

# BMJ Open A protocol for a systematic review and meta-analysis to identify measures of breakthrough pain and evaluate their psychometric properties

Katie Greenfield <sup>1</sup>, Simone Holley <sup>1</sup>, Daniel Eric Schoth <sup>1</sup>, Julie Bayliss,<sup>2</sup> Anna-Karenia Anderson <sup>3</sup>, Satbir Jassal,<sup>4</sup> Dilini Rajapakse,<sup>2</sup> Lorna Katharine Fraser <sup>5</sup>, Christine Mott,<sup>6</sup> Margaret Johnson,<sup>7</sup> Ian Wong <sup>8</sup>, Richard Howard,<sup>9</sup> Emily Harrop <sup>10,11</sup> Christina Liossi <sup>1,9</sup>

**To cite:** Greenfield K, Holley S, Schoth DE, *et al.* A protocol for a systematic review and meta-analysis to identify measures of breakthrough pain and evaluate their psychometric properties. *BMJ Open* 2020;**10**:e035541. doi:10.1136/bmjopen-2019-035541

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-035541>).

Received 06 November 2019  
Revised 14 February 2020  
Accepted 17 February 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Professor Christina Liossi;  
[cliossi@soton.ac.uk](mailto:cliossi@soton.ac.uk)

## ABSTRACT

**Introduction** Breakthrough pain is common in children and adults with cancer and other conditions, including those approaching end-of-life, although it is often poorly managed, possibly partly due to a lack of validated assessment tools. This review aims to (1) identify all available instruments measuring breakthrough pain in infants, children, adolescents or adults and (2) critically appraise, compare and summarise the quality of the psychometric properties of the identified instruments using COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria.

**Methods and analysis** Two searches will be carried out between October 2019 and January 2020, one for each aim of the review. The Cochrane Library, International Prospective Register of Systematic Reviews, Embase, Cumulative Index of Nursing and Allied Health Literature, Medical Literature Analysis and Retrieval System Online (MEDLINE), PsycINFO, Web of Science Core Collection, Google Scholar, the ProQuest Dissertations & Theses Database, Evidence Search and OpenGrey databases will be searched from database inception until the date the search is conducted. Reference lists of eligible articles will be screened and authors in the field contacted. For search 1, articles will be screened by two reviewers by abstract, and full-text where necessary, to identify if a breakthrough pain assessment was used. Search 2 will then be conducted to identify studies evaluating measurement properties of these assessments. Two reviewers will screen articles from search 2 by title and abstract. All potentially relevant studies will be screened by full text by both reviewers. For search 2, data will be extracted in parallel with the quality assessment process, as recommended by COSMIN. Two reviewers will assess methodological quality using the COSMIN Risk of Bias checklist and the COSMIN updated criteria for good measurement properties. Findings will be summarised and, if possible, data will be pooled using meta-analysis. The quality of the evidence will be graded and summarised using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) guidelines.

**Ethics and dissemination** Results of this review will be submitted for publication in a peer review journal and presented at conferences.

## Strengths and limitations of this study

- This review will follow both the Preferred Reporting Items for Systematic Review and Meta-Analysis and the COnsensus-based Standards for the selection of health Measurement INstruments methodologies for conducting systematic reviews.
- Eleven databases will be searched, including those that store doctoral dissertations and other types of grey literature.
- The use of English language restriction may introduce a language bias and lead to erroneous conclusions.
- The limited literature on breakthrough pain assessment tools may prevent quantitative pooling of data (ie, meta-analysis).
- Patient and public involvement is ensured at key stages of the review, that is, development of the research question, helping to determine the appropriate inclusion and exclusion criteria and identifying search terms and relevant outcomes, and peer reviewing academic papers.

**PROSPERO registration number** CRD42019155583.

## INTRODUCTION

Even though the diagnosis, assessment, and treatment, of breakthrough pain (BTP; [figure 1](#)) has dramatically improved in recent years, BTP continues to present a significant clinical challenge for patients, their caregivers, healthcare professionals and healthcare systems. While BTP is often experienced by people with cancer<sup>1 2</sup> it is also seen in patients with other conditions,<sup>2 3</sup> including those with terminal diseases approaching end-of-life.<sup>4</sup> The reported prevalence of BTP varies significantly across studies<sup>5</sup> primarily due to variability in the definitions of BTP and methods used to assess it, study designs,



