

RESEARCH ARTICLE

Development of a core outcome set for use in community-based bipolar trials—A qualitative study and modified Delphi

Ameeta Retzer¹, Ruth Sayers², Vanessa Pinfold², John Gibson^{2,3}, Thomas Keeley⁴, Gemma Taylor⁵, Humera Plappert³, Bliss Gibbons⁶, Peter Huxley⁷, Jonathan Mathers⁸, Maximillian Birchwood^{9,10}, Melanie Calvert^{10,11*}

1 Centre for Patient Reported Outcomes Research (CPROR), Institute of Applied Health Research, and Birmingham Health Partners Centre for Regulatory Science and Innovation, University of Birmingham, Birmingham, United Kingdom, **2** The McPin Foundation, London, United Kingdom, **3** Institute for Mental Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom, **4** GlaxoSmithKline (formerly of CPROR, University of Birmingham), London, United Kingdom, **5** Addiction and Mental Health Group (AIM), Department of Psychology, University of Bath, Bath, United Kingdom, **6** Coventry and Warwickshire Partnership NHS Trust and Warwick Medical School, University of Warwick, Warwick, United Kingdom, **7** Centre for Mental Health and Society, Bangor University, Bangor, United Kingdom, **8** Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom, **9** Mental Health and Wellbeing, Warwick Medical School, University of Warwick, Coventry, United Kingdom, **10** School of Psychology, University of Birmingham, Birmingham, United Kingdom, **11** NIHR Birmingham Biomedical Research Centre, NIHR Surgical Reconstruction and Microbiology Research Centre and NIHR Applied Research Collaboration West Midlands, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, United Kingdom

* m.calvert@bham.ac.uk



OPEN ACCESS

Citation: Retzer A, Sayers R, Pinfold V, Gibson J, Keeley T, Taylor G, et al. (2020) Development of a core outcome set for use in community-based bipolar trials—A qualitative study and modified Delphi. PLoS ONE 15(10): e0240518. <https://doi.org/10.1371/journal.pone.0240518>

Editor: Rosemary Frey, University of Auckland, NEW ZEALAND

Received: April 2, 2019

Accepted: September 28, 2020

Published: October 28, 2020

Copyright: © 2020 Retzer et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Relevant data are within the manuscript and its supporting information files.

Funding: This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research programme awarded to MD as lead applicant (RP-PG-0611-20004). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. GlaxoSmithKline provided

Abstract

Background

A core outcome set (COS) is a standardised collection of outcomes to be collected and reported in all trials within a research area. A COS can reduce reporting bias and facilitate evidence synthesis. This is currently unavailable for use in community-based bipolar trials. This research aimed to develop such a COS, with input from a full range of stakeholders.

Methods

A co-production approach was used throughout. A longlist of outcomes was derived from focus groups with people with a bipolar diagnosis and carers, interviews with healthcare professionals and a rapid review of outcomes listed in bipolar trials on the Cochrane database. An expert panel with personal and/or professional experience of bipolar participated in a modified Delphi process and the COS was finalised at a consensus meeting.

Results

Fifty participants rated the importance of each outcome. Sixty-six outcomes were included in Round 1 of the questionnaire; 13 outcomes were added by Round 1 participants and were rated in Round 2. Seventy-six percent of participants (n = 38) returned to Round 2 and 60 outcomes, including 4 outcomes added by participants in Round 1, received a rating of 7–9

support in the form of a salary for TK. The specific roles of TK are articulated in the 'author contributions' section. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have read the journal's policy and have the following competing interests: VP is the research director at McPin Foundation, a registered charity (1117336). She is a board member of the MRC developing mind programme, Rethink Mental Illness clinical advisory group, THIS Institute (Cambridge University) involvement and engagement panel. McPin receives funding from a variety of sources including NIHR / UKRI research grants, evaluation commissions from other charities and universities. MC receives funding from the National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, the NIHR Surgical Reconstruction and Microbiology Research Centre at the University Hospitals Birmingham NHS Foundation Trust, Health Data Research UK, Innovate UK (part of UK Research and Innovation), Macmillan Cancer Support, UCB Pharma and the Patient-Centered Outcomes Research Institute (PCORI). MC has received personal fees from Astellas, Daiichi Sankyo, Glaukos, Takeda, Merck, and Ferring outside the submitted work, and chairs the ISOQOL (International Society for Quality of Life Research) Best Practice for PROs in Trials Taskforce. TK is an employee of GlaxoSmithKline and completed work on this project while employed at Centre for Patient Reported Outcomes Research (CPROR), University of Birmingham (UoB). GT is currently funded by a Cancer Research UK Population Researcher Postdoctoral Fellowship award (reference: C56067/A21330). This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents, products in development or marketed products associated with this research to declare.

by >70% and 1–3 by <25% of the sample. Fourteen participants finalised a COS containing 11 outcomes at the consensus meeting: personal recovery; connectedness; clinical recovery of bipolar symptoms; mental health and wellbeing; physical health; self-monitoring and management; medication effects; quality of life; service outcomes; experience of care; and use of coercion.

Conclusions

This COS is recommended for use in community-based bipolar trials to ensure stakeholder-relevant outcomes, facilitate data synthesis, and transparent reporting. The COS includes guidance notes for each outcome to allow the identification of suitable measurement instruments. Further validation is recommended for use with a wide range of communities and to achieve standardised measurement.

Introduction

This article describes the development of a core outcome set recommended for use in community-based trials for adults with bipolar, as part of the PARTNERS2 study. The PARTNERS2 study aims to help integrate primary care and community-based mental health services, further methodological detail and rationale for methodological decisions are available in the published protocol [1]. The term “bipolar” is used throughout this paper and in this research in place of “bipolar affective disorder” [2], as the preferred term of research team members with lived experience of bipolar and Bipolar UK, the leading charity in this area. This has been understood in this research in accordance with the definition and scope defined in the International Classification for Diseases (ICD-10) [2] (the ICD-11 will come into effect in 2022 [3]).

Bipolar is defined as “two or more episodes in which the [person]’s mood and activity levels are significantly disturbed, this disturbance consisting on some occasions of an elevation of mood and increased energy and activity (hypomania or mania) and on others of a lowering of mood and decreased energy and activity (depression)”. Life expectancy among people with a diagnosis of bipolar is reduced when compared with the general population [4–6] and the mortality gap appears to be widening [7]. In addition, individuals with bipolar experience increased unemployment and stigma [8], and bipolar research is under-researched when compared to other mental health research [9]. There is a growing primary evidence base that indicates the efficacy of psychological treatments (e.g. non-pharmaceutical interventions such as cognitive-behavioural therapy, family interventions, and psycho-education) for bipolar and its long-term management; however, meta-analyses are undermined by poor quality evidence [10]. A better understanding of interventions and to make comparisons between them requires data from well-designed and conducted randomised controlled trials (RCTs) [11] using a unified approach to outcome selection.

