

Documentation of breakthrough pain in narrative clinical records of children with life-limiting conditions: Feasibility of a retrospective review

Journal of Child Health Care

1–15

© The Author(s) 2018



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1367493518807312

journals.sagepub.com/home/chc

Linda JM Oostendorp¹ , Dilini Rajapakse², Paula Kelly², Joanna Crocker^{1,3}, Andrew Dinsdale², Lorna Fraser⁴ and Myra Bluebond-Langner¹

Abstract

This study explored the feasibility of generating reliable information on the frequency, nature and management of breakthrough pain (BTP) in children with life-limiting conditions and life-threatening illnesses (LTIs) from narrative clinical records. In the absence of standardized ways for documenting BTP, we conducted a consensus exercise to develop a glossary of terms that could denote BTP in the records. Thirteen clinicians who contributed to the records reached consensus on 45 terms which could denote BTP, while emphasizing the importance of contextual information. The results of this approach together with guidance for improving the reliability of retrospective reviews informed a data extraction instrument. A pilot test of this instrument showed poor agreement between raters. Given the challenges encountered, we do not recommend a retrospective review of BTP using narrative records. This study highlighted challenges of data extraction for complex symptoms such as BTP from narrative clinical records. For both clinical and research purposes, the recording of complex symptoms such as BTP would benefit

¹ Louis Dundas Centre for Children's Palliative Care, UCL Institute of Child Health, University College London, London, UK

² Louis Dundas Centre for Oncology Outreach and Children's Palliative Care, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

³ Health Experiences Institute, University of Oxford, London, UK

⁴ Department of Health Sciences, University of York, York, UK

Corresponding author:

Myra Bluebond-Langner, UCL Great Ormond Street Institute of Child Health, Faculty of Population Health Sciences, Louis Dundas Centre for Children's Palliative Care, 30 Guilford Street, London WC1N 1EH, UK.

Email: Bluebond@ucl.ac.uk

from clear criteria for applying definitions, a more structured format and the inclusion of validated assessment tools. This study also showed the value of consensus exercises in improving understanding and interpretation of clinical notes within a service.

Keywords

Breakthrough pain, child, data collection/instrumentation, life-limiting conditions, retrospective review

Introduction

Retrospective reviews of clinical records are a relatively inexpensive way to generate greater insight in unexplored areas of medical research and inform subsequent prospective studies (Gearing et al., 2006; Hellings, 2004). Extracting useful information from clinical records, however, presents various challenges including individual variation in documentation and inconsistent use of terminology (Samuels, 2012; Worster and Haines, 2004). In emerging areas of practice, such as paediatric palliative care, shared descriptors of signs and symptoms may not even exist (Craig et al., 2015; Liben et al., 2008; Siden et al., 2012).

Reports from professionals and parents indicate that pain is frequently unsatisfactorily treated in children with life-limiting conditions (LLCs) or life-threatening illnesses (LTIs), particularly towards the end of life (Drake et al., 2003; Goldman et al., 2006; Tomlinson et al., 2011; Wolfe et al., 2000). One of the most difficult types of pain to manage is breakthrough pain (BTP). Even when an active pain management strategy is in place to alleviate background pain, children can experience BTP. The WHO definition of BTP is provided in Figure 1.

In brief, BTP is characterized as ‘a temporary increase in the severity of pain over and above the pre-existing baseline pain level’ (World Health Organization (WHO), 2012). Importantly, most evidence on pain assessment and management in children relates to post-operative pain and caution is required in extrapolating this evidence to children with LLCs where clinicians anticipate pain escalation as the disease progresses. The single study investigating BTP in a population of children with LLCs (Friedrichsdorf et al., 2007) showed that 57% out of 27 hospitalized children with cancer who had controlled background pain reported to have experienced BTP in the previous 24 hours. While this study is a useful starting point, further larger-scale prospective studies on the frequency, nature and management of BTP in the entire population of children with LLCs or LTIs are needed to develop effective interventions.

Clinical records detailing children’s pain symptoms and the strategies used to alleviate these are a potential source of data for baseline information. As a first step in expanding the evidence base on the frequency, nature and management of BTP in children with LLCs or LTIs, we undertook a retrospective review of the clinical records of children cared for by the largest specialist paediatric palliative care service in the United Kingdom, giving access to a large and diverse cohort of children. While clinical records are recorded electronically within a bespoke Access database and comprise structured elements, the service has not implemented structured electronic health records yet. Documentation of clinical assessments carried out by team members (clinical nurse specialists, medical registrars and consultants) in the home, hospital, or hospice, or via telephone consultations are reported in a traditional narrative format. In the absence of standardized measures and shared descriptors for documenting BTP, we decided to conduct a consensus exercise, using features from