**Figure 1** Graphical representation of breakthrough pain.

settings and disease groups. In a recent systematic review of 19 studies and more than 6000 adult patients with cancer pain, the prevalence of BTP ranged from 33% to 95%, with an overall pooled prevalence of 59.2%.<sup>6</sup> That is comparable to the reported prevalence of BTP pain in adults with chronic non-cancer pain.<sup>5,7</sup> Similarly, BTP is common in children with cancer or life-limiting conditions (LLCs),<sup>8,9</sup> including those with communication and/or learning difficulties as well as pre-verbal infants, for which assessment of BTP is a particular challenge. The single small study investigating BTP in a population of children with LLCs<sup>10</sup> showed that 57% out of 27 hospitalised children with cancer who had controlled background pain reported having experienced BTP in the previous 24 hours. A retrospective review of clinical records showed that younger children (aged 7–12 years) had a significantly higher risk of BTP compared with teenagers and children and that children can experience BTP even when following a pain management strategy to control background pain.<sup>9</sup>

Despite ongoing international efforts for over 20 years, consensus on the definition of BTP has not yet been achieved.<sup>11</sup> In the cancer literature, Portenoy and Hagen<sup>12</sup> proposed the following working definition for breakthrough cancer pain (BTCP): ‘a transitory exacerbation of pain experienced by patients undergoing long-term opioid treatment for cancer-related pain whose baseline pain is adequately controlled’. However, more recent definitions do not include regular opioid medication<sup>13</sup> or background pain as prerequisites for BTCP.<sup>14</sup> The expert working group of the European Association for Palliative Care recommended the use of a simpler term such as ‘episodic’ or ‘transient’ pain, as the English term ‘breakthrough’ has no literal translation in many other languages. The WHO defines BTP in broader terms as ‘a temporary increase in the severity of pain over and above the pre-existing baseline pain level’. According to the WHO,<sup>15</sup> BTP is usually sudden, severe and occurs in brief episodes that can happen multiple times a day.<sup>15</sup> In the 11th edition of the International Classification of Diseases (ICD-11), a task force of the International Association for the Study of Pain describes cancer pain as continuous (background pain) or intermittent (episodic) pain.<sup>16</sup> The latter is defined as either related to movement or activity (incident pain) or unrelated (spontaneous pain). Nonetheless, BTCP specifically is not listed in the ICD-11. BTP can have various pathophysiologies (nociceptive,

neuropathic, mixed) and multiple causes or comorbidities. Adults with BTP report markedly reduced quality of life, functioning and productivity compared with patients with chronic pain and no BTP, and those without clinically significant pain.<sup>5</sup> Evidence suggests that this type of pain may also detrimentally affect the mental and physical well-being of patients and caregivers.<sup>17,18</sup> There is a paucity of research on short/longer-term effects of BTP in infants, children and young people.<sup>8</sup>

A recent systematic review concluded that current guidelines agree on many aspects of the management of BTCP. However, the evidence to support current guidelines remains low grade, and more research is needed in this area of care along with an international consensus on the definition and diagnostic criteria of BTCP.<sup>19</sup> It is important that patients are assessed to ensure that background pain and BTP are differentiated<sup>20,21</sup> and that BTP is not misdiagnosed as ‘end-of-dose failure’ of around-the-clock treatment, or as pain occurring when opioid medication is started or titrated.<sup>21,22</sup> While there are a number of validated pain assessment instruments, there is a lack of measures for BTP specifically.<sup>23,24</sup> These are needed since the unique features of BTP are often not picked up by existing pain assessment tools (eg, its temporal features and triggers).<sup>25</sup>

Existing BTP assessment tools for use in adults include the Breakthrough Pain Questionnaire (BPQ)<sup>12</sup>; the Alberta Breakthrough Pain Assessment Tool for Cancer Patients (ABPAT),<sup>26</sup> the Breakthrough Pain Assessment Tool (BAT)<sup>27</sup> and the Italian Questionnaire for BTP (IQ-BTP).<sup>7</sup> The ABPAT is completed by a clinician with the patient, while the BPQ, BAT and IQ-BTP are self-reported questionnaires. The ABPAT was developed primarily for use in research and is considerably longer (26 questions) than the BAT (14 questions), which was designed for use in clinical settings. However, the BAT only assesses the characteristics of previously diagnosed BTP, rather than diagnosing BTP. Oostendorp *et al*<sup>8</sup> aimed at generating reliable information on the nature and management of paediatric BTP from narrative clinical records. The authors piloted a BTP data extraction instrument based on a glossary of BTP terms developed by 13 clinicians. However, this showed poor agreement between raters, highlighting the need for clear BTP criteria, a structured way to document BTP, and validated assessment tools for use with children. BTP assessment tools designed for adults may not be suitable for children due primarily to developmental and disease differences.

