain that most PROM items mapped to, was emotional function, assessed in 18 of the 21 PROMs. Other commonly assessed health domains were ‘pain/pain-related swallowing’ (assessed in 14 different PROMs), ‘physical activity/activities of daily life’ (in 13 PROMs) and ‘appetite/eating/taste’ (in 12 PROMs). Uncommon domains were ‘spiritual issues’ (assessed in one PROM) and ‘dizziness/
dumping’ (assessed in two PROMs). Non-validated questions predominantly focused on eating and therefore were mapped onto the ‘appetite/eating/taste’ domain. A patient advocate checked the categorisation of items into health domains and there were no difference of opinion.

**Discussion**

This study comprehensively analysed PROs from studies in radical treatment for oesophageal cancer. Some 116 scales and 32 single items were identified from 21 validated PROMs. As many as 94 different verbatim names were used to describe PRO scales and single items and although many names were similar, content examination revealed component questions did not always address comparable issues. In-depth examination and categorisation of PROM contents concluded that together they addressed 32 different health domains demonstrating the vast overlap between PROMs.

Our findings show how evidence synthesis of oesophageal cancer PROs may be hampered because of the range of PROMs used in trials and the multiple scales and single items within them, often with inconsistent and non-transparent terminology. Core outcome sets aim to reduce this problem by identifying and prioritising the important health domains to be measured in all studies. The development of core outcome sets in other clinical areas has

| Table 2 Verbatim names of PROM scales and single items |
|--------------------------------------------------------|
| 1 Activities of daily living                          | 33 Extended family relationships | 65 Professional care providers |
| 2 Activity level                                       | 34 Fatigue                       | 66 Psychological symptoms     |
| 3 Anger-hostility                                     | 35 Fatigue-inertia                | 67 Psychological distress     |
| 4 Anxiety                                             | 36 Financial difficulties        | 68 Psychological impact       |
| 5 Appetite loss                                       | 37 Financial impact              | 69 Psychological issues       |
| 6 Bodily pain                                         | 38 Fullness following meals      | 70 Reflux                     |
| 7 Body image                                          | 39 Functional wellbeing          | 71 Role emotional             |
| 8 Choking                                             | 40 General health                | 72 Role function              |
| 9 Cognitive function                                  | 41 Global evaluations            | 73 Role physical              |
| 10 Confusion-bewilderment                             | 42 Global QOL                    | 74 Sexual relationships       |
| 11 Constipation                                       | 43 Global satisfaction           | 75 Sleep disturbances         |
| 12 Cough                                              | 44 Global life satisfaction      | 76 Social activities          |
| 13 Deglutition                                        | 45 Hair loss                     | 77 Social environment         |
| 14 Deglutition disturbances                           | 46 Health perceptions            | 78 Social function            |
| 15 Depression                                         | 47 Healthcare orientation        | 79 Social issues              |
| 16 Depression-dejection                               | 48 Heartburn                     | 80 Social/family wellbeing     |
| 17 Diarrhoea                                          | 49 Indigestion                   | 81 Speech                     |
| 18 Diarrhoea/constipation                              | 50 Information needs             | 82 Spiritual issues           |
| 19 Difficulty in swallowing                           | 51 Insomnia                      | 83 Stool formation            |
| 20 Domestic environment                               | 52 Loss of independence          | 84 Swallowing problems        |
| 21 Dry mouth                                          | 53 Loss of weight                | 85 Symptoms                   |
| 22 Dumping                                            | 54 Medical treatment             | 86 Symptoms of GERD           |
| 23 Dysphagia                                          | 55 Mental health                 | 87 Taste                      |
| 24 Dyspnoea                                           | 56 Nausea & vomiting             | 88 Tension-anxiety            |
| 25 Eating                                             | 57 Odynophagia                   | 89 Trouble swallowing saliva  |
| 26 Eating restrictions                                | 58 Pain                          | 90 Vigour-activity            |
| 27 Eating with others                                 | 59 Physical symptoms             | 91 Vitality                   |
| 28 Emotion                                            | 60 Physical activity             | 92 Vocational environment     |
| 29 Emotional function                                 | 61 Physical distress             | 93 Weight                     |
| 30 Emotional problems (anxiety)                       | 62 Physical function             | 94 Worry/fear of recurrence   |
| 31 Emotional wellbeing                                | 63 Physical problems             |                             |
| 32 Esophageal cancer scale                            | 64 Physical wellbeing            |                             |