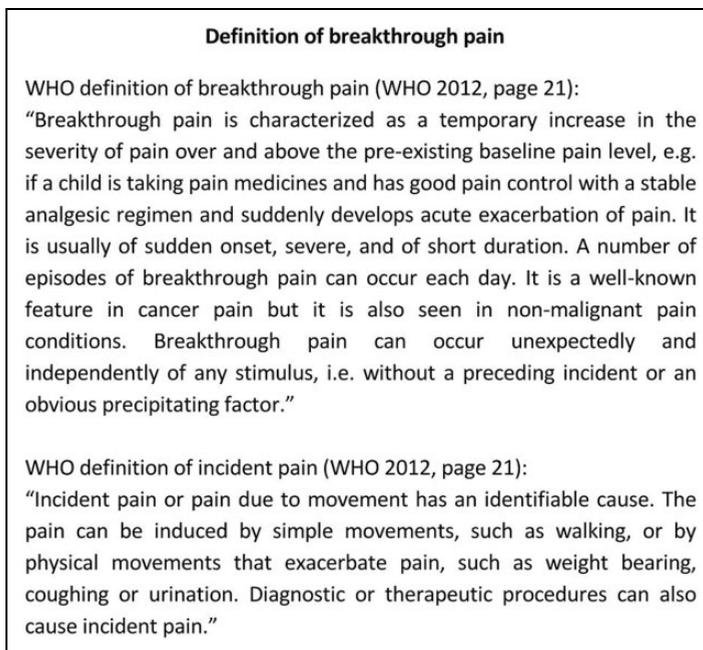


Figure 1. Definition of BTP as used in this study.

both the Delphi consensus technique and the nominal group technique, to develop a glossary of terms that could denote BTP in the records of this service. In this article, we report on the results of the consensus exercise as it relates to the development of a data extraction instrument for a retrospective review of clinical notes. We also report on the feasibility of generating reliable information about BTP when applying this instrument to clinical records. Finally, we report on how such consensus exercises can inform both current practice within a clinical team and design of prospective studies on a particular symptom or set of symptoms.

Methods

Design

The aim of the study was to explore the feasibility of generating reliable information on the frequency, nature and management of BTP in children with LLCs or LTIs from narrative clinical records to inform current practice and the design of a prospective study. The study consisted of three parts, and the first part was a consensus exercise among clinicians to create a glossary of terms that could denote BTP in the narrative clinical records of children with LLCs or LTIs. Importantly, we did not aim to develop a definition of BTP itself (we used the WHO definition as described above) rather we sought consensus on how clinicians interpret and apply this definition and how they may document BTP in narrative clinical records. In the second part of the study, we developed a data extraction instrument informed by the results of the consensus exercise and published guidance on data extraction instruments (Eder et al., 2005; Engel et al., 2009; Gregory and Radovinsky, 2012; Jansen et al., 2005). In part three of the study, a pilot test of the data

extraction instrument was performed to assess interrater reliability (Lilford et al., 2007; Yawn and Wollan, 2005).

Participants

Consensus exercise. Since we aimed to conduct a retrospective review of the records kept within a particular service, the most appropriate participants for the consensus exercise were deemed to be clinicians contributing to these records. All six nurses and seven out of eight doctors from the service were invited to participate in the consensus exercise (one doctor was a member of the study team).

Pilot test of the data extraction instrument. Two raters independently used the data extraction instrument to identify episodes of BTP in the records of children with LLCs or LTIs. This included a research nurse and a palliative care consultant (DR), both experienced in caring for children with LLCs or LTIs. Data were extracted from the records of 52 children with an LLC or LTI, selected from the service's caseload between 1 July 2012 and 30 June 2014.

Questionnaire development

An initial list of terms and phrases that could potentially denote BTP in the clinical records was prepared for the clinicians to give feedback on. Two clinicians working within the service (DR and Liz Rasdall) and a member of the study team with both clinical and research experience (PK) independently reviewed the records of 20 children to extract any terms and phrases potentially used to denote BTP. These 20 children were a sample of the children who had died within the last three years, from malignant and non-malignant conditions. A total of 63 terms and phrases were identified and divided into five categories: (i) symptoms reported by patients, parents or healthcare professionals (comprising a list of pain descriptors) (15 items); (ii) behaviours noted in the child (18 items); (iii) neurological symptoms (3 items); (iv) medications (24 items); and (v) other treatments (3 items) (see Online Supplemental File 1).

In the online questionnaire (1A/B), clinicians were asked to rate whether they felt the presence of each of these items would indicate BTP ('yes', 'no', 'possibly') (Online Supplemental File 2). Clinicians could also suggest additional terms/phrases. Participants were provided with a definition of BTP encompassing the WHO's definition of both BTP and incident pain in children with medical illness suffering from persistent pain (WHO, 2012). In including incident pain in our definition, we considered that definitions vary as to whether incident pain is included (Davies et al., 2009; Friedrichsdorf and Postier, 2014) or not (WHO, 2012) and decided to perform an inclusive review of BTP in children with LLCs, where identifying the underlying cause of the pain can inform selection of a management strategy (see Figure 1 for the full definitions).

In the second questionnaire (2A/B), we included contextual information alongside the items selected from the first questionnaire (see Figure 2).