Although recent reviews have investigated the prevalence<sup>6</sup> and/or management<sup>19,24,28</sup> of BTP, and one 2010 review<sup>23</sup> looked at BTP classification and assessment, to our knowledge, there are no recent published systematic reviews assessing the measurement properties of BTP assessment tools for people of all ages. It is clear that for pain to be successfully managed and treated, it is important that it is assessed, monitored and re-assessed effectively. As such, the main objectives of this systematic review are to

1. Identify all available instruments assessing self/caregiver-reported BTP in infants, children or adults.
2. Critically appraise, compare and summarise the quality of the measurement properties of the identified instruments using COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria.<sup>29</sup>

## METHODS

This protocol follows the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA) guidelines.<sup>30 31</sup> The systematic review follows the PRISMA and the COSMIN guidelines.<sup>32 33</sup> Details of the protocol are registered on the International Prospective Register of Systematic Reviews (PROSPERO) registry. Any protocol amendments will be reported in the systematic review publication. The planned start date for this review is 22 September 2019 and the planned end date is 22 March 2020.

### Eligibility criteria

Our study selection criteria were guided by the COSMIN group guidelines.<sup>29</sup>

### Inclusion criteria

1. All studies and PhD theses (since these are formally peer-reviewed) published from the inception of each database, which involve any type of BTP assessment instrument or measure (eg, questionnaire, inventory, self-report, caregiver-report) for objective 1; or that aim to develop and/or evaluate one or more measurement properties of a BTP assessment instrument for objective 2.
2. Measures might include duration, intensity and/or frequency of BTP or its impact on the individuals, for example, functional disability. All definitions of BTP will be included since there is a lack of consensus on an established definition.
3. Studies carried out worldwide that involve patient or caregiver assessment of BTP in infants, children or adults of all ages. Caregivers will include carers, family, parents and healthcare professionals.

### Exclusion criteria

1. Articles written in any language other than English.
2. Masters theses, conference abstracts, reviews.
3. Studies that involve pain assessments but do not assess BTP specifically.

### Searches

Two searches will be run, one for each objective of the review. Search 1 will aim to identify all available instruments measuring self/caregiver-reported BTP in infants, children or adults. Search 2 will aim to identify all studies that evaluate measurement properties included in the main domains (reliability, validity, responsiveness and interpretability) of the COSMIN taxonomy<sup>34</sup> (see [table 1](#)).

### Electronic sources

The Cochrane Library and PROSPERO will be searched initially to check for any existing systematic reviews on this topic. Embase, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Medical Literature Analysis and Retrieval System Online (MEDLINE) (both via Ebsco), PsycINFO, the Web of Science Core Collection and the advanced Google Scholar search facility (pdfs only, first 50 records) will then be searched. To identify any additional unpublished work, the ProQuest Dissertations and Theses Database, Evidence Search and OpenGrey will also be searched. The search strategy will include hand searching of reference lists of relevant studies for additional records. No limits will be placed on publication date when conducting the searches, with each database searched from database inception until the date on which the search is conducted. Finally, active researchers in the field who have contributed to this literature will be contacted. All searches will be run between October 2019 and January 2020.

### Search terms

Search 1 will comprise of two blocks: (1) BTP and (2) assessments. Search 1—block 1 will include the following terms: ‘breakthrough pain’ OR ‘break through pain’ OR ‘break-through pain’ OR ‘BTP’ OR ‘incident pain’ OR ‘incidental pain’ OR ‘episodic pain’ OR ‘transient pain’ OR ‘transitory pain’ OR ‘spontaneous pain’ or ‘pain flare’.

Search 1—block 2 will include terms describing assessments such as: apprais\* OR report\* OR rated OR rating\* OR assess\* OR index OR indices OR instrument\* OR measure\* OR questionnaire\* OR profile\* OR scale\* OR score\* OR status OR survey\* OR construct\* OR development\*.

Combinations of keywords, text words, Medical Subject Headings (MeSH) and other terms relevant to the review question will be chosen for each database to optimise the search sensitivity and specificity. The search strategy will be piloted and adapted for each database.

Search 2 will include two blocks: (1) terms to identify the BTP assessments found in search 1 and (2) measurement properties of these assessments. Once the BTP assessments have been identified from search 1, the terms used in search 2—block 1 will be developed.

Search 2—block 2 will include the COSMIN search filters for identifying studies on measurement properties in Medline, Embase, PsycINFO and CINAHL. Since the remainder of the databases do not have a COSMIN search filter, only block 1 will be used for the remaining databases.

### Data management

All records and data will be saved to Endnote X8.<sup>35</sup> This software will be used to identify potential duplicates. The researchers will check these and remove all confirmed duplicated references.