GERD: Gastroesophageal reflux disease; QOL: Quality of life.
been undertaken using a range of methods, in particular the approach to including PROs [86-89]. In rheumatoid arthritis, for example, the initial American College of Rheumatology (ACR) core set was developed by a committee of experts (16 professionals in rheumatoid arthritis trials, health services research and biostatistics) who reviewed the literature on the validity of trial outcomes (for example, sensitivity to change or how well it predicted/correlated with a definite clinical change) and used a nominal group process to recommend and reach consensus on a list of core outcomes. The list was presented and finalised at a specialist international conference (OMERACT: Outcome Measures for Rheumatology in Clinical Trials) and contained both clinical and PROs, although patients were not involved in the consensus process. Outcomes were specific (for example, number of swollen joints) or more general domains (for example, functional status), with recommendations on how to measure the outcomes decided later [86]. Subsequent OMERACT conference discussions and workshops deliberately involved patients and

| Scale name | PROMs using identical scale name | Scales with similar* names (origin PROM) |
|------------|---------------------------------|----------------------------------------|
| Pain       | QLQ C30                          | Bodily pain (SF 36)                     |
|            | QLQ C36                          |                                        |
|            | QLQ OES18                         |                                        |
|            | QLQ OES24                         |                                        |
|            | QLQ OG25                          |                                        |
|            | DAUG 32                           |                                        |
| Physical function | MOS SF20                       | Physical wellbeing (FACT-E)        |
|            | QLQ C30                          | Physical problems (PNPC)              |
|            | QLQ C36                          | Physical distress (RSCL)              |
|            | SF 36                            | Physical activity (DAUG 32)           |
|            | EQOL                             | Role physical (SF 36)                 |
|            | G IQLI                           |                                        |
| Social function | QLQ C30                        | Social activities (PNPC)             |
|            | QLQ C36                          | Social environment (PAIS)             |
|            | SF 36                            | Social issues (PNPC)                  |
|            | EQOL                             | Social/family wellbeing (FACT-E)      |
|            | G IQLI                           |                                        |
| Activities of daily living | PNPC                           |                                        |
|            | EQOL                             |                                        |
|            | RSCL                             |                                        |
| Dysphagia  | QLQ OES18                         |                                        |
|            | QLQ OES24                         |                                        |
|            | QLQ OG25                          |                                        |
| Emotional function | QLQ C30                    | Emotional wellbeing (FACT-E)         |
|            | QLQ C36                          | Emotional problems (anxiety) (QLQ OES24) |
|            | EQOL                             | Emotion (G IQLI)                      |
|            |                                  | Role emotional (SF 36)                 |
| Global QOL | QLQ C30                          | Global life satisfaction (LAGS)        |
|            | QLQ C36                          |                                        |
| Nausea & Vomiting | DAUG 32                | Psychological issues (PNPC)           |
|            | QLQ C30                          | Psychological impact (LAGS)           |
|            | QLQ C36                          | Psychological symptoms (RSCL, adapted) |
| Psychological distress | RSCL                        |                                        |
|            | (adapted)                        |                                        |
|            | PAIS                             |                                        |
| Reflux     | DAUG 32                          |                                        |
|            | QLQ OG25                          |                                        |
|            | QLQ OES18                         |                                        |