Trialists evaluate the effectiveness of an intervention in a clinical trial by choosing outcomes that reflect any beneficial or harmful effects—these can be specific, such as change in weight, or broad constructs, such as pain [12]. Outcomes can be measured in several ways, including the use of laboratory findings, biomarkers, or mortality; or they can be reported by observers, clinicians, or the patient themselves. RCTs can provide robust evidence to inform clinical decision making and health care policy development [13], but the inconsistent use of highly varied

trial outcomes within the same research area can undermine evidence synthesis. Additionally, for outcome data to be useful, the outcomes used must be of relevance to a range of stakeholders [14], including people with a bipolar diagnosis, carers and healthcare professionals [12, 15]. Commonly, the outcomes used in bipolar research have focused on clinical outcomes, such as change in symptoms as assessed by clinicians. However a mounting view in mental health research suggests that a broader set of outcomes may better suit the goals that people with diagnoses seek to achieve during treatment. This has been indicated in the case of schizophrenia [16, 17], and the Bipolar Priority Setting partnership suggests this is also the case for those with a bipolar diagnosis [18].

A core outcome set (COS) is a standardised collection of outcomes recommended to be reported in all controlled trials within a research area [12]. A COS represents the minimum outcomes to be measured and reported when undertaking a trial [19]. A COS for use in trials for those receiving non-pharmaceutical community-based interventions for bipolar (rather than as a hospital in-patient) could reduce reporting bias and enable evidence synthesis. The aim of this research is to develop such a COS.

This is the first study to develop a COS for community-based bipolar trials but builds upon an effort in the field to unify the outcomes and priorities within bipolar research. In 2010, the development of two “core sets” for bipolar based on the International Classification of Functioning, Disability and Health (ICF) guide [20] began. Focusing on functioning, the core sets use the ICF guide for bipolar, and while they may be used as outcome measures in research settings, the main intention of these core sets are for use in clinical practice. In 2015 a set of “patient important outcomes” [21] were published, aiming to investigate the relative importance of bipolar outcomes from the perspective of patients. The set of “patient important outcomes” considered the views of a single stakeholder group, those of people with a bipolar diagnosis, whereas the COS developed in our research included a range of information sources and engaged the views of different key stakeholders, allowing for the potential identification of gaps [22]. In comparison with the COS developed in our research, the ICF study constructed a shortlist of treatment outcomes relevant in the evaluation and selection of pharmacological treatments and was not conceptualised with intention for use in community-based bipolar trials. Similarly, the ICF core sets [23] were not developed with the sole intention of its use to be in community-based bipolar trials, as has the COS described here. In 2016, the James Lind Alliance (JLA) published a Priority Setting Partnership (PSP) about bipolar [24]. This provides researchers with 10 new bipolar research priorities into which the COS could be adopted and demonstrates increasing interest in bipolar research. The Royal College of Psychiatrists have also provided an overview of outcome measures for use in adult psychiatry [25].

Methods

Ethics statement

Ethical approval was sought and granted from the National Research Ethics Service (NRES) West Midlands—Edgbaston (Reference Number: 14/WM/0052). The research was registered with the COMET (Core Outcome Measures in Effectiveness Trials) Initiative [26] and is reported in adherence with the COS-STAR Statement [27] and the GRIPP2-SF [28]. The study was conceived prior to publication of the COS Standards for Development [29] recommendations but is in alignment with its standards.

A co-production approach, drawing upon the expertise of academics, healthcare professionals, people with bipolar, and carers, was used in this research. Experiential expertise from people with a bipolar diagnosis and carers was facilitated through the PARTNERS2 patient and public involvement (PPI) programme. Peer researchers with a diagnosis of bipolar were

employed within the research team and three Lived Experience Advisory Panels (LEAPs) were established, consisting of an average of 5 people with schizophrenia and bipolar diagnoses and family members with experience of caring for individuals with mental health diagnoses. Research team members and advisors with lived experience advised on and provided strategies to ensure the research phases and output would be relevant and useful to those with bipolar. Research team members with lived experience were involved in all of the tasks required for the fulfilment of the research, including collection and analysis of data, in the same manner as those without lived experience. However, research team members with lived experience had an additional role whereby they would provide ongoing advice to colleagues that would ensure the research and the manner in which it was conducted remained relevant and appropriate to people with bipolar, for whom the COS would be finally intended. In this paper, references to the “research team”, this should always be taken to include those research team members with lived experience. The LEAPs discussed and advised on the phases of COS development on 8 occasions between February 2015 and November 2016. The LEAPs often discussed the same phases at multiple meetings. Work undertaken by LEAP members included commenting and advising on consent and information materials (in particular ensuring use of accessible language), recruitment strategies, and delivery of research. Meetings were held with individual members of the LEAP on two occasions to pilot the Delphi interface.

This article details the process through which an outcome longlist was developed and subjected to rating and refining via a two round Delphi survey and a stakeholder consensus meeting, resulting in a COS recommended for use in community-based bipolar trials. The COS was developed via three phases: 1) identifying a longlist of outcomes from focus group discussions, one-to-one interviews, and a rapid review, 2) refining the outcome longlist using Delphi methodology, and 3) finalising the COS in a consensus meeting as the last part of the Delphi (see Fig 1). The three phases involved input from several stakeholders (see Fig 2).

The authors can confirm that the development of the bipolar COS reported was undertaken in adherence to the published protocol (steps 1–3), with the exception of use of 2 rounds of Delphi survey instead of 3, to mediate sample attrition. The latter steps of the published protocol, 4 and 5, will be completed and reported separately. The schizophrenia COS is still in development.

Phase 1: Outcome identification

Outcomes were identified through a qualitative study involving focus group discussions with people with a bipolar diagnosis and carers with experience of bipolar; one-to-one interviews with healthcare professionals and researchers; and a rapid review of the Cochrane database.