**Table 1** Measurement properties included in the main domains of the COSMIN taxonomy

Domain	Psychometric property	Definition
Reliability		The extent that the measurement is free from measurement error such that scores for patients who have not changed are the same under repeated measurements
	Internal consistency	The extent that items are inter-related
	Reliability	The proportion of the total variance in the measurements that is due to 'true' differences between patients (as opposed to error)
	Measurement error	Error in a participant's score that is not attributed to the construct being measured
Validity		The extent that an assessment measures what it aims to measure
	Content validity	The extent that an assessment's content reflects the construct being measured
	Face validity	The extent that an assessment looks like it reflects the construct being measured
	Construct validity	The extent that an assessment's scores are consistent with hypotheses based on the assumption that the tool measures what it purports to measure
	Structural validity	The extent that an assessment's scores reflect the dimensionality of the construct being measured
	Hypothesis testing	Construct validity for the items of an assessment
	Cross-cultural validity	The extent that items on a translated or culturally modified assessment reflect the original items
	Criterion validity	The extent that an assessment's scores represent the 'gold standard'
Responsiveness		An assessment and/or its items' ability to detect change over time in the construct being measured
Interpretability*		The extent that clinical or everyday understanding can be applied to an assessment's scores

\*Interpretability is not considered a measurement property, but an important characteristic of a measurement instrument. COSMIN, COnsensus-based Standards for the selection of health Measurement INstruments.

### Selection process

For search 1, two reviewers (KG and SH) will identify if a BTP assessment was used in each article. This will involve screening each article by abstract and by full text where necessary.

For search 2, articles will be screened by title and abstract by two reviewers (KG and SH) independently. At this stage, articles will be judged as either (a) 'not relevant' or (b) 'potentially relevant'. Inter-coder agreement will be evaluated using Cohen's kappa coefficient. A minimum kappa value of 0.75 will be taken to represent high agreement.<sup>36</sup> We will retrieve the full text of all remaining potentially relevant articles. If the relevance of a study cannot be ascertained from the abstract, then we will obtain the full text. These will be read by two reviewers (KG and SH) to make the final decision regarding inclusion. If needed, further details will be sought from authors and any uncertainties will be resolved via discussion with a third reviewer (CL).

### Data collection process and quality assessments

For search 1, if a BTP assessment was used in an article, both the article and the name of the assessment will be recorded. The article will not be assessed further unless it meets the criteria for the second objective (studies that evaluate one or more properties of a BTP instrument). For

search 2, data will be extracted in parallel with the quality assessment process, as recommended by COSMIN.<sup>29</sup>

For all included studies found from search 2, the following information will be extracted into a piloted data collection form: study aims, used definition of BTP, study population characteristics (including setting, country, language, disease characteristics, sample size, age range and gender), characteristics of the BTP assessment (including name, intended context for use, administration method, number of scales or subscales, number of items, response options, recall period and information about interpretability and feasibility) and the methodological quality ratings of each study, per measurement property, per BTP assessment tool. KG will extract this information which will be subsequently checked by DS, with any disagreements resolved through discussion with CL.

Both KG and DS will independently undertake a quality appraisal of studies included from search 2, after reading and familiarisation with the COSMIN checklist development, manual and scoring system.<sup>29 34</sup> Disagreements will be resolved by discussion between KG and DS, and CL if required. In line with the COSMIN guidelines for conducting systematic reviews of patient-reported outcome measures (PROMs),<sup>29</sup> the following steps will be taken to assess quality:

1. Methodological quality of each individual study on a measurement property will be assessed using the COSMIN Risk of Bias checklist.<sup>37</sup>
2. The result of each study on a measurement property will be assessed using the COSMIN updated criteria for good measurement properties.<sup>34</sup>
3. The quality of the evidence per measurement property per assessment will be graded and summarised.

#### Assessing risk of bias

For step 1, we will use the COSMIN Risk of Bias checklist,<sup>37</sup> which is comprised of 10 items (or boxes). One box assesses standards for PROMs and nine assess the following measurement properties: content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity (including sensitivity and specificity for dichotomous scores), hypotheses testing for construct validity and responsiveness (see [table 1](#) for definitions). Only checklist items that are assessed in each included study will be completed since not all measurement properties are evaluated in all articles. Studies will be given an overall rating of very good, adequate, doubtful or inadequate quality.

#### Assessing psychometric properties

The psychometric property of each BTP assessment will then be rated as sufficient, insufficient or inconsistent using the updated criteria for good measurement properties (see [table 2](#)).<sup>34</sup>

#### Summarising and grading the quality of the evidence

Step 3 focuses on the quality of each BTP assessment as a whole as opposed to the quality of each study, which was the focus of steps 1 and 2. This first involves reviewing the extracted data to establish whether the results of all studies of the same BTP assessment are consistent.

#### Meta-analysis of psychometric properties

Following scoping searches of current available literature together with the heterogeneity of the clinical population, pooling of data in a meta-analysis may not be possible; however, it will not be ruled out and will be assessed prior to synthesis and conducted if possible.