*‘similar’ = one or more identical word.
### Table 4 Categorised PRO health domains showing number of items in existing PROMs assessing each domain

| PRO domain                      | Specific measures for gastrointestinal disease | Cancer specific measures | Generic measures | Frequency assessed in PROMs (total number of items) |
|--------------------------------|-----------------------------------------------|--------------------------|------------------|---------------------------------------------------|
| EQOL                           | FACT-E                                       | GIQL                     | DAUGS 32         | LAGS                                              | QLQ-OG25   | QLQ-OES18  | QLQ-OES24   | Adapted RSCL⁴ | Adapted RSCL⁵ | RSCL        | PNPC        | QLQ-C30 | QLQ-C36 | WOCS | MOS | SF-20 | PAIS | POMS | SF-36 | HADS |
|                                |                                               |                          |                  |                                                   |            |            |            |                 |                     |             |             |             |            |          |       |      |       |      |       |       |       |
| **Disease specific**           |                                               |                          |                  |                                                   |            |            |            |                 |                     |             |             |             |            |          |       |      |       |      |       |       |       |       |
| Pain/pain-related swallowing  | 3                                             | 1                        | 1                | 2                                                | 1          | 3          | 3          | 3                | 5                    | 6           | 1          | 2          | 1          | 1       | 2    |      |       |      |       |       |       |       |
| Appetite/eating/taste         | 4                                             | 3                        | 4                | 5                                                | 1          | 6          | 4          | 4                | 3                    | 1           | 1          | 1          | 1          |         |      |      |       |      |       |       |       |       |
| Fatigue                       | 1                                             | 1                        | 2                | 1                                                | 1          | 2          | 1          | 3                | 3                    | 15          | 4          |           |            |         |      |      |       |      |       |       |       |       |
| Dysphagia/swallowing saliva    | 1                                             | 4                        | 1                | 2                                                | 1          | 4          | 4          | 4                | 2                    | 1           |            |            |            |         |      |      |       |      |       |       |       |       |
| Regurgitation/vomiting         | 1                                             | 1                        | 5                | 1                                                | 1          | 1          | 1          | 1                | 1                    | 2           | 1          | 1          | 1          |         |      |      |       |      |       |       |       |       |
| Reflux/heartburn               | 1                                             | 6                        | 1                | 2                                                | 1          | 1          | 1          | 1                | 1                    | 1           |            |            |            |         |      |      |       |      |       |       |       |       |
| Nausea                         | 1                                             | 1                        | 1                | 1                                                | 1          | 1          | 1          | 1                | 1                    | 1           |            |            |            |         |      |      |       |      |       |       |       |       |
| Belching/bloating/gas          | 1                                             | 1                        | 4                | 3                                                | 1          |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Diarrhoea/frequent bowel movements | 5                                             | 2                        | 1                |                                                   |            | 1          |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Cough                          | 1                                             |                           | 1                | 1                                                | 1          | 1          |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Choking                        | 1                                             |                           | 1                | 1                                                | 1          | 1          |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Dry mouth                      | 1                                             |                           | 1                | 1                                                | 1          | 1          |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Breathing                      | 1                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Weight                         | 1                                             | 1                        | 1                | 1                                                | 1          |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Communication/speech difficulties | 2                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Eating - social impact         | 1                                             |                           | 1                | 1                                                |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Sexual function                | 1                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Constipation                   | 1                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Hair loss                      | 1                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Body image                     | 1                                             |                           |                  |                                                   |            |            |            | 1                | 1                    |              |            |            |            |         |      |      |       |      |       |       |       |       |
| Dizziness/dumping             | 5                                             | 2                        |                   |                                                   |            |            |            | 1                |                       |              |            |            |            |         |      |      |       |      |       |       |       |       |

*Adapted RSCL a* and *Adapted RSCL b* refer to modified versions of the RSCL instrument.
| Generic | Emotional function | 2 | 8 | 6 | 3 | 2 | 3 | 7 | 5 | 7 | 19 | 4 | 8 | 4 | 5 | 6 | 43 | 5 | 14 | 18 (151) |
|---------|-------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Role physical/ADLs | 2 | 2 | 1 | 2 | 8 | 8 | 9 | 3 | 4 | 3 | 4 | 5 | 13 (59) |
| Physical function | 2 | 3 | 1 | 25 | 3 | 4 | 4 | 6 | 5 | 1 | 9 | 11 (63) |
| Social function | 1 | 7 | 1 | 19 | 2 | 2 | 1 | 17 | 2 | 9 | 52 (52) |
| Generic health | 1 | 1 | 1 | 1 | 5 | 2 | 6 | 7 | 17 |
| Sleep | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 | 7 |
| Global QOL | 1 | 1 | 4 | 1 | 1 | 2 | 6 | 10 |
| Cognition | 1 | 3 | 2 | 1 | 7 | 5 | 14 |
| Role emotional | 2 | 4 | 2 | 3 | 4 | 11 |
| Financial issues | 5 | 1 | 1 | 1 | 4 | 8 |
| Spiritual issues | 2 | | | | | | |