Qualitative sampling, recruitment and data collection. People with a bipolar diagnosis and carers were recruited in the West Midlands and Lancashire regions of the UK through local support groups, electronic advertisement via third sector organisations, and snowball sampling. Eligible participants self-identified as having a bipolar diagnosis (current or previous); receiving/having received mental health treatment in a community setting; being aged between 18–65 years, as recommended by clinical research team members; and fluent in English. Eligible carers self-identified as having supported individuals with bipolar diagnoses as a family carer; being aged between 18–65 years; and fluent in English. Consent was taken at entry to the study and re-taken at the focus group. Focus group discussions were co-facilitated by two research team members, one of whom is a “peer researcher” with experience of mental health problems in addition to having a research background. This approach is intended to build rapport and trust with participants [30] through self-disclosure. Participants were asked open questions [31] (see S1 File) and produced written descriptions of the effects of their

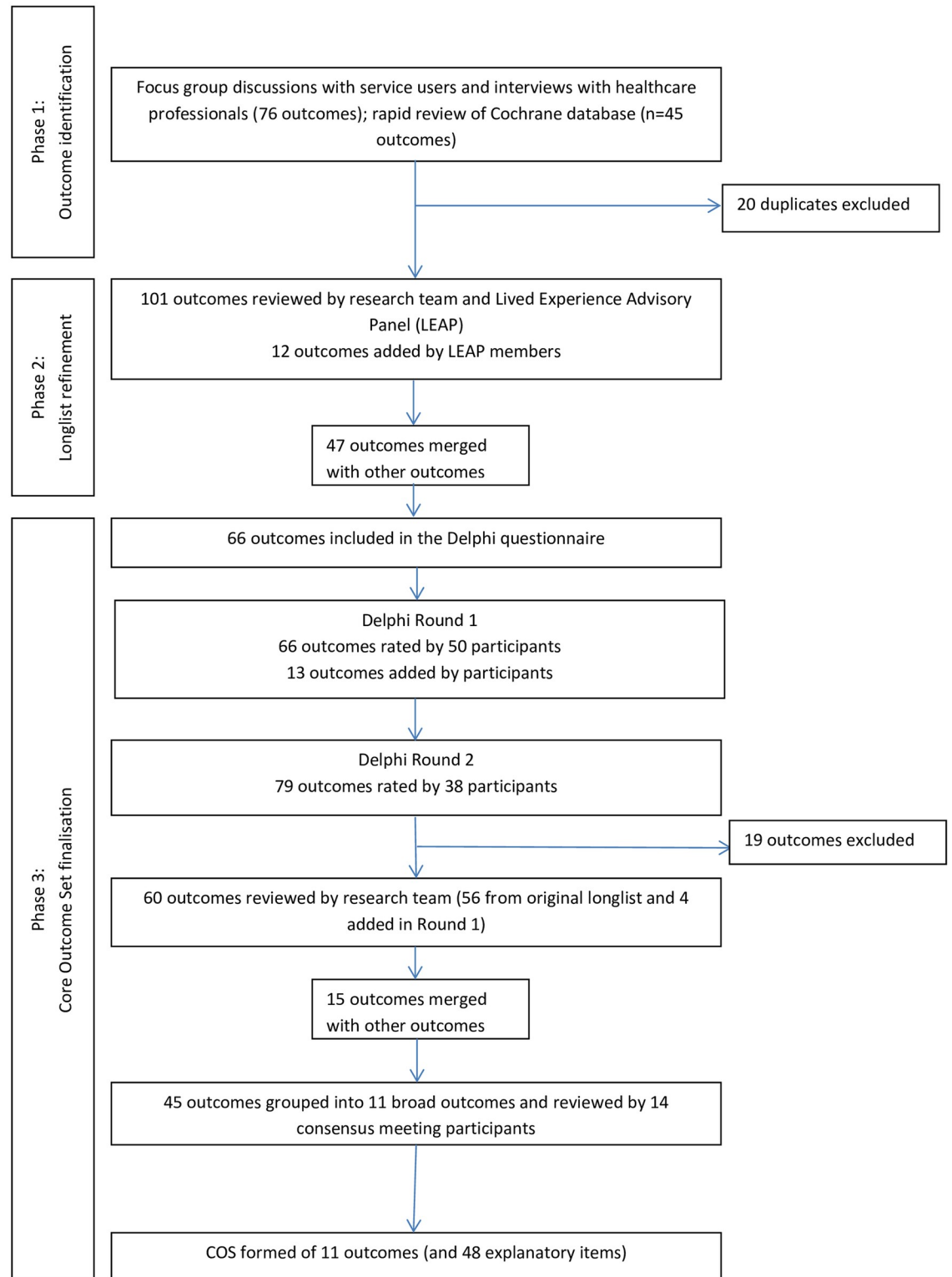


Fig 1. Illustration of core outcome set development process.