Initially we considered psychometric meta-analysis<sup>38 39</sup>; however, we opted for conventional meta-analytic techniques due to current disagreements in the field regarding various aspects of the technique and that simulations have reported trivial differences between the two in most cases.<sup>40</sup> Briefly, Hunter and Schmidt's meta-analytic approach<sup>38 39</sup> is based on psychometric principles and contends that a substantial portion of the variability observed in an X-Y relationship across primary-level studies is the result of artifactual sources of variance (eg, sampling error, measurement error in the dependent or criterion variable, range restriction). Consequently, to estimate better moderators of the X-Y relationship in the population, meta-analysts should correct for artifactual across-study variability by controlling it via research design and/or subtracting it from the total observed

variance. Many artefact-correction procedures and techniques have been developed and are applied with some disagreements to the optimum ones. Other differences to conventional meta-analysis include, among others, emphasis on credulity intervals rather than overall means and use of correlation coefficients instead of Fisher's Z transformed values, and the 75% rule to assess heterogeneity.

We selected Fisher's Z, which ranges from  $-\infty$  to  $+\infty$  and can be interpreted similar to a correlation coefficient, as the standardised common effect size. Depending on which measurement properties have been assessed and which estimates reported, intraclass correlations, as well as Pearson and Spearman correlations, will be converted to Fisher's Z using Fisher's variance stabilising transformation.<sup>41 42</sup> F-ratios and unstandardised beta coefficients will be converted to *r* and then to Fisher's Z.<sup>41 42</sup> Exact *p* values, if no other test statistic is available, will be converted to a standard normal deviate (Z-score), then to *r*, and then to Fisher's Z.<sup>41</sup> Studies reporting effect sizes derived from >2 follow-up assessments will be coded for both the entire study follow-up length (ie, one effect size per psychometric property, per study) and shorter, non-overlapping follow-up intervals (eg, 2–4 weeks, 4–6 weeks). Analyses will be age-stratified due to noted methodological differences between studies with younger and older children and adults.

All analyses will be completed using Comprehensive Meta-Analysis V.3.<sup>43</sup> Random-effects models will be used, which assume the average effect size varies between studies and therefore heterogeneity is to be expected. To examine heterogeneity, two statistics will be reported and examined: the Cochrane's Q test and the I<sup>2</sup> statistic. With the Cochrane's Q test, a significant result is indicative of heterogeneity of effects. Caution is advocated in the interpretation of this statistic, however, since it has low power when analyses are performed with few studies or with small sample sizes.<sup>44</sup> Instead of testing whether heterogeneity is present, the I<sup>2</sup> statistic describes the percentage of variability in effect estimates due to heterogeneity as opposed to sampling error.<sup>45</sup> A rough guide for the interpretation of the I<sup>2</sup> statistic is provided by the Cochrane Handbook: 0%–40%: might not be important; 30%–60%: may represent moderate heterogeneity; 50%–90%: may represent substantial heterogeneity; 75%–100%: considerable heterogeneity.<sup>44</sup> If evidence of significant heterogeneity exists, then, where appropriate, sensitivity analyses will be conducted to verify the robustness of meta-analytic results. Funnel plots will also be created and examined to detect potential reporting biases and small-study effects for all analyses that include data from more than 10 studies.<sup>44</sup>

If statistical pooling of results is not possible, the findings from different studies on each measurement property per BTP assessment will be qualitatively summarised, which will include a summarised result (eg, for hypothesis testing for construct validity, the percentage of results that are in accordance with the hypothesis will be listed).

**Table 2** COSMIN updated criteria for good measurement properties<sup>34</sup>

Measurement property	Rating	Criteria
Structural validity	Sufficient (+)	<b>Classical test theory (CTT):</b> Confirmatory factor analysis: comparative fit index (CFI) or Tucker-Lewis index (TLI) or comparable measure >0.95 or root mean square error of approximation (RMSEA) <0.06 or standardised root mean residuals (SRMR) <0.08 <b>Item response theory (IRT)/Rasch:</b> All of the following ▶ No violation of unidimensionality: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08) ▶ No violation of local independence: residual correlations among the items after controlling for the dominant factor <0.20 OR Q3's<0.37 ▶ No violation of monotonicity: adequate looking graphs or item scalability >0.30 and adequate model fit: IRT: $\chi^2 > 0.01$ ; Rasch: infit and outfit mean squares $\geq 0.5$ and $\leq 1.5$ OR Z standardised values > -2 and <2
	Indeterminate (?)	CTT: Not all information required for a sufficient rating is reported IRT/Rasch: Model fit not reported
	Insufficient (-)	Criteria for sufficient rating not met
Internal consistency	+	At least low evidence for sufficient structural validity (assessed using GRADE) <sup>46</sup> and Cronbach's alpha(s) $\geq 0.70$ for each unidimensional scale or subscale)
	?	Criteria for low evidence or higher not met
	-	At least low evidence for sufficient structural validity and Cronbach's alpha(s) <0.70 for each unidimensional scale or subscale
Reliability	+	Intraclass correlation coefficient (ICC) or weighted Kappa $\geq 0.70$
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa <0.70
Measurement error	+	Smallest detectable change (SDC) or limits of agreement (LoA) <minimal important change (MIC)
	?	MIC not defined
	-	SDC or LoA >MIC
Hypotheses testing for construct validity	+	75% or more of the results from all studies are in accordance with the hypotheses
	?	No hypotheses defined (by review team)
	-	Results are not in accordance with the hypotheses
Cross-cultural validity/measurement invariance	+	No important differences found between group factors (eg, age) in multiple group factor analysis or no important differential item functioning (DIF) for group factors (McFadden's $R^2 < 0.02$ )
	?	No multiple group factor analysis OR DIF analysis performed
	-	Important differences between group factors or DIF found
Criterion validity	+	Correlation with gold standard $\geq 0.70$ OR area under curve (AUC) $\geq 0.70$
	?	Not all information for sufficient rating reported
	-	Correlation with gold standard <0.70 OR AUC <0.70
Responsiveness	+	Result is in accordance with the hypothesis or AUC $\geq 0.70$
	?	No hypotheses defined (by review team)
	-	Results are not in accordance with the hypotheses or AUC <0.70