\*Adapted RSCL - de Boer study. 
\*Adapted RCSL - van Knippenberg study 
ADLs: Activities of daily life.
led to the addition of fatigue in the ACR core set [87,90], and continued work using interviews with patients, identified further important PROs [5]. This led to the development of a ‘patient core set’ of disease-specific and global outcome domains solely derived from patient opinion to complement the professional ACR core set [88]. Our current study methodology ensures that the patient perspective and relevant PROs inform the development of a core set of outcome domains from an early stage, because it examines the content of validated PROMs which are developed with significant patient involvement. The identified PRO domains will be prioritised using a Delphi method to reach consensus on the important to include in the core outcome set, alongside clinical outcomes [1], and is the focus of future work. Patients, surgeons and clinical nurse specialists will be surveyed to ensure the opinions of all key stakeholders are sought, a recommended approach by the COMET (Core Outcome Measures in Effectiveness Trials) initiative [4].

This study included a detailed systematic search to identify PROs measured in oesophageal cancer studies and used rigorous methodology to identify health domains, however, it does have weaknesses. The categorisation of question items into health domains was performed by two experts and independently checked by other members of the research team, including a patient advocate, but it is possible that others may have categorised items differently. Inter-rater reliability statistics could have been recorded to describe agreement between the experts when categorising items. Future work therefore is needed to standardised and validate this method. In addition, presentation of the methodology to a greater number of patients or patient representatives could strengthen the robustness and reliability of the categorisation process.

Conclusion
In summary, this study demonstrates there is diversity in the PROMs selected to evaluate radical treatment for oesophageal cancer. Within and between PROMs there is a lack of clarity between named scales and items and the underlying health domains being assessed meaning data synthesis is limited. A methodology for identifying important PRO health domains is proposed which can be used to inform the development of a core set of health domains. Following this it will be necessary to determine accurate and efficient ways to measure these core domains, drawing on items banks developed by initiatives such as PROMIS (Patient Reported Outcomes Measurement Information System) and COSMIN (Consensus-based Standards for the selection of health status Measurement Instruments) [91,92].

Additional file

Additional file 1: Search strategy as applied to MEDLINE (OVID).

Abbreviations
COS: Core outcome set; HRQL: Health-related quality of life; PRO: Patient-reported outcome; PROM: Patient-reported outcome measure; RCT: Randomised controlled trial.

Competing interests
The authors declare they have no competing interests.
Authors’ contributions
RM conducted and coordinated the review and drafted the manuscript. MJ was involved in the systematic databases searches, data extraction and contributed to the draft manuscript. IK and JD contributed to the systematic search for studies. RW helped to screen abstracts for the review. ST contributed to the design of the study. MJ checked the data extraction, grouped data into PRO domains and contributed to the draft manuscript. JB conceived the idea for the review, was involved in the design of the study and data extraction, grouped data into PRO domains, contributed to the draft manuscript and supervised the project. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by the UK Medical Research Council Network of Hubs for Trials Methodology Research (ConDuCT Hub - Collaboration and Innovation in Difficult and complex randomised Controlled Trials - G0800860) and the National Institute for Health Research (NIHR) Patient Benefit Program PB-PG-0807–14131. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The authors thank Mrs Jackie Elliott, representative of the Gastro/ Oesophageal Support & Help (GOSH) group, Bristol, for checking the categorisation of items.

Author details
1School of Social and Community Medicine, University of Bristol, Canyny Hall, 39 Whatley Road, Bristol BS8 2PS, UK. 2Department of Medical Psychology, Academic Medical Center/University of Amsterdam, Meibergdreef 5, Amsterdam NL 1105 AZ, Netherlands. 3Department of Public Health, Erasmus MC, P.O. Box 2040, Rotterdam NL 3000 CA, Netherlands. 4Division of Surgery, Head and Neck, University Hospitals Bristol NHS Foundation Trust, Level 3, Dolphin House, Bristol Royal Infirmry, Marlborough Street, Bristol BS2 8HW, UK.