<https://doi.org/10.1371/journal.pone.0240518.g001>

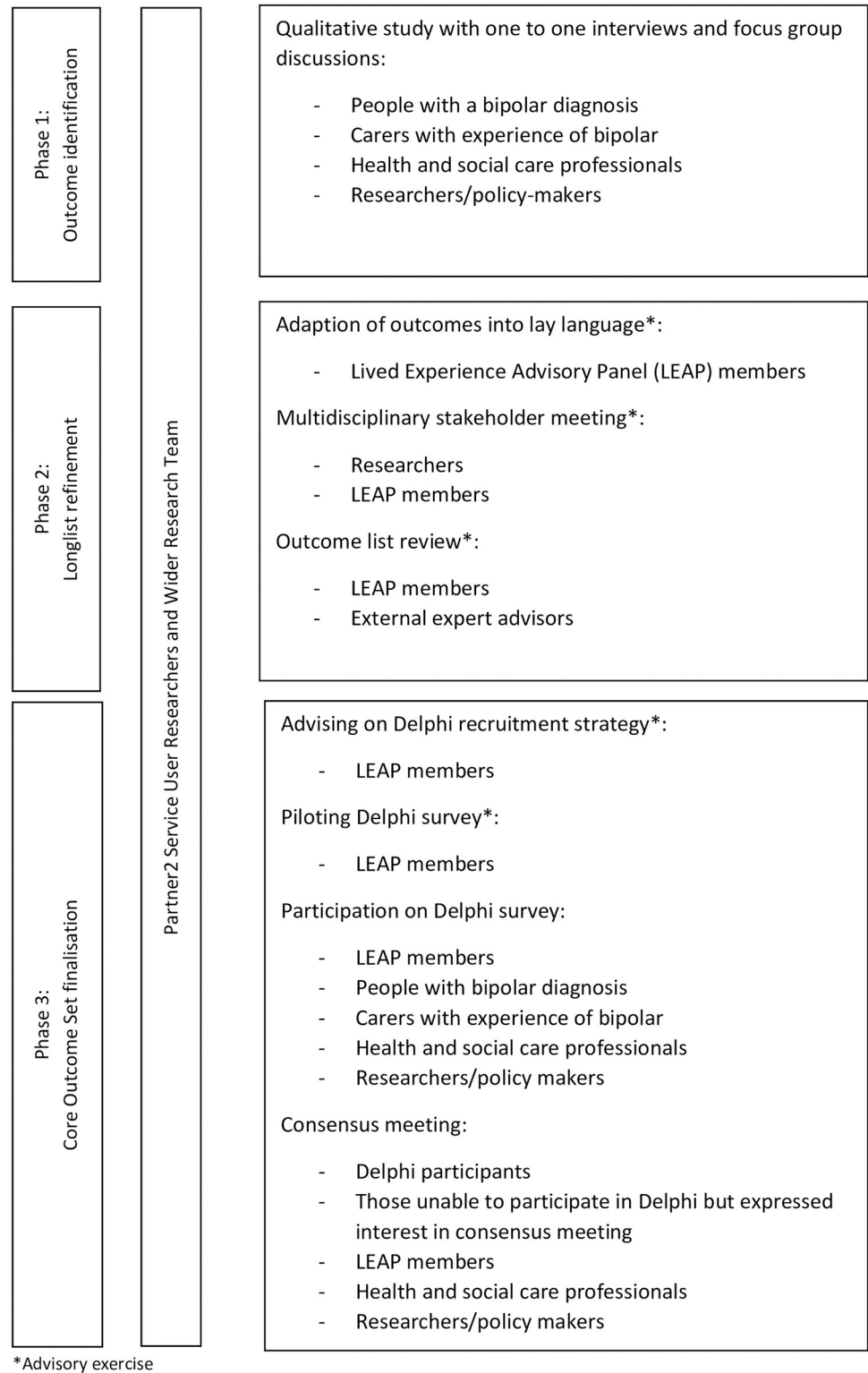


Fig 2. Overview of stakeholder involvement per phase.

<https://doi.org/10.1371/journal.pone.0240518.g002>

mental health problems on daily living which were placed on the walls of the room, allowing their categorisation under broad headings. This exercise of “concept mapping” [32] facilitated discussion and engagement from all participants, and informed subsequent analyses.

Researchers and healthcare professionals were recruited through existing professional contacts of the research team, aided by the prominent members of the bipolar research field being known to the team. Purposive sampling was used to capture a range of professional roles including health care professionals, social care professionals, commissioners, researchers and policy makers (see [S2 File](#)). Eligible participants were professionally involved with people with bipolar diagnoses. Semi-structured telephone interviews were undertaken by TK (see [S1 File](#)). Verbal consent was taken at the start, within the digital audio recording of the interview, and participants returned completed consent forms to the researcher after the interview.

Analysis of focus group discussions and one-to-one interviews. Focus group and interview recordings were transcribed verbatim and checked for accuracy by TK through concordant listening and reading. Data from focus group discussions and one-to-one interviews were analysed together due to the purpose of analysis being to identify all possible outcomes. Transcripts were uploaded to Dedoose online qualitative data management software [33] to manage and support data analysis. Dedoose was used to organise the qualitative data collected during the focus groups and one-to-one interviews to generate the outcome longlist. Descriptive accounts of the interviews and focus group discussions were written (TK), focusing on outcome identification. The first iteration of the coding structure was developed following thorough reading of the early transcripts and written descriptions from the focus group discussions (TK). These were open coded, line by line, and these codes were grouped into categories. Over-arching themes were identified from the categories to generate detailed outcome lists. Code formation was completed collaboratively by peer researchers and other research team members (VP, RS), drawing on personal experiences and prompting reflexive discussion and detailed conversations within the wider team. A 20 percent code application check was completed (RS).

Rapid review. A pragmatic approach was used to identify outcomes collected in bipolar research in community settings. Two researchers (TK, GT) independently performed a complete search of all pre-categorised titles listed under the bipolar reviews on the Cochrane database for systematic reviews. Cochrane reviews follow a rigorous structure and outcomes listed under the “Primary Outcomes and Secondary Outcomes” sub-headings in each review’s methods section were identified. Data relating to measures used and monitoring adverse effects or safety (unless these were included in the primary and/or secondary outcomes listed) were not extracted. Review protocols were not included in this data extraction. The lists of outcomes formed by each researcher were compared for completeness and differences in the categorisation were resolved through discussion (Database accessed in March 2015).

Phase 2: Longlist refinement

The outcomes identified in Stage 1 were checked for duplication though detail was favoured at this stage. With the input of the LEAPs, the outcomes were adapted into lay language, organised under broad headings, and merged to minimise overlap. The outcome list was then reviewed during a multi-disciplinary stakeholder meeting composed of four mental health researchers including two with personal experience of bipolar, three outcome measurement researchers, and LEAP members including three people with a bipolar diagnosis and a carer. The resulting outcome list was then reviewed by the wider PARTNERS2 research team, LEAP members and external expert advisors to consider the merging decisions to ensure the list was comprehensive to the best of their knowledge.

Phase 3: Core outcome set finalisation

The outcome longlist developed during Stages 1 and 2 were subjected to a two-round Delphi survey and a final consensus meeting.

Delphi survey. Participants were recruited from the UK only. People with a bipolar diagnosis and carers were recruited nationally through local support groups, electronic advertisement via third sector organisations and social media. The LEAPs advised on recruitment strategies and circulated recruitment materials via their own networks. Health and social care professionals and researchers were recruited through the professional networks of the PARTNERS2 research team. Purposive sampling was used to capture a range of professional roles and supplemented as required through snowball sampling.

The main eligibility criteria were that participants had experience of bipolar, due to receiving a diagnosis themselves, caring for someone who had a diagnosis, working in a professional capacity with those with bipolar, and/or a research background in bipolar, and that they could take part in both rounds of the Delphi. Following advice from the research team members with lived experience during the course of the research, a screening tool was also used to promote representativeness in the Delphi sample so that it would be more typical of the diverse population of those with bipolar (see [S2 File](#)). With their input, the screening tool was developed with particular focus on diversity of age, ethnicity, gender, and history of mental health support for participants with lived experience of bipolar. This was used to monitor sample diversity and inform and direct recruitment. To promote inclusion, a paper-based version of the survey was available upon request.

The Delphi survey was hosted by Delphi Manager software [34] and piloted with two LEAP members using cognitive appraisal techniques [35] (AR), resulting in changes to wording and providing insight into how questions were interpreted. The length of time taken to complete each round of the survey was noted to be 30 minutes, and this was included in consent information provided to potential participants. The survey design presented participants with an outcome label and an option to read a description of the outcome. These descriptions were generated by the LEAPs and meant participants could choose how much information they needed to read (see [S3 File](#)). During development of the outcome list, the outcomes were presented in domains: recovery, connectedness, mental health, physical health, self-management, medication, quality of life, service outcomes. These were developed by the research team, including research team members with lived experience, and were approved by the LEAPs. The purpose of the domains was to organise the outcomes and to promote accessibility during development. It was decided by the research team and LEAP advisors that the outcomes should be presented in these domains in the survey, for the same purpose.