For example, whereas questionnaires 1A/1B showed terms such as 'achiness', contextual factors were added to the terms and phrases in questionnaires 2A/2B to facilitate discrimination of BTP (e.g. 'sudden occurrence of achiness'; see Online Supplemental Material File 1 for an overview of the contextual factors added). Participants were asked to record for each of the terms/phrases whether they would indicate BTP ('yes' or 'no') in the context provided (Online Supplemental File 3).

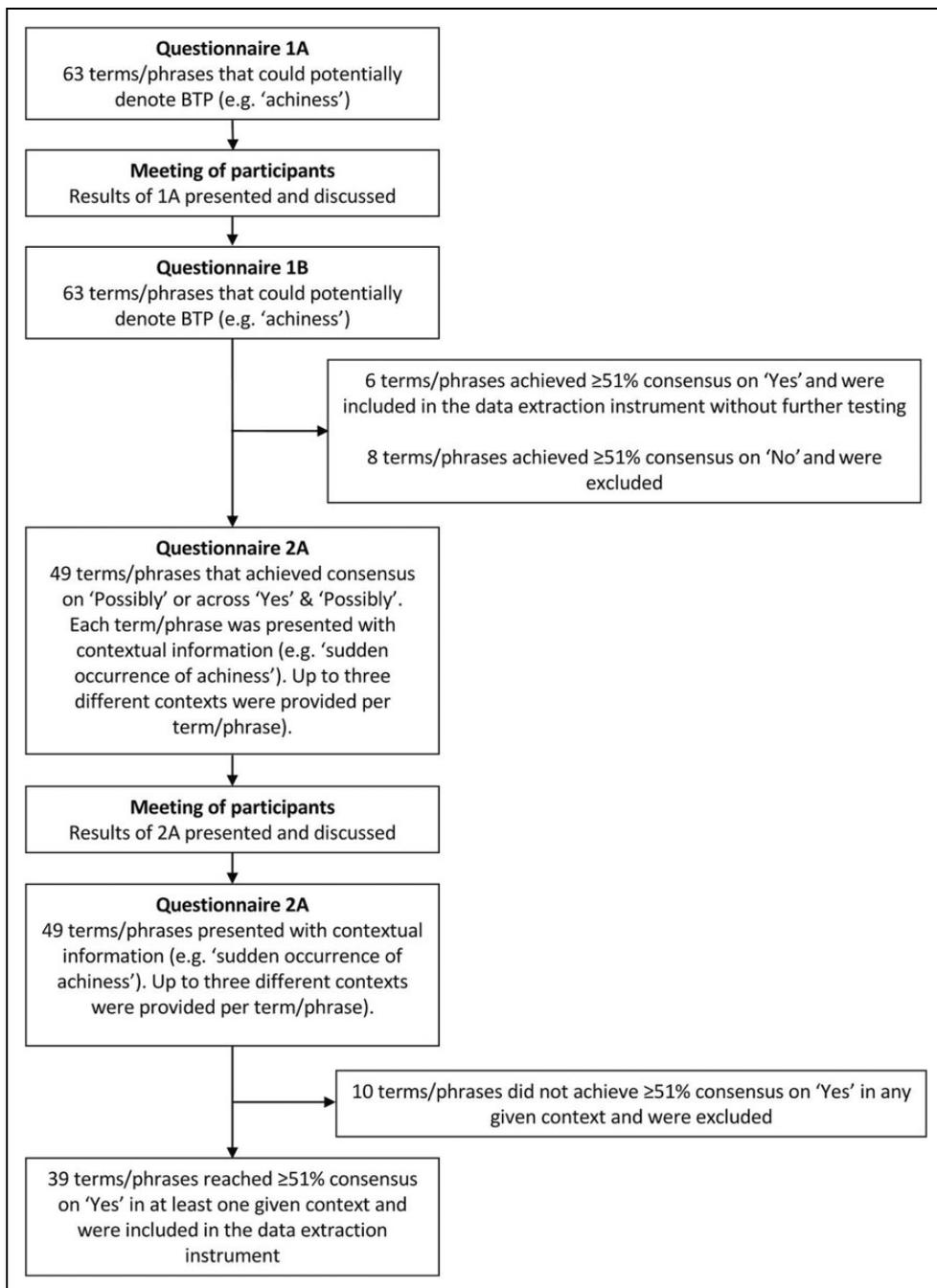


Figure 2. Flow diagram of consensus process.

Development of data extraction instrument

An electronic data extracting instrument (see Online Supplemental File 4) and accompanying manual (see Online Supplemental File 5) were developed. The data extraction instrument was designed to identify occurrences of BTP in three steps (see Figure 3): (i) identification of pain episodes using a glossary of 28 pain descriptors developed through the consensus exercise; (ii) selection of pain episodes with evidence of controlled background pain; and (iii) confirmation of the presence of BTP through identification of contextual information (derived from the consensus exercise results).

The electronic instrument contained prompts to aid data extraction and provide explanations, and fields were made mandatory where applicable. Wherever possible, dropdown lists were generated to reduce typing and errors.