Received: 3 October 2013 Accepted: 17 January 2014 Published: 5 February 2014

References
1. Blencowe NS, Strong S, McNair AGK, Brookes ST, Crosby T, Griffin SM, Blazeby JM: Reporting of short-term clinical outcomes after esophagectomy: a systematic review. Ann Surg 2012, 255:658–666.
2. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, Decullier E, COMET (Core Outcome Measures in Effectiveness Trials) Initiative. Publication bias and outcome reporting bias. Plos One 2008, 3(3):831.
3. Williamson PM, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P: Developing core outcome sets for clinical trials: issues to consider. Trials 2012, 13:132.
4. COMET (Core Outcome Measures in Effectiveness Trials) Initiative. http://www.comet-initiative.org/.
5. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S: What outcomes from pharmacologic treatments are important to people with rheumatoid arthritis? Creating the basis of a patient core set. Arthri Care Res 2010, 62(s4):649–666.
6. US Department of Health and Human Services Food and Drug Administration: Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims; 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
7. Jacobs M, Macfie RD, Blazeby JMK, Van Berge Henegouwen MI, De Haes HC, Smets EM, Sprangers MA: Systematic review reveals limitations of studies evaluating health-related quality of life after potentially curative treatment for esophageal cancer. Qual Life Res 2012, 21:1787–1803.
8. Parameswaran R, McNair A, Avery KNL, Berrisford RG, Wajed SA, Sprangers MAG, Blazeby JM: The role of health-related quality of life outcomes in clinical decision-making in surgery for esophageal cancer: a systematic review. Ann Surg Oncol 2008, 15:227–239.
9. Jacobs M, Macfie RC, Elbers RG, Sitnikova K, Korfage I, Smets EM, Henselmans I, Van Berge Henegouwen MI, De Haes JC, Blazeby JM, Sprangers MA: Meta-analysis shows clinically relevant and long-lasting deterioration in health-related quality of life after esophageal cancer surgery. Qual Life Res 2013. [Epub ahead of print].
10. Guyatt GH, Feeny DH, Patrick DL: Measuring health-related quality of life. Ann Intern Med 1993, 118:622–629.
11. Ariga H, Nemoto K, Miyazaki S, Yoshioka T, Ogawa Y, SakaYachi T, Jingu K, Miyata G, Onodera K, Ichikawa H, Kamel T, Kato S, Ishioka C, Satomi S, Yamada S: Prospective comparison of surgery alone and chemoradiotherapy with selective surgery in resectable squamous cell carcinoma of the esophagus. Int J Radiat Oncol 2009, 75:348–356.
12. Avery KNL, Metcalfe C, Barham CP, Alderson D, Falk SJ, Blazeby JM: Quality of life during potentially curative treatment for locally advanced oesophageal cancer. British J Surg 2007, 94:1369–1376.
13. Barbour AP, Lagergren P, Hughes R, Alderson D, Barham CP, Blazeby JM: Health-related quality of life among patients with adenocarcinoma of the gastro-oesophageal junction treated by gastrectomy or oesophagectomy. Br J Surg 2008, 95:80–84.
14. Blazeby JM, Farnond JR, Donovan J, Alderson D: A prospective longitudinal study examining the quality of life of patients with esophageal cancer. Cancer 2000, 88:1781–1787.
15. Blazeby JM, Sanford E, Falk SJ, Alderson D, Donovan JL: Health-related quality of life during neoadjuvant treatment and surgery for localized esophageal carcinoma. Cancer 2005, 103:1791–1799.