During Delphi registration, participants assigned themselves to one of four stakeholder groups, 1) person with a bipolar diagnosis, 2) carer, 3) health/social care professional, 4) researcher/policy maker. Participants were requested to choose the group to which they most identified, though it was acknowledged that they may belong to more than one. They were invited to rate each of the outcomes on the longlist on a nine-point Likert scale (where nine indicated the highest level of importance and one indicated the lowest). Participants were also invited to suggest outcomes they considered were absent from the Stage 1 and 2 longlist. These were automatically included, verbatim, for rating in the Round 2. Following closure of Round 1, the software internally calculated the ratings of each outcome by stakeholder group. Participants returning for Round 2 were presented percentage distribution of scoring for each point on the scale from 1–9 from the previous round, along with their own scores for each outcome. Round 2 participants were invited to review their own ratings from the first round and consider whether they wished to change their initial score for each outcome, using the same scale.

All original outcomes presented in Round 1 were presented in Round 2. Further details pertaining to methodology can be found in the published protocol [1].

Analysis of Delphi survey data. The conditions and means for determining inclusion and exclusion were defined in advance [36]. For each outcome presented in Round 2, the proportion of participants rating 1–3, 4–6, and 7–9 on the Likert scale was calculated. Outcomes rated as 7–9 by >70% of participants and 1–3 by <25% of participants were pre-specified in our protocol to be automatically included in the COS. Outcomes automatically excluded from the COS were those which >70% of participants rated as 1–3 and <25% rated as 7–9. “Disagreement” occurred when >33% of participants scored an outcome as 1–3 and >33% scored the same outcome 7–9. These outcomes underwent additional analysis whereby their mean scores were calculated and those outcomes with a mean above 4.5 were included in the COS and those with a mean less than 4.5 were excluded. These criteria are comparable to those used throughout COS methodology [37–39].

Consensus meeting. Delphi participants were approached and invited to participate in the consensus meeting, as were those who had been unable to participate in the Delphi but had expressed an interest in attending the consensus meeting. LEAP members, members of the wider PARTNERS2 research team with professional experience as mental health professionals who had limited involvement in the COS development to date, and known contacts of the research team were also invited to participate. A screening tool was used to promote diversity (see [S2 File](#)).

The original aim of the consensus meeting was for attendees to discuss those outcomes that were in “disagreement”. Due to the large volume of outcomes that were rated as “important” by Delphi participants, the outcomes automatically included in the COS following the Delphi analysis were provisionally grouped by the research team using the domain headings in which they were presented during outcome list development and then in the Delphi survey. This is an adaption to the standard COS methodology. The proposed grouping of outcomes was finalised at a consensus meeting. Attendees voted on grouping of outcomes and rearranged them as they saw fit. Decisions made during the consensus meeting were subject to anonymous voting using TurningPoint software [40] and only those decisions sanctioned by >70% of the group attendees were ratified. However, in all cases where there was less than 100% consensus, the decisions were discussed further until those who were in disagreement were satisfied that their views had been considered and that the decision could proceed.

Results

Phase 1

Three focus group discussions were held, two with people with a bipolar diagnosis and one with carers. The groups ranged in size from 4–8 people and lasted between 96–120 minutes. Recruitment and data collection took place between July 2014 and March 2015. Fifteen people with a bipolar diagnosis with an average age of 46 years participated; 9 identified as female and 6 as male; 12 identified as White British and 3 identified as British Asian or Asian. Seven carers with an average age of 59 years participated; five of these identified as female and 2 as male; five identified as White British, 1 as British Asian or Asian, and 1 did not specify their ethnic background.

Telephone interviews were carried out with 16 healthcare professionals and researchers. Interview length ranged from 25–47 minutes. Recruitment and data collection took place between July and November 2014. Participants held multiple professional roles: 2 clinical commissioners, 2 non-clinical commissioners; 4 general practitioners; 4 healthcare management/

mental health leads; 1 mental health nurse; 5 psychiatrists; 6 researchers; 2 social workers; and 2 third sector employees.

Data were extracted from 17 bipolar reviews contained within the Cochrane database and 45 independent outcomes were identified from the bipolar database. Outcomes were classed as “independent” if the terminology used in the Cochrane database showed a clear difference. If the outcome terminology showed notable similarity, such as mortality and mortality rates, this was classed as one independent outcome. The majority of these outcomes were used in multiple reviews. Seventy-six outcomes were identified through the focus group discussions and interviews. Twenty of these were removed due to duplication.

Phase 2

One hundred and one outcomes were reviewed by the research team and LEAPs, resulting in the addition of 12 outcomes and the merging of 47 (see [S4 File](#)).

Phase 3

Fifty Delphi participants were recruited. Delphi participants were recruited between September and December 2016, during which Round 1 was open. All participants participated via the online survey. Round 2 was open from December 2016–February 2017. Ninety-three individuals were contacted via known and referred contacts of the research team and a 32% ($n = 30$) recruitment rate was achieved. Twenty participants were recruited through the LEAPs, support groups, social media, and third party organisations.

A process of monitoring and reminders was used to ensure completion, resulting in a 76% return rate ($n = 38$) to Round 2 of the Delphi (see [S5 File](#)). Fifteen people with a bipolar diagnosis participated, 2 of whom identified as male and 13 as female; 12 identified as White British, 1 as British Asian or Asian, 1 as mixed heritage, and 1 did not specify. While the screening tool was used to promote sample representativeness, nobody was excluded from the Delphi on this basis (see [S6 File](#)). Four carers participated, 1 identified as male and 3 as female; 3 identified as White British and 1 as British Asian or Asian. Twenty-three healthcare professionals participated, 12 identified as male and 11 as female; 19 identified as White British, 2 as British Asian or Asian, and 2 did not specify. Eight researchers participated, all identifying as White British, 1 identifying as male and 7 as female.

Sixty-six outcomes were included in the Delphi questionnaire and 13 outcomes were added by participants during Round 1. Three outcomes were suggested by Round 1 Delphi participants and rated as important by participants in Round 2, so were included in the consensus meeting discussion. A total of 60 outcomes met the pre-specified criteria for automatic inclusion into the COS (see [S7 File](#)), 56 from original longlist and 4 added in Round 1 by Delphi participants. Fifteen of these outcomes were merged with other outcomes by the research group and then the remaining 45 outcomes were provisionally grouped into 11 outcome domains, (see [S8 File](#)) to ensure the final COS would be feasible for use in future trials.

The consensus meeting, co-chaired by an outcome measurement researcher and a peer researcher, took place in September 2017 and was attended by 14 people: 6 healthcare professionals, 5 people with a bipolar diagnosis, 2 carers and 1 researcher. [Table 1](#) shows the results of the voting and discussion.

Final core outcome set

The outcome identification, refining and finalisation are illustrated in [Fig 1](#). The final COS included 11 outcome domains. The adapted process meant moving from detailed outcome items in the Delphi, to grouped outcomes in the consensus workshop.