COSMIN, COnsensus-based Standards for the selection of health Measurement INstruments.

These results (quantitative and/or qualitative) will be displayed in a summary of findings table.

If the results of studies that involve the same BTP assessment are inconsistent, explanations for this will be explored and results will be summarised per subgroup of consistent results where possible. For example, an assessment may be sufficiently reliable for use in adults but not in children so results will be summarised by population age group. If no explanation for inconsistent results can be found, the rating of the majority of the results will be used.

The measurement property of each BTP assessment will then be rated as sufficient, insufficient or inconsistent using the updated criteria for good measurement properties.<sup>34</sup> Finally, the quality of the evidence (our confidence that the summarised result is trustworthy) will be rated using a modified version of the GRADE guidelines.<sup>46</sup> This involves grading the evidence as high, moderate, low or very low quality based on four factors: risk of bias, inconsistency, indirectness and imprecision. These overall ratings and GRADE assessments will be included in the summary of findings table.

## Outcomes

The main outcomes are to identify, critically appraise and compare, and summarise all available instruments assessing BTP in infants, children or adults, using COSMIN criteria.

A table of the identified BTP instruments will be presented along with a summary of findings table per measurement property, which will include the pooled results of the measurement properties, an overall rating and a grade of the quality of the evidence.

The interpretability (the extent to which clinical or commonly understood meanings can be assigned to scores) and feasibility (ease of application) of these BTP assessment tools will be described. We will then formulate recommendations for the most appropriate BTP assessment tool(s) based on the earlier findings. Where possible, these will be formulated by disease and age groups (children and adults) separately. The included assessments will be categorised into (1) the most suitable assessments (ie, those with high evidence of sufficient content validity and low evidence for internal inconsistency); (2) assessments that need further validation before they can be recommended and (3) assessments that are not recommended (ie, due to high evidence for an inadequate measurement property).

## Patient and public involvement

Funding bodies and the wider scientific community currently expect that researchers actively involve patients and the public in their research, including systematic reviews. Active patient and public involvement (PPI) within systematic reviews, in particular, has been proposed as a way to enhance the actual and perceived usefulness of synthesised research evidence, therefore addressing barriers to the uptake of evidence into practice.<sup>47</sup>

PPI is represented in the authorship of this systematic review. MJ is involved at key stages of this systematic review, that is, development of the research question, helping to determine the appropriate inclusion and exclusion criteria and identifying search terms and relevant outcomes, and peer reviewing academic papers.

MJ is one of the PPI representatives at the UK National Institute for Health Research Pain and Palliative Care Clinical Studies Group-Children and is a member of the PPI group for the Primary Care Unit Palliative Care and End of Life Research Group at the University of Cambridge Department of Public Health and Primary Care. MJ is also a trustee and board member of The Childhood Cancer Parents Alliance and a member of the National Cancer Research Institute Consumer Forum.

## DISCUSSION

To our knowledge, this will be the first systematic review to identify all available BTP instruments and review and summarise the psychometric properties of these assessments. For BTP to be effectively managed, reliable ongoing assessment is critical.

The dearth and/or heterogeneity of the included studies, which may not report on the same psychometric properties, or may lack sufficient detail on these, could limit the overall review and subsequent recommendations. The findings will be used to inform healthcare professionals on the most appropriate BTP assessment tools for different populations and will highlight gaps in BTP assessment research.

## Ethics and dissemination

Due to no patient data being accessed or collated, no research ethics is required for this systematic review. Results of this review will be submitted for publication in a peer review journal and presented at conferences.