16. Blazeby JM, Williams MH, Brookes ST, Alderson D, Farnond JR: Quality-of-lifemeasurement in patients with esophageal cancer. Gut 1995, 37:505–508.
17. Brooks JA, Kesler KA, Johnson CS, Ciaccia D, Brown JW: Prospective analysis of quality of life after surgical resection for esophageal cancer: Preliminary results. J Surg Oncol 2002, 81:185–194.
18. Cense HA, Visser MRM, Van Sandick JW, de Boer AGEM, Lamme B, Obertop H, Van Lanschot JIB: Quality of life after colon interposition by necessity for esophageal cancer replacement. J Surg Oncol 2004, 88:32–38.
19. Chang LC, Oelschlagker BK, Quiroga E, Parra JD, Mulligan M, Wood DE, Pellegrini CA: Long-term outcome of esophagectomy for high-grade dysplasia or cancer found during surveillance for Barrett’s esophagus. J Gastrointest Surg 2006, 10:341–346.
20. de Boer AGEM, Genovesi PIO, Sprangers MAG, van Sandick JW, Obertop H, Van Lanschot JIB: Quality of life in long-term survivors after curative transhiatal oesophagectomy for oesophageal carcinoma. Brit J Surg 2000, 87:1716–1721.
21. de Boer AGEM, van Lanschot JIB, van Sandick JW, Hulstiger IJB, Stalmeire FMM, de Haes JCJM, Tilanus HW, Obertop H, Sprangers MAG: Quality of life after transhiatal compared with extended transthoracic resection for adenocarcinoma of the esophagus. J Clin Oncol 2004, 22:4020–4028.
22. Djan T, Lagergren J, Blazeby JM, Lagergren P: Long-term health-related quality of life following surgery for oesophageal cancer. Brit J Surg 2008, 95:1121–1126.
23. Egberts JH, Schwinbkind B, Bestmannn B, Schafmayer C, Egberts F, Frensch F, Kuechler T, Tepel J: Impact of the site of anastomosis after oncologic esophagectomy on quality of life – A prospective, longitudinal outcome study. Ann Surg Oncol 2008, 15:556–575.
24. Gillham CM, Aheme N, Rowley S, Moore J, Hollywood D, O’Byrne K, Reynolds JV: Quality of life and survival in patients treated with radical chemoradiation alone for oesophageal cancer. Clin Oncol-Uk 2008, 20:227–233.
25. Gockel I, Gonner U, Domeyer M, Lang H, Junginger T: Long-term survivors of esophageal cancer: disease-specific quality of life, general health and complications. J Surg Oncol 2010, 102:516–522.
26. Gradauskas P, Rubikas R, Safers V: Changes in quality of life after esophageal resections for carcinoma. Medicina 2006, 42:187–194.
27. Hallas CN, Patel N, Do A, Jackson M, Murphy P, Drakeley ML, Socra A, Page RD: Five-year survival following esophageal cancer resection: psychosocial functioning and quality of life. Psychol Health Med 2001, 6:85–94.
28. Headrick JR, Nichols FC, Miller DL, Allen MS, Trastek VF, Deschamps C, Long-term survivors of esophagectomy on quality of life A prospective, longitudinal outcome study examining the quality of life of patients with esophageal cancer. Ann Thorac Surg 2002, 73:1697–1702.
29. Higuchi A, Minamide J, Ota Y, Takeda K, Aoyama N: Evaluation of the quality of life after surgical treatment for thoracic esophageal carcinoma. Esophagus 2006, 3:53–59.
30. Hummuzlu M, Arastad H, Arastad AKH, Hjerstedt M, Viste A: Health-related quality of life in long-term survivors after high-dose chemoradiotherapy followed by surgery in esophageal cancer. Dis Esophagus 2011, 24:39–47.
50. Staal EFWC, van Sandick JW, van Tinteren H, Cats A, Aaronson NK: Health-related quality of life among patients cured by surgery for esophageal cancer. Cancer Care Soc. 2007: 110:588–693.