Table 1. Final core outcome set, voting rounds and scores.

Outcome and explanatory items	Total number of voting rounds	Final voting results		
		Keep	Change	Discuss further
Personal Recovery Achieving goals; having a sense of identity; hope; meaning in life; empowerment; coping with self-stigma; wellbeing; self-esteem (which may overlap with “mental health and wellbeing”); and being able to build an everyday life	2	93%	7%	0%
Connectedness Satisfaction with social networks; trust; relationships with friends, family and others; social support via a person’s own social contacts; social isolation; and loneliness (“loneliness” was considered an important concept that could be an outcome itself or could overlap with “mental health and wellbeing”)	1	79%	0%	21%
Clinical recovery of bipolar symptoms A person’s increased or reduced experience of paranoia; delusions; anxiety; depression; unusual behaviour; elevated mood; and a person’s relapse or recovery response	1	100%	0%	0%
Mental health and wellbeing A person’s experience of psychological distress; and guilt and shame	1	79%	7%	14%
Physical health Relates particularly to the health concerns for people with bipolar including cardiovascular disease, metabolism concerns, or substance use but the focus of this will differ from trial to trial	1	100%	0%	0%
Self-monitoring and management Self-management and understanding of diagnosis; self-management of medication; medication adherence underpinned by satisfaction with medication; mood control and stabilisation; increasing healthy behaviours and reducing unhealthy behaviours insofar as they are linked to their impact upon bipolar	1	100%	0%	0%
Medication effects Side-effects; coping with side-effects; and satisfaction with medication	3	79%	21%	0%
Quality of life Health-related quality of life; meaningful occupation and activities; being in control of finances; personal safety and security; home living conditions and organisation; and vulnerability to harm	1	93%	7%	0%
Service outcomes There being a relapse plan in place; timely and accurate diagnosis; and number of days between referral and subsequent assessments	1	86%	7%	7%
Experience of care Dignity and respect; a person’s overall satisfaction with service; shared decision-making and control; a trusting patient and healthcare professional relationship; and active involvement of the person in their in treatment and care plan	1	100%	0%	0%
Use of coercion The use of measures such as sectioning, restraint, isolation, or seclusion to manage distress during hospital admission	1	93%	7%	0%

<https://doi.org/10.1371/journal.pone.0240518.t001>

The consensus meeting participants finalised the COS and mapped individual long list items to each outcome as follows, providing a guide for future use (Table 1).

Recommended safety indicators were: all-cause mortality, self-harm, attempted suicide, self-harm, use of emergency care and recommended outcomes for health economics evaluation included all health service use including hospital admission, home treatment, and outpatient use. All (n = 14, 100%) consensus meeting attendees voted in favour of these (see S9 File).

Discussion

The final COS consists of 11 outcome domains: personal recovery; connectedness; clinical recovery of bipolar symptoms; mental health and wellbeing; physical health; self-monitoring

and management; medication effects; service outcomes; service user experience of care; and use of coercion. The development of the COS has drawn upon several sources, including a rapid review of the Cochrane database and qualitative work with key stakeholders. We recommend that researchers use this COS to inform their selection of measures used in future community-based bipolar trials. There are a range of measures currently available for use, however, we would recommend that when choosing measures, teams do so with consultation from key stakeholders including methodologists, clinicians, and those with lived experience, in addition to considering psychometric properties of measures and their alignment with research aims. Further research is required to identify which measures should be recommended for each COS outcome, however these would require regular review and may vary due to the particular requirements of each study.

The longlist of 66 outcomes first included in the Delphi survey were rated highly by participants, and as a result, 56 (85%) of these were automatically included in the proposed COS arrangement discussed at the consensus meeting. This suggests the process of outcome identification and refining used in this work has generated a large number of outcomes that were relevant to the stakeholders involved. Bipolar is a condition that impacts on every aspect of a person's life, and thus the detailed outcome identification undertaken within the COS process reflected this extensive process and all-encompassing impact. However, use of the COS would need to be feasible in trials while retaining all outcomes rated as important by participants. Grouping items into higher-level outcome domains during the consensus meeting served this purpose. The large number of items included in the final COS was discussed at length amongst the research team and at the consensus meeting. It was felt it was important to adhere to our protocol and retain items that met the pre-specified threshold, particularly as stakeholders had strongly indicated their importance. There is potential for several items to be assessed together, for example, using health related quality of life or satisfaction questionnaires. Core outcomes sets recommend 'what' to measure; further research is required to evaluate the measurement properties of assessment tools, map these to the outcomes identified and reach consensus on the optimal ways to assess these outcomes in a standardised way.

Ongoing validation is necessary to ensure external validity; the continued relevance and importance of the outcomes; to evaluate implementation; and engage additional stakeholders [1]. This research was undertaken with inclusion of participants based in England, Wales and Scotland. Further research is required to assess the validity of the COS in specific populations, such as black, Asian and minority ethnic (BAME) communities in the UK, and to achieve greater consensus on its applicability in international settings, involving expert panels and stakeholders from the widest possible range of nations and communities. Efforts are also required to ensure the adoption and endorsement of the COS with funders, journals, and others involved in the development, facilitation and undertaking of bipolar research. Uptake of the COS [41], its implementation [42], and the consistency of its measurement [43] can be assessed through review of future community-based bipolar trials and their publication outputs.

The strengths of this research include that stakeholders were given multiple opportunities throughout to identify and remedy any gaps in the outcomes longlist and proposed core outcome set as well as the consensus generating nature of the Delphi. The relevance and importance of the COS is greatly strengthened overall through extensive lived experience input at each stage. In addition to the contribution of people with a bipolar diagnosis and carers as research participants in the qualitative component, this work has drawn upon the expertise of peer researchers employed on PARTNERS2 and LEAPs throughout the process, including in the analysis of results. This ensures the COS has validity for people directly experiencing bipolar and seeking support and treatment, as well as for health care professionals, researchers and

commissioners. Of the researchers gathering initial qualitative data, one had personal experience of bipolar; three LEAPs were consulted regularly and redefined and reworded the outcome descriptors for the Delphi survey; the Delphi and consensus meeting included equal numbers of professionals and people with lived experience of bipolar; and the consensus meeting was co-chaired by a peer-researcher with a bipolar diagnosis. Deliberations involving this full range of stakeholders on equal terms led to a more accessible and salient COS. Each outcome included in the COS is accompanied by an explanatory guide to aid interpretation and facilitate the selection of suitable measures without ambiguity about the intended meaning as understood by our participants, team members and advisors. The resulting COS aligns with accepted definitions of outcomes, indicated notably in its inclusion of outcomes relating patient experience [25].