Data collection

Consensus exercise. The consensus process is depicted in Figure 2. Participants in the consensus exercise both completed questionnaires (as also used in the Delphi consensus technique) and participated in face-to-face meetings (which had features of the nominal group technique; McMillan et al., 2016; see Figure 2). Clinicians completed four anonymous online questionnaires at their convenience (rounds 1A/1B/2A/2B). Two face-to-face meetings were held to enable clinicians to discuss areas of disagreement and help the study team to better understand their views (James and Warren-Forward, 2015; Jones and Hunter, 1995; Vakil, 2011). These meetings were facilitated by an external clinical psychologist and highly structured to minimize the risk of one or more participants dominating discussions and to make sure all relevant issues were presented and discussed. To optimize attendance, these meetings were part of the clinicians' weekly handover meeting. After these meetings, participants were invited to recomplete the questionnaire (rounds 1B and 2B). Participants were provided with a copy of the results emphasizing that they need not conform to the prevailing view. An observer (JC/PK) took field notes to capture the content and tenor of discussions.

Pilot test of the data extraction instrument. In the pilot test, two raters independently extracted data using the data extraction instrument to assess interrater reliability and identify any issues that could impede a full retrospective review. In part 1, raters extracted data from 21 clinical records that were selected randomly, stratified by year to factor in changes to documentation styles, and were checked to ensure both malignant and non-malignant conditions were included. The study team, including the two raters, discussed disagreements and agreed refinements of the data extraction instrument to reflect the insights gained. The same two clinicians extracted data from another 31 randomly selected clinical records to conduct a further test of interrater reliability. A sample size calculation showed that, assuming the null hypothesis value of κ to be 0.0, 22–30 cases would be required to detect a statistically significant κ of 0.6 or 0.7 with a power of 90% (Sim and Wright, 2005).

Analysis

Consensus exercise. After each round of the consensus process, a decision was made whether to include items in the data extraction instrument (see Figure 2). Items were included in the instrument if at least 51% of participants selected 'yes' in round 1B or 2B. A low threshold of 51% was

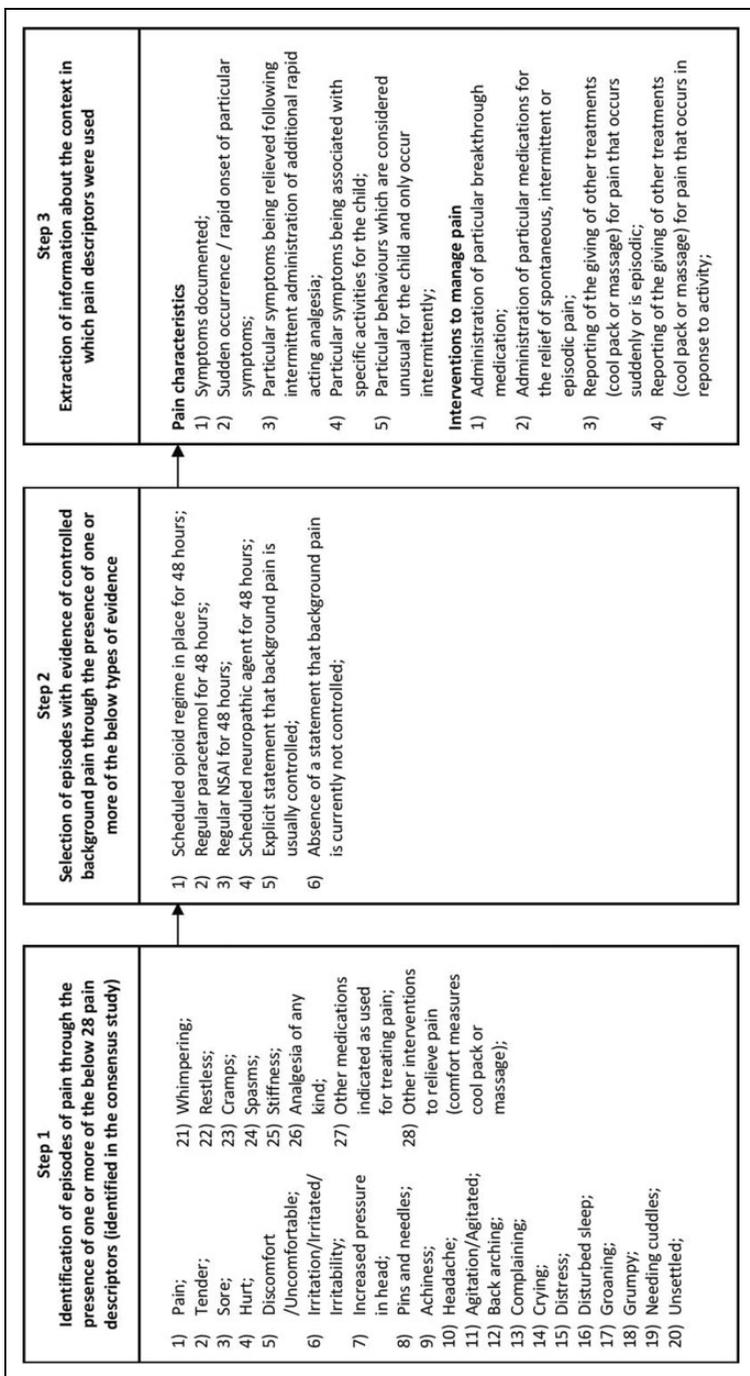


Figure 3. Overview of data extraction procedure. This Figure shows the data extraction instrument that was used in the second part of the pilot study, after refinements were made following the first part of the pilot study.

considered appropriate, as we aimed to develop an inclusive instrument with a high sensitivity to detect potential occurrences of BTP.