31. Leibman S, Smithers BM, Gottle DC, Martin I, Thomas J: Minimally invasive esophagectomy. Surg Endosc 2006: 20:428–433.

32. Lulettch, JD, Veloso-Rivera M, Buenaventura PO, Christie NA, McLawson J, Little VR, Schuwer PR, Close JM, Fernando HC: Minimally invasive esophagectomy outcomes. Ann Surg 2003 238:486–494.

33. Martin L, Lagergren J, Lindblad M, Rouvelas I, Lagergren P: Malnutrition after esophageal cancer surgery in Sweden. Brit J Surg 2007: 94:1966–1950.

34. McLarty AJ, Deschamps C, Trastek VF, Allen MS, Paivorelo PC, Hamssen WS: Esophageal resection for cancer of the esophagus: long-term function and quality of life. Ann Thorac Surg 1997: 63:1568–1572.

35. Moraca RJ, Low DE: Outcomes and health-related quality of life for esophagectomy for high-grade dysplasia and intramucosal cancer. Arch Surg-Chicago 2006: 141:545–549.

36. Nakamura M, Kido Y, Hosoya Y, Yano M, Nagai H, Monden M: Postoperative gastrointestinal dysfunction after 2-field versus 3-field lymph node dissection in patients with esophageal cancer. Surg Today 2007: 37:379–382.

37. Ohanlon DM, Harkin M, Karat D, Sergeant T, Hayes N, Griffin SM: Quality of life assessment in patients undergoing treatment for esophageal carcinoma. Brit J Surg 1995: 82:1682–1685.

38. Olens MF, Larsson M, Hammered E, Lundell L: Physical function and quality of life after thoracoabdominal esophagectomy resection: results of a longitudinal follow-up study. Digest Surg 2005: 22:53–68.

39. Paransennan P, Blazey JM, Hughes R, Mitchell K, Bensford RG, Wajed SA: Health-related quality of life after minimally invasive esophagectomy. Brit J Surg 2010: 97:525–531.

40. Reynolds IV, McLauhlin R, Moore J, Rowley S, Ravi N, Byrne PJ: Prospective evaluation of quality of life in patients with localized esophageal cancer treated by multimodal therapy or surgery alone. Brit J Surg 2006: 93:1084–1090.

41. Rosmolen WD, Boer KB, de Leeuw R, Stolp C, Henegouwen MN, Bergman JGIM, Sprangers MAG: Quality of life and fear of cancer recurrence after endoscopic and surgical treatment for early neoplasia in Barrett's esophagus. Endoscopy 2010: 42:525–531.

42. Rutegaud M, Lagergren J, Ronvels I, Lindblad M, Blazey JM, Lagergren P: Population-based study of surgical factors in relation to health-related quality of life after esophageal cancer resection. Brit J Surg 2008: 95:592–601.

43. Safiedine N, Xu W, Quadri SM, Hsu W, Takeda F: Prospective evaluation of quality of life of patients with locoregional esophageal cancer after definitive chemoradiation therapy followed by surgery. Qual Life Res 2007: 16:825–830.

44. Schmidt CE, Bestmann E, Kuchler T, Schmid A, Kremer B: Quality of life associated with surgery for esophageal cancer: differences between collar and intrathoracic anastomoses. World J Surg 2004: 28:355–360.

45. Schneider L, Hartwig W, Aulmann S, Lenzen C, Strobel O, Fritz S, Hackert T, Keller M, Buen R, Werner J: Quality of life after emergency vs. elective esophagectomy with cervical reconstruction. Scand J Surg 2009: 99:3–8.

46. Seragi-Vernant S, Bossert M, Crehange G, Buffet-Diny J, Puyraveau M, Maingon P, Mercier M, Bossert JF: Feasibility of chemoradiotherapy for oesophageal cancer in elderly patients aged ≥75 years. A prospective, single-arm Phase II study. Drugs Aging 2009: 26:255–262.

47. Spector NM, Hicks FD, Pickleson JI: Quality of life and symptoms after surgery for gastroesophageal cancer. Gastroenterol Nurs 2002: 25:120–125.

48. Staal EFWC, van Sandick JW, van Tinteren H, Cats A, Aaronson NK: Health-related quality of life in long-term esophageal cancer survivors after potentially curative treatment. J Thorac Cardiovasc Surg 2010: 140:777–783.

49. Stein HU, Feith M, Mueller J, Werner M, Sievert JR: Limited resection for early (T1) esophageal carcinoma in Barrett's esophagus. Ann Surg 2000: 232:733–740.