Limitations include that the rapid Cochrane review identified outcomes used in systematic reviews which had the benefit of allowing a rapid and practical review, however the elicitation of outcomes through this methodology may not be complete as systematic reviews will not include all outcomes used in research within a given field. In addition to this, the outcomes identified in this way may not have been chosen by a range of stakeholders including those with lived experience, clinicians, and researchers.

Further limitations relate to the Delphi sample. The first is that the final sample 50 is relatively small, particularly given that four stakeholder groups were recruited. Secondly, is the issue of participant attrition between rounds during the Delphi survey. This is, however, mitigated because the data collected in the first round was included in the summarised results presented to Round 2 participants regardless of whether the corresponding participant had returned, and would have been used to inform the Round 2 ratings. Additionally, Delphi participants were asked to self-assign to one of four stakeholder groups during survey registration— 1) person with a bipolar diagnosis; 2) carer; 3) health/social care professional; 4) researcher/policy maker. During the Delphi development it was indicated that participants may identify with more than one of these categories and as such, they were invited to assign to whichever group with which they identified primarily. Open text allowed participants to further elaborate about the breadth of their experience, however these were not able to be included in analyses. Further to this, while efforts were made to ensure the diversity of the sample of individuals recruited, men and those from BAME communities, specifically those of Black, African, Caribbean, Black British origin, are underrepresented. This may be addressed with additional validation exercises with further groups and communities.

This research has used robust methodology to develop a COS for community-based bipolar trials. Its adoption in future studies will enable the generation of coherent, stakeholder-relevant outcome data that may strengthen meta-analyses and promote the value of bipolar research and the integration of subsequent findings into clinical practice.

Supporting information

S1 File.

(DOCX)

S2 File.

(DOCX)

S3 File.

(DOCX)

S4 File.

(XLSX)

S5 File.
(DOCX)

S6 File.
(DOCX)

S7 File.
(DOCX)

S8 File.
(DOCX)

S9 File.
(DOCX)

Acknowledgments

Thanks to Paula Williamson for her methodological advice throughout; Dr Ben Gray for his contribution into the qualitative data analysis; the Lived Experience Advisory Panels and the wider PARTNERS2 team for their input and support of this work.

Author Contributions

Conceptualization: Vanessa Pinfold, Thomas Keeley, Maximillian Birchwood, Melanie Calvert.

Data curation: John Gibson, Thomas Keeley, Gemma Taylor, Bliss Gibbons.

Formal analysis: Ameeta Retzer, Ruth Sayers, Vanessa Pinfold, Gemma Taylor.

Methodology: Ameeta Retzer, Vanessa Pinfold, Peter Huxley, Jonathan Mathers, Maximillian Birchwood, Melanie Calvert.

Project administration: Humera Plappert.

Supervision: Vanessa Pinfold, Humera Plappert, Jonathan Mathers, Maximillian Birchwood, Melanie Calvert.

Writing – original draft: Ameeta Retzer, Vanessa Pinfold, Thomas Keeley.

Writing – review & editing: Ameeta Retzer, Ruth Sayers, Vanessa Pinfold, John Gibson, Thomas Keeley, Gemma Taylor, Bliss Gibbons, Peter Huxley, Jonathan Mathers, Maximillian Birchwood, Melanie Calvert.

References

1. Keeley T, Khan H, Pinfold V, Williamson P, Mathers J, Davies L, et al. Core outcome sets for use in effectiveness trials involving people with bipolar and schizophrenia in a community-based setting (PARTNERS2): study protocol for the development of two core outcome sets. *Trials*. 2015; 16:47. <https://doi.org/10.1186/s13063-015-0553-0> PMID: 25887033
2. World Health Organisation: International Classification of Diseases– 10 (Version: 2016)–Bipolar Affective Disorder. <https://icd.who.int/browse10/2016/en#/F31> [Accessed 19.11.19].
3. World Health Organisation: WHO releases new International Classification of Diseases (ICD 11). [https://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-\(icd-11\)](https://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-(icd-11)) [Accessed 19.11.19].
4. Laursen TM. Life expectancy with schizophrenia or bipolar affective disorder. *Schizophrenia Research* 2011; 131: 101–4. <https://doi.org/10.1016/j.schres.2011.06.008> PMID: 21741216