## Author affiliations

<sup>1</sup>School of Psychology, University of Southampton, Southampton, UK

<sup>2</sup>The Louis Dundas Centre, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

<sup>3</sup>Paediatric Palliative Care, Royal Marsden Hospital NHS Trust, London, UK

<sup>4</sup>Paediatric Palliative Care, Rainbows Hospice for Children and Young People, Loughborough, UK

<sup>5</sup>Martin House Research Centre, Department of Health Sciences, University of York, York, UK

<sup>6</sup>Paediatric Palliative Care, Hummingbird House Hospice, Brisbane, Queensland, Australia

<sup>7</sup>Patient & Public Representative c/o Public Health and Primary Care, University of Cambridge, Cambridge, UK

<sup>8</sup>School of Pharmacy, University College London, London, UK

<sup>9</sup>Department of Anaesthesia and Pain Medicine, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

<sup>10</sup>Paediatric Palliative Care, Helen & Douglas House Hospice, Oxford, UK

<sup>11</sup>Paediatric Palliative Care, Oxford University Hospitals NHS Trust, Oxford, UK

**Acknowledgements** We thank Dr Lynda Brook for her contributions to this manuscript.

**Contributors** CL conceived the idea, KG, SH, DS and CL planned and designed the study protocol, data extraction and statistical analysis and wrote the first draft; ICKW, JB, SJ, DR, RH, MJ, EH, A-KA, CM and LKF provided critical insights. All authors have approved and contributed to the final written manuscript.

**Funding** This review is supported by Great Ormond Street Children's Charity and SPARKS (Grant number: V5118).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iDs