50. Sweeney MR, Schieck L, Barsevick A, Babb JS, Goldberg M: Quality of life after esophagectomy for caner. Oncol Nusr Forum 2002: 29:1127–1131.

51. Tan QY, Wang RW, Jiang YG, Fan SZ, Hsin MKY, Gong TO, Zhou JH, Zhao YP: Lung volume reduction surgery allows esophageal tumor resection in selected esophageal carcinoma with severe emphysema. Ann Thorac Surg 1999: 82:1849–1856.
75. Osse BHP, Vernooij M, Schade E, Grof R: Towards a new clinical tool for needs assessment in the palliative care of cancer patients: The PNPC instrument. J Pain Symptom Manage 2004, 28:329–341.
76. Northouse LL: Mastectomy patients and the fear of cancer recurrence. Cancer Nurs 1981, 4:213–220.
77. Zigmond AS, Snaith RP: The hospital anxiety and depression scale. Acta Psychiatr Scand 1983, 67:361–370.
78. Derogatis LR: The psychosocial adjustment to illness scales (PAIS). J Psychosom Res 1986, 30:77–91.
79. Stewart AL, Hays RD, Ware JE: The MOS Short-form general health survey - reliability and validity in a patient population. Med Care 1988, 26:724–732.
80. Shachar S: A shortened version of the Profile of Mood States. J Pers Assess 1983, 47:305–310.
81. Ware JE, Sherbourne CD: The MOS 36-item short-form health survey (SF-36). Concepts, development, and use. Med Care 1992, 30:473–483.
82. Aaronson N, Ahmedzai S, Bullinger M: The EORTC core quality of life questionnaire: interim results of an international field study. In Effect of cancer on quality of life. Edited by Osoba D. Boca Raton FL: CRC Press Ltd; 1991.
83. de Boer A, Genovesi P, Sprangers MAG, van Sandick JW, Obertop H, van Lanschot JB: Quality of life in long-term survivors after curative transhiatal oesophagectomy for oesophageal carcinoma. British J Surg 2000, 87:1716–1721.
84. Lagergren P, Fayers P, Conroy T, Stein HJ, Sezer O, Hardwick R, Hammerlid E, de Boer A, Genovesi P, Sprangers MAG, van Sandick JW, Obertop H, van Lanschot JB: Quality of life in long-term survivors after curative transhiatal oesophagectomy for oesophageal carcinoma. British J Surg 2000, 87:1716–1721.
85. Clifton JC, Finley RJ, Gelfand G, Graham AJ, Inculet R, Malthaner R, Tan L, Lim J, Singer J, Lovato C: Development and validation of a disease-specific quality of life questionnaire: interim results of an international field study. In Effect of cancer on quality of life. Edited by Osoba D. Boca Raton FL: CRC Press Ltd; 1991.
86. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, Furst D, Goldsmith C, Kieszak S, Lightfoot R, Paulus H, Tugwell P, Weinblatt M, Widmark R, Williams J, Wolfe F: The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993, 36:29–740.
87. Kirwan JR, Hewlett SE, Helberg T, Hughes RA, Carr M, Hehir M, Kvien TK, Minnock P, Newman SP, Quest EM, Taal E, Wale J: Incorporating the patient perspective into outcome assessment in rheumatoid arthritis—progress at OMERACT 7. J Rheumatol 2005, 32:2250–2256.
88. Sanderson T, Morris M, Calinon R, Richards P, Hewlett S:Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis Care Res 2010, 62:647–656.
89. Sinha IP, Smyth RL, Williamson PR: Using the Delphi Technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. Plos Medicine 2011, 8(6).
90. Kirwan JR, Minnock P, Adebajo A, Benshan B, Choy E, de Wit M, Hazes M, Richards P, Saag KS, Suarez-Almazor M, Wells G, Hewlett S: Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007, 34:1174–1177.
91. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M: PROMIS Cooperative Group: The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH roadmap cooperative group during its first two years. Med Care 2007, 45:S3–51.
92. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bourne LM, de Vet HCW: Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012, 21:651–657.

Cite this article as: Macefield et al.: Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). Trials 2014 15:49.