5. Hoang U, Stewart R, Goldacre MJ. Mortality after hospital discharge for people with schizophrenia or bipolar disorder: retrospective study of linked English hospital episode statistics, 1999–2006. *BMJ* 2011; 343: d5422. <https://doi.org/10.1136/bmj.d5422> PMID: 21914766
6. Chang CK, Hayes RD, Perera G, Broadbent MT, Fernandes AC, Lee WE, et al. Life expectancy at birth for people with serious mental illness and other major disorders from a secondary mental health care case register in London. *PLoS One* 2011; 6: e19590. <https://doi.org/10.1371/journal.pone.0019590> PMID: 21611123
7. Hayes JF, Marston L, Walters K, King MB, Osborn DPJ. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000–2014. *The British Journal of Psychiatry* 2017; 211: 175–181. <https://doi.org/10.1192/bjp.bp.117.202606> PMID: 28684403
8. Farrelly S, Clement S, Gabbidon J, Jeffrey D, Dockery L, Lassman F, et al. MIRIAD study group. Anticipated and experienced discrimination amongst people with schizophrenia, bipolar disorder and major depressive disorder: a cross sectional study. *BMC Psychiatry*. 2014; 14:157. <https://doi.org/10.1186/1471-244X-14-157> PMID: 24885144
9. Transforming mental health through research: UK mental health research funding (2014–2017). <https://s3.eu-central-1.amazonaws.com/www.joinmq.org/UK+Mental+Health+Research+Funding+2014-2017+digital.pdf> [Accessed 27.3.19].
10. Matthijs Oud, Mayo-Wilson E, Braidwood R, Schulte P, Jones SH, Morriss R, et al. Psychological interventions for adults with bipolar disorder: systematic review and meta-analysis. *The British Journal of Psychiatry* 2016; 208:213–222. <https://doi.org/10.1192/bjp.bp.114.157123> PMID: 26932483
11. Coulman KD, Hopkins J, Brookes ST, Chalmers K, Main B, Owen-Smith A, et al. A Core Outcome Set for the Benefits and Adverse Events of Bariatric and Metabolic Surgery: The BARIACT Project. *PLoS Medicine* 2016; 13(11):e1002187. <https://doi.org/10.1371/journal.pmed.1002187> PMID: 27898680
12. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012; 13:132. <https://doi.org/10.1186/1745-6215-13-132> PMID: 22867278
13. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials*. 4th ed. New York: Springer; 2010 2010.
14. Smith H, Horobin A, Fackrell K, et al. Defining and evaluating novel procedures for involving patients in Core Outcome Set research: creating a meaningful long list of candidate outcome domains. *Research Involvement and Engagement* (2018) 4:8 <https://doi.org/10.1186/s40900-018-0091-5> PMID: 29507772
15. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Stat Med*. 1999; 18(15):1905–42. PMID: 10532877
16. Mortimer AM. Symptom rating scales and outcome in schizophrenia. *The British journal of psychiatry Supplement*. 2007; 50:s7–14. <https://doi.org/10.1192/bjp.191.50.s7> PMID: 18019038
17. Bellack AS. Scientific and consumer models of recovery in schizophrenia: concordance, contrasts, and implications. *Schizophrenia bulletin*. 2006; 32(3):432–42. <https://doi.org/10.1093/schbul/sbj044> PMID: 16461575
18. Oxford University Hospitals: Bipolar Priority Setting. <https://www.ouh.nhs.uk/research/patients/priority-setting-partnerships/bipolar/default.aspx> [Accessed 18.11.19].
19. Sinha IP, Altman DG, Beresford MW, Boers M, Clarke M, Craig J, et al. StaR Child Health Group. Standard 5: Selection, measurement, and reporting of outcomes in clinical trials in children. *Pediatrics*. 2012; 129:S146–52. <https://doi.org/10.1542/peds.2012-0055H> PMID: 22661761
20. ICF Research Branch. ICF Core Sets for Bipolar Disorders. www.icf-research-branch.org/icf-core-sets-projects2/mental-health/development-of-the-icf-core-sets-for-bipolar-disorders [Accessed 06.03.18].
21. Eiring Ø, Nylenna M, Nytrøen K. Patient-Important Outcomes in the Long-Term Treatment of Bipolar Disorder: A Mixed-Methods Approach Investigating Relative Preferences and a Proposed Taxonomy. *Patient*. 2016; 9:91–102. <https://doi.org/10.1007/s40271-015-0128-x> PMID: 25990222
22. Franklin KK, Hart JK. Idea Generation and Exploration: Benefits and Limitations of the Policy Delphi Research Method. *Innovative Higher Education* 2007; 31:237–246.
23. Ayuso-Mateos JL, Avila CC, Anaya C, Cieza A, Vieta E, and the Bipolar Disorders Core Sets Expert Group. Development of the International Classification of Functioning, Disability and Health core sets for bipolar disorders: results of an international consensus process. *Disability and Rehabilitation* 2012; 35:2138–46.
24. James Lind Alliance: Priority Setting Partnerships. Bipolar. <http://www.jla.nihr.ac.uk/priority-setting-partnerships/bipolar/> [Accessed 06.03.18].
25. Royal College of Psychiatrists (RCP) (2010) Recommended Outcome Measures for Use in Adult Psychiatry: Draft for Consultation. RCP: London.

26. Birchwood M, Calvert M, Keeley T, Pinfold V, Davies L, England E, et al. PARTNERS2L Development of a core outcome set for use in mental health trials involving people with schizophrenia or bipolar disorder in a community setting. <http://www.comet-initiative.org/studies/details/606> [Accessed 14.12.17].
27. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. *PLoS Med* 2016; 13(10): e1002148. <https://doi.org/10.1371/journal.pmed.1002148> PMID: 27755541
28. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research, *BMJ* 2017; 358:j3453. <https://doi.org/10.1136/bmj.j3453> PMID: 28768629
29. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. (2017) Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med* 14(11): e1002447. <https://doi.org/10.1371/journal.pmed.1002447> PMID: 29145404
30. Ryan L, Kofman E, Aaron P. Insiders and outsiders: working with peer researchers in researching Muslim communities. *International Journal of Social Research Methodology* 2011, 14:1, 49–60, <https://doi.org/10.1080/13645579.2010.481835>
31. Keeley T, Williamson P, Callery P, Jones LL, Mathers J, Jones J, et al. The use of qualitative methods to inform Delphi surveys in core outcome set development. *Trials*. 2016; 17(1):230. <https://doi.org/10.1186/s13063-016-1356-7> PMID: 27142835
32. Huxley P., Evans S., Madge S., Webber M., Burchardt T, McDaid D, et al. Development of a social inclusion index to capture subjective and objective life domains (phase II): psychometric development study. *Health Technology Assessment* 2012, 16 (1). pp. 1–248. ISSN 1366-5278 <https://doi.org/10.3310/hta16010> PMID: 22260923
33. Dedoose: Great Research Made Easy. <http://www.dedoose.com/> [Accessed 09.01.18].
34. COMET Initiative: Delphi Manager. <http://www.comet-initiative.org/delphimanager/> [Accessed 22.12.17].
35. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing*. 2003; 42(1); 57–63. <https://doi.org/10.1046/j.1365-2648.2003.02579.x> PMID: 12641812
36. Sinha I, 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 Med*. 2011; 8(1):e1000393. <https://doi.org/10.1371/journal.pmed.1000393> PMID: 21283604
37. Morris C, Dunkley C, Gibbon FM, Currier K, Roberts D, Rogers M et al. Core Health Outcomes In Childhood Epilepsy (CHOICE): protocol for the selection of a core outcome set. *Trials* 2017; 18:572 <https://doi.org/10.1186/s13063-017-2323-7> PMID: 29183384
38. Alkhaffaf B, Glenny A, Blazeby J, Williamson P, Bruce AB. Standardising the reporting of outcomes in gastric cancer surgery trials: protocol for the development of a core outcome set and accompanying outcome measurement instrument set (the GASTROS study). *Trials* 2017; 18:370 <https://doi.org/10.1186/s13063-017-2100-7> PMID: 28793921
39. MacLennan S, Williamson PR, Bekema H, Campbell M, Ramsay C, N'Dow J, et al. A core outcome set for localised prostate cancer effectiveness trials. *BJU International* 2017; 120: E64–E79 <https://doi.org/10.1111/bju.13854> PMID: 28346770
40. Turning Technologies. TurningPoint Software. www.turningtechnologies.co.uk [Accessed 09.01.18].
41. Kirkham JJ, Clarke M, Williamson PR. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. *BMJ* 2017; 357:j2262. <https://doi.org/10.1136/bmj.j2262> PMID: 28515234
42. Bautista-Molano W, Navarro-Compán V, Landewé RB, Boers M, Kirkham JJ, van der Heijde D. How well are the ASAS/OMERACT Core Outcome Sets for Ankylosing Spondylitis implemented in randomized clinical trials? A systematic literature review. *Clin Rheumatol*. 2014; 33:1313–22. <https://doi.org/10.1007/s10067-014-2728-6> PMID: 24970597
43. Kirkham JJ, Boers M, Tugwell P, Clarke M, Williamson PR. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials* 2013; 14:324. <https://doi.org/10.1186/1745-6215-14-324> PMID: 24103529