Pilot test of the data extraction instrument. In the pilot test, agreement between raters was assessed using Cohen's κ . The analysis included variables related to patient characteristics and the three steps in identifying BTP (see Figure 3). While the minimum required value of κ is arbitrary, values of 0.6 and 0.7 (moderate to substantial agreement) are often reported (Landis and Koch, 1977). We used a value of 0.7 for episodes of pain identified, the presence of evidence of controlled background pain and the presence of contextual information to confirm that the pain was actually BTP. A more lenient value of 0.6 was used for agreement on more detailed information, such as specific pain descriptors.

Ethical considerations

The study was approved by the Hospital and University Joint Research and Development Office (13LC02; 12-12-2013). Written informed consent was sought from clinicians participating in the consensus exercise. Data were extracted from children's records by members of the palliative care team and were pseudonymized before being transferred to researchers outside of the care team for analysis. Seeking parental consent was not deemed necessary for this retrospective review.

Results

Consensus exercise

Participants. Out of 13 clinicians invited, 7–11 clinicians (54–85%) completed questionnaires and attended facilitated discussions. These included three consultants, three specialist registrars, one locum specialist, five clinical specialist nurses and one nurse consultant.

Round 1. In round 1A, participants reached consensus for 49 out of 63 items (78%; see Online Supplemental File 1). Two participants provided comments, including 'sudden onset of pain' and 'all words are dependent on a thorough history and on who is reporting and who is taking history therefore entirely subjective'.

In the facilitated discussion participants commented that they needed more contextual information to determine whether a term or phrase would indicate BTP. This contextual information included temporal terms, underlying cause of pain, age, baseline behaviour, symptom management and presence of (other) pain descriptors.

In round 1B, participants reached consensus on all items; 6 items (10%) would denote BTP, 8 items (13%) would not denote BTP and 49 items (78%) would possibly denote BTP, depending on context. Two comments were provided, including 'Consider temporal terms in relation to pain' and 'Using a temporal relationship to pain descriptors may help to define terminology'.

Round 2. Contextual information was added to the remaining 49 items (see Online Supplemental File 1). In round 2A, clinicians reached consensus that all items in one or more given contexts would denote BTP. Five comments were provided, all stating that while a term could possibly denote BTP, it was difficult to say this with certainty and more information about the context was required. Two contradictory comments were made about the definition of BTP as used in this

study, one recommending to include incident pain in the definition of BTP and the other recommending against it.

The facilitated discussion provided further insight into the challenges still present even with additional contextual information. Clinicians indicated deciding between yes and no was challenging, with some stating they had responded to uncertainty with 'yes'. In addition, some distinction was made between phenomena that could evoke pain (e.g. spasms) and those where pain was accepted as the underlying causal mechanism for the behaviour demonstrated. Some respondents advocated more contextual information (e.g. further information on pain history). As with the first facilitated discussion, there was considerable discussion on symptoms associated with neuropathic pain.

In round 2B, consensus was achieved such that 39 items could denote BTP in at least one given context (80%). Ten items (20%) did not reach consensus on 'yes' in any given context and were excluded.

Pilot test of data extraction instrument

In the first part of the pilot test, data were extracted from the clinical records of 21 children with a malignant (24%) or non-malignant (76%) condition. The two extractors identified 82 episodes of pain; only 28 were identified by both (34% agreement; $\kappa = -0.475$; $p < 0.001$). Following discussions, the data extraction instrument was refined by explicating the first pain descriptor 'pain of any description', amalgamating similar pain descriptors and adding a list of relevant medications and interventions to the database (see Online Supplemental File 4).

In the second part of the pilot test, data were extracted from the clinical records of 31 children with a malignant (23%) or non-malignant (77%) condition. The two extractors identified 121 episodes of pain, of which 58 were identified by both (48% agreement; $\kappa = -0.346$; $p = 0.001$).

Discussion

Main findings

In this study, we investigated the feasibility of a retrospective review of the rich information in the records of children with LLCs or LTIs to obtain more insight into BTP. Using a structured consensus approach, we found that clinicians recognized a wide range of terms and phrases that could be used to document BTP in narrative records, although single word descriptors were usually considered insufficient without further contextual elements. A pilot test of a structured data extraction instrument based on the terms and phrases identified showed that even with detailed guidance in place, interrater reliability in identifying episodes of BTP from the narrative clinical records was unsatisfactory. Data extracted under these conditions are unlikely to provide reliable information about the frequency, nature and management of BTP (Landis and Koch, 1977), therefore we concluded that a retrospective review of narrative clinical records was not feasible.