Katie Greenfield <http://orcid.org/0000-0001-8827-6543>

Simone Holley <http://orcid.org/0000-0003-4631-2862>

Daniel Eric Schoth <http://orcid.org/0000-0002-9144-8067>

Anna-Karenia Anderson <http://orcid.org/0000-0001-8542-4942>

Lorna Katharine Fraser <http://orcid.org/0000-0002-1360-4191>

Ian Wong <http://orcid.org/0000-0001-8242-0014>

Emily Harrop <http://orcid.org/0000-0002-2480-2062>

Christina Lioffi <http://orcid.org/0000-0003-0627-6377>

## REFERENCES

- 1 Alarcón M<sup>a</sup> Dolores López, Estévez FV, Cabezón-Gutiérrez L, *et al.* Expert consensus on the management of breakthrough cancer pain in older patients. A Delphi study. *J Geriatr Oncol* 2019;10:643–52.
- 2 Esparza-Miñana J-M. Diagnosis and medical treatment of breakthrough pain. *Medicina Clínica* 2018;150:114–8.
- 3 Svendsen KB, Andersen S, Arnason S, *et al.* Breakthrough pain in malignant and non-malignant diseases: a review of prevalence, characteristics and mechanisms. *Eur J Pain* 2005;9:195–206.
- 4 Pathmawathi S, Beng TS, Li LM, *et al.* Satisfaction with and perception of pain management among palliative patients with breakthrough pain: a qualitative study. *Pain Manag Nurs* 2015;16:552–60.
- 5 Narayana A, Katz N, Shillington AC, *et al.* National breakthrough pain study: prevalence, characteristics, and associations with health outcomes. *Pain* 2015;156:252–9.
- 6 Deandrea S, Corli O, Consonni D, *et al.* Prevalence of breakthrough cancer pain: a systematic review and a pooled analysis of published literature. *J Pain Symptom Manage* 2014;47:57–76.
- 7 Samolsky Dekel BG, Remondini F, Gori A, *et al.* Development, validation and psychometric properties of a diagnostic/prognostic tool for breakthrough pain in mixed chronic-pain patients. *Clin Neurol Neurosurg* 2016;141:23–9.
- 8 Friedrichsdorf SJ, Postier A. Management of breakthrough pain in children with cancer. *J Pain Res* 2014;7:117–23.
- 9 Oostendorp LJ, Rajapakse D, Kelly P, *et al.* Documentation of breakthrough pain in narrative clinical records of children with life-limiting conditions: feasibility of a retrospective review. *J Child Health Care* 2019;23:564–78.
- 10 Friedrichsdorf SJ, Finney D, Bergin M, *et al.* Breakthrough pain in children with cancer. *J Pain Symptom Manage* 2007;34:209–16.
- 11 O'Hagan P, Mercadante S, O'Hagan P. Breakthrough cancer pain: the importance of the right treatment at the right time. *Eur J Pain* 2018;22:1362–74.
- 12 Portenoy RK, Hagen NA. Definition breakthrough pain: prevalence and characteristics. *Pain* 1990;41:273–81.
- 13 Mercadante S, Lazzari M, Reale C, *et al.* Italian oncological pain survey (IOPS): a multicentre Italian study of breakthrough pain performed in different settings. *Clin J Pain* 2015;31:214–21.
- 14 Lohre ET, Klepstad P, Bennett MI, *et al.* From "Breakthrough" to "Episodic" Cancer Pain? A European Association for Palliative Care Research Network Expert Delphi Survey Toward a Common Terminology and Classification of Transient Cancer Pain Exacerbations. *J Pain Symptom Manage* 2016;51:1013–9.
- 15 World Health Organization. *Who guidelines on the pharmacological treatment of persisting pain in children with medical illnesses*. Geneva: World Health Organization, 2012: 166.
- 16 Bennett MI, Kaasa S, Barke A, *et al.* The IASP classification of chronic pain for ICD-11: chronic cancer-related pain. *Pain* 2019;160:38–44.
- 17 Bunn R, Griffiths M. Understanding and managing breakthrough pain. *J Community Nurs* 2011;25:25–9.
- 18 Hjermstad MJ, Kaasa S, Caraceni A, *et al.* Characteristics of breakthrough cancer pain and its influence on quality of life in an international cohort of patients with cancer. *BMJ Support Palliat Care* 2016;6:344–52.
- 19 Davies AN, Elsner F, Filbet MJ, *et al.* Breakthrough cancer pain (BTCP) management: a review of international and national guidelines. *BMJ Support Palliat Care* 2018;8:241–9.
- 20 Davies AN, Dickman A, Reid C, *et al.* The management of cancer-related breakthrough pain: recommendations of a task group of the science Committee of the association for palliative medicine of great Britain and Ireland. *Eur J Pain* 2009;13:331–8.
- 21 Daeninck P, Gagnon B, Gallagher R, *et al.* Canadian recommendations for the management of breakthrough cancer pain. *Curr Oncol* 2016;23:96–108.
- 22 Simmonds MA. Management of breakthrough pain due to cancer. *Palliative & Supportive Care* 1999;13:1103–8.
- 23 Haugen DF, Hjermstad MJ, Hagen N, *et al.* Assessment and classification of cancer breakthrough pain: a systematic literature review. *Pain* 2010;149:476–82.
- 24 Mercadante S, Marchetti P, Cuomo A. *Breakthrough pain and its treatment: critical review and recommendations of IOPS (Italian oncologic pain survey) expert group*, 2015.
- 25 Vellucci R, Fanelli G, *et al.* Working Group Nientemale DEI. What to do, and what not to do, when diagnosing and treating breakthrough cancer pain (BTCP): expert opinion. *Drugs* 2016;76:315–30.
- 26 Hagen NA, Stiles C, Nekolaichuk C, *et al.* The Alberta breakthrough pain assessment tool for cancer patients: a validation study using a Delphi process and patient think-aloud interviews. *J Pain Symptom Manage* 2008;35:136–52.
- 27 Webber K, Davies AN, Zeppetella G, *et al.* Development and validation of the breakthrough pain assessment tool (bat) in cancer patients. *J Pain Symptom Manage* 2014;48:619–31.
- 28 Brant JM, Rodgers BB, Gallagher E, *et al.* Breakthrough Cancer Pain: A Systematic Review of Pharmacologic Management. *Clin J Oncol Nurs* 2017;21:71–80.
- 29 Prinsen CAC, Mokkink LB, Bouter LM, *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
- 30 Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
- 31 Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
- 32 Mokkink LB, Prinsen CAC, Bouter LM, *et al.* The consensus-based standards for the selection of health measurement instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105–13.
- 33 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
- 34 Mokkink LB, Prinsen C, Patrick DL, *et al.* COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual* 2018;78:1.
- 35 Endnote Referencing Software. [program] Version X8. New York, NY: Clarivate Analytics, 2017.
- 36 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- 37 Mokkink LB, de Vet HCW, Prinsen CAC, *et al.* COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171–9.
- 38 Hunter JE, Schmidt FL. *Methods of meta-analysis: correcting error and bias in research findings*. Newbury Park, CA: Sage, 2004.
- 39 Hunter JE, Schmidt F. *Methods of correcting error and bias in research findings*. Newbury Park, CA: Sage, 1990.
- 40 Borenstein M, Hedges LV, Higgins JP, *et al.* *Introduction to meta-analysis*. John Wiley & Sons, 2011.
- 41 Metawin 2.0. *User's manual: statistical software for meta-analysis*. Sinauer Associates, 2000.
- 42 Rosenthal R, Cooper H, Hedges L. Parametric Measures of Effect Size. In: Cooper H, Hedges LV, eds. *The Handbook of research synthesis*. New York: Sage, 1994: 231–44.
- 43 Comprehensive meta-analysis. [program] Version 3: Biostat, 2015.
- 44 The Cochrane Collaboration. *Cochrane Handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. John Wiley & Sons, 2011.
- 45 Higgins JPT, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
- 46 Guyatt GH, Oxman AD, Schünemann HJ, *et al.* Grade guidelines: a new series of articles in the Journal of clinical epidemiology. *J Clin Epidemiol* 2011;64:380–2.
- 47 Wallace J, Nwosu B, Clarke M. Barriers to the uptake of evidence from systematic reviews and meta-analyses: a systematic review of decision makers' perceptions. *BMJ Open* 2012;2:e001220.