Findings in context of literature

Assessing interrater reliability should be a routine component of any data extraction study but is often overlooked (Allison et al., 2000). Recent retrospective reviews in the population of children with LLCs and LTIs include reviews of symptoms (Jagt-van Kampen et al., 2015), pain medication (Orsey et al., 2009; Schiessl et al., 2008) and care provided (Feudtner et al., 2003; Ho and

Straatman, 2013). While some of these studies described measures taken to improve data quality, none described a reliability assessment. Several retrospective reviews in adult patient populations that did include an assessment of interrater reliability have demonstrated problems (Goulet et al., 2007; Meschia et al., 2006; Reeves et al., 2008; Thomas et al., 2002). Achieving satisfactory interrater reliability relies on three key components: (i) the data extraction instrument, (ii) the data extractors and (iii) the source of information from which the data are being extracted.

- (i) The content of the data extraction instrument was based on the results of the consensus exercise and the format followed published guidance (Eder et al., 2005; Engel et al., 2009; Gregory and Radovinsky, 2012; Jansen et al., 2005). Terms and phrases to be used in round 1 were derived from a review of children's clinical records. A Canadian study among hospitalized children provided a much more extensive list of terms and phrases to describe any pain (e.g. including physiological indicators), but there was considerable overlap with the terms and phrases identified in our study (Rashotte et al., 2013). The fact that such a wide range of terms/phrases was identified may have contributed to the difficulties in reliably extracting episodes of BTP. The use of ambiguous or complicated definitions as a cause of poor interrater reliability has been suggested previously in retrospective reviews in adult populations (Reeves et al., 2008).
- (ii) The pilot test to assess the reliability of extracted data involved the research nurse and a consultant, who worked in the clinical team and was responsible for writing notes. Both data extractors were clinically experienced, received training and recorded any issues regarding data extraction. In recent retrospective reviews in the population of children with LLCs or LTI, data were either extracted by one person (with no reliability assessment reported; Ho and Straatman, 2013; Jagt-van Kampen et al., 2015), or no information on who extracted data was provided (Chang et al., 2015; Feudtner et al., 2003; Orsey et al., 2009; Schiessl et al., 2008). Researchers seem to underestimate the challenges involved in extracting high quality data from clinical records (Allison et al., 2000). More methodological details need to be reported to allow readers to assess the quality of retrospective reviews and to enable researchers to learn from the experiences of others.
- (iii) The narrative clinical notes in this study were from a specialist paediatric palliative care team that cares for children with LLCs or LTIs both in hospital and in the community. Clinical notes may be completed when a team member has assessed a child in hospital, at home or in a hospice. They may also be completed after a telephone consultation with a parent or healthcare professional. These proxy assessments result in children's pain experiences being further filtered through other reporters (Twycross et al., 2015) prior to documentation in the clinical records. Similar issues have been noted in other retrospective reviews in this population, including discontinuity in terms of location of care, and lack of documentation who assessed the child using what methods, and whether consultations were performed by phone or face to face (Jagt-van Kampen et al., 2015; Kelly et al., 2018; Thrane et al., 2017). Suboptimal documentation was also discussed in retrospective reviews in adult populations, and it was suggested that this could be improved by using additional data sources (e.g. meeting the clinician who recorded the notes; Goulet et al., 2007), but this may not be feasible.

An additional challenge is the complexity of BTP. It is by definition intermittent and often unpredictable such that reporting and recording will vary. As highlighted in the consensus exercise,

documentation of contextual factors was regarded as essential in determining the presence of BTP. Another factor possibly contributing to poor interrater reliability are the inconsistencies in the narrative nature of the notes and the lack of a routinely used structured pain assessment rating scale (Chang et al., 2015; Hunt, 2012). Correspondingly, a lack of structure has been highlighted in previous studies on documentation in palliative care (Curtis et al., 2018; Furuno et al., 2017; Gunhardsson et al., 2008; McEvoy, 2000; Stewart et al., 2017; Yawn and Wollan, 2005). In one study, pain was commonly noted but different terminology was used, with pain characteristics often lacking (Gunhardsson et al., 2008). The authors emphasized the impact of accurate documentation on quality of care, symptom control and effective handover (Gunhardsson et al., 2008; McEvoy, 2000). The use of guidelines, checklists and more structured forms have been suggested to improve documentation (de la Cruz et al., 2016; Gunhardsson et al., 2008; McEvoy, 2000; Stewart et al., 2017). A study among children's nurses showed that the majority of these nurses agreed that the introduction of pain assessment tools across the hospital would improve documentation (Simons and Macdonald, 2004).

Strengths and limitations

While recognizing the potential value of the information contained in clinical records for research purposes, we were aware of the challenges in extracting reliable data and decided to use a structured process in developing a data extraction instrument, using clinician feedback and published guidance, and assess the reliability of data extracted. While extracting reliable data on BTP from children's narrative clinical records remains elusive, this study is one of the first to investigate the description of pain in children's clinical records. A Canadian study has investigated the description of pain in children's clinical records but this included any pain, not specifically BTP, and the study was limited to a 24-hour period for inpatients in Canadian paediatric hospitals (Rashotte et al., 2013). Our study specifically focused on BTP in children with malignant and non-malignant life-limiting illnesses, cared for by a paediatric palliative care team, adding the complexity of multiple settings. Admittedly, this study was confined to a single tertiary palliative care service and therefore terms and phrases identified in this service may not be recognized or utilized in narrative records by other healthcare professionals in this field. Moreover, other services may have adopted structured electronic health records, which would have implications for the structure of notes and consequently the design of retrospective reviews. However, there are indications that even when structured electronic records are used, clinicians may still resort to narratives to document pain episodes in detail (Rashotte et al., 2013). Finally, one could argue that to develop a truly inclusive data extraction instrument, a threshold higher than 51% should have been used for excluding terms/phrases. However, for most of the 18 items that were excluded, at least 70% of participants did not think that this term/phrase in the given context would indicate BTP.

What this study adds

Several lessons can be learned from this feasibility study. First, the consensus exercise yielded important insights into BTP that can benefit future studies. Even with a WHO definition of BTP available, clinicians within a single service differed in their interpretation and application of this definition, for example, whether incident pain should be regarded as a subtype of BTP. As a result, clinicians identified a large number and wide variety of terms and phrases that could be used to document BTP in children's narrative notes. Clinicians found it particularly challenging to decide

whether neurological symptoms (e.g. dystonia) and neuropathic pain and its management strategies could be regarded as BTP.

Second, this study illustrates several methodological challenges related to using narrative clinical records for research purposes. While narrative clinical records provide a richer description of pain experiences than pain scores or records of analgesia, the potential to identify a complex pain symptom through the presence of specific words or phrases was challenged by palliative care clinicians who emphasized the subjective and contextual aspects of assessments. In addition, our study also demonstrated that clinical records serve several purposes including an assessment of the child, a report of actual and planned interventions and are often written to illustrate the rationale for clinical decisions. They are not neutral reports but rather a constructed account of complex events.

Implications for practice

Given the challenges encountered, we do not recommend a further retrospective review of BTP using narrative records. While it was not possible to generate reliable information from a retrospective review of narrative clinical records, challenges identified and lessons learned can give direction to future documentation in clinical records and research. Clear and consistent recording of notes is crucial for the clinical care of children with LLCs or LTIs, especially because data from a wide variety of sources and care settings are used in care and treatment. Narrative clinical records could benefit from a more structured format, the inclusion of validated assessment tools and consistent recording of contextual factors. Opportunities to debate and develop a consensus view among clinicians are also critical in improving the consistency of notes and as such should be a feature of practice review. Better consistency of documentation will enable researchers to make use of the information that is already available to obtain a better understanding of complex symptoms and, ultimately, develop effective interventions. Currently available guidance for management of BTP in children is based on clinical experience rather than research evidence (Friedrichsdorf and Postier, 2014). The best way forward would be a prospective study recording episodes of BTP and pain management interventions based on a clear definition of all components and utilizing structured pain assessment tools.

Acknowledgements

The authors would like to thank all clinicians who participated in the consensus study for their invaluable feedback and suggestions. We would like to thank Lorraine Hackett for her assistance with the pilot study of the data extraction instrument. We are grateful to Liz Rasdall for her contribution to identifying terms and phrases for the first round of the consensus study, to Dr Alison Twycross, Dr Emily Harrop and Professor Faith Gibson for reviewing a draft of the study protocol, to Dr Michelle Koh for her contribution to piloting the questionnaire and to Dr Penny Titman who facilitated group discussions in the consensus study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was partly funded by a charitable trust who wishes to

remain anonymous. The researchers were funded by the Louis Dundas Centre for Children's Palliative Care (grant number 2LGB/B (PK); 2LGB/C (JC and LO)) and the True Colours Trust (grant number 2LGA) (MBL). The research was supported by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London. The funders had no role in study design, in the collection, analysis and interpretation of data, in the writing of the manuscript and in the decision to submit the article for publication.

ORCID iD

Linda JM Oostendorp  <https://orcid.org/0000-0001-5544-2672>

Supplemental material

Supplemental material for this article is available online.

References

- Allison JJ, Wall TC, Spettell CM, et al. (2000) The art and science of chart review. *Journal on Quality Improvement* 26: 115–136.
- Chang J, Versloot J, Fashler SR, et al. (2015) Pain assessment in children: validity of facial expression items in observational pain scales. *The Clinical Journal of Pain* 31: 189–197.
- Craig F, Henderson EM and Bluebond-Langner M (2015) Management of respiratory symptoms in paediatric palliative care. *Current Opinion in Supportive and Palliative Care* 9: 217–226.
- Curtis JR, Sathitratanaheewin S, Starks H, et al. (2018) Using electronic health records for quality measurement and accountability in care of the seriously ill: opportunities and challenges. *Journal of Palliative Medicine* 21: S52–S60.
- Davies AN, Dickman A, Reid C, et al. (2009) 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. *European Journal of Pain* 13: 331–338.
- de la Cruz M, Reddy A, Vidal M, et al. (2016) Impact of a palliative care checklist on clinical documentation. *Journal of Oncology Practice* 12: e241–e247.
- Drake R, Frost J and Collins JJ (2003) The symptoms of dying children. *Journal of Pain and Symptom Management* 26: 594–603.
- Eder C, Fullerton J, Benroth R, et al. (2005) Pragmatic strategies that enhance the reliability of data abstracted from medical records. *Applied Nursing Research* 18: 50–54.
- Engel L, Henderson C, Fergenbaum J, et al. (2009) Medical record review conduction model for improving interrater reliability of abstracting medical-related information. *Evaluation & The Health Professions* 32: 281–298.
- Feudtner C, DiGiuseppe DL and Neff JM (2003) Hospital care for children and young adults in the last year of life: a population-based study. *BMC Medicine* 1: 3.
- Friedrichsdorf SJ, Finney D, Bergin M, et al. (2007) Breakthrough pain in children with cancer. *Journal of Pain and Symptom Management* 34: 209–216.
- Friedrichsdorf SJ and Postier A (2014) Management of breakthrough pain in children with cancer. *Journal of Pain Research* 7: 117–123.
- Furuno JP, Noble BN, Bearden DT, et al. (2017) Feasibility of retrospective pharmacovigilance studies in hospice care: a case study of antibiotics for the treatment of urinary tract infections. *Journal of Palliative Medicine* 20: 316–317.

- Gearing RE, Mian IA, Barber J, et al. (2006) A methodology for conducting retrospective chart review research in child and adolescent psychiatry. *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 15: 126–134.
- Goldman A, Hewitt M, Collins GS, et al. (2006) Symptoms in children/young people with progressive malignant disease: United Kingdom Children's Cancer Study Group/Paediatric Oncology Nurses Forum survey. *Pediatrics* 117: e1179–e1186.
- Goulet F, Jacques A, Gagnon R, et al. (2007) Assessment of family physicians' performance using patient charts: interrater reliability and concordance with chart-stimulated recall interview. *Evaluation & The Health Professions* 30: 376–392.
- Gregory KE and Radovinsky L (2012) Research strategies that result in optimal data collection from the patient medical record. *Applied Nursing Research* 25: 108–116.
- Gunhardsson I, Svensson A and Bertero C (2008) Documentation in palliative care: nursing documentation in a palliative care unit—a pilot study. *American Journal of Hospice and Palliative Medicine*® 25: 45–51.
- Hellings P (2004) A rich source of clinical research data: medical records and telephone logs. *Journal of Pediatric Health Care* 18: 154–155.
- Ho C and Straatman L (2013) A review of pediatric palliative care service utilization in children with a progressive neuromuscular disease who died on a palliative care program. *Journal of Child Neurology* 28: 40–44.
- Hunt A (2012) Pain assessment. In: Goldman A, Hain R and Liben S (eds) *Oxford Textbook of Palliative Care for Children*, 2nd ed. Oxford: Oxford University Press, pp. 204–217.
- Jagt-van Kampen CT, van de Wetering MD and Schouten-van Meeteren AYN (2015) The timing, duration, and management of symptoms of children with an incurable brain tumor: a retrospective study of the palliative phase. *Neuro-Oncology Practice* 2: 70–77.
- James D and Warren-Forward H (2015) Research methods for formal consensus development. *Nursing Research* 22: 35–40.
- Jansen AC, van Aalst-Cohen ES, Hutten BA, et al. (2005) Guidelines were developed for data collection from medical records for use in retrospective analyses. *Journal of Clinical Epidemiology* 58: 269–274.
- Jones J and Hunter D (1995) Consensus methods for medical and health services research. *British Medical Journal* 311: 376–380.
- Kelly J, Ritchie J, Donovan L, et al. (2018) A retrospective review of resuscitation planning at a children's hospital. *Children (Basel)* 5: 9.
- Landis JR and Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174.
- Liben S, Papadatou D and Wolfe J (2008) Paediatric palliative care: challenges and emerging ideas. *Lancet* 371: 852–864.
- Lilford R, Edwards A, Girling A, et al. (2007) Inter-rater reliability of case-note audit: a systematic review. *Journal of Health Services Research & Policy* 12: 173–180.
- McEvoy M (2000) Development of a new approach to palliative care documentation. *International Journal of Palliative Nursing* 6: 288–297.
- McMillan SS, King M and Tully MP (2016) How to use the nominal group and Delphi techniques. *International Journal of Clinical Pharmacy* 38: 655–662.
- Meschia JF, Barrett KM, Chukwudelunzu F, et al. (2006) Interobserver agreement in the trial of org 10172 in acute stroke treatment classification of stroke based on retrospective medical record review. *Journal of Stroke and Cerebrovascular Diseases* 15: 266–272.
- Orsey AD, Belasco JB, Ellenberg JH, et al. (2009) Variation in receipt of opioids by pediatric oncology patients who died in children's hospitals. *Pediatric Blood & Cancer* 52: 761–766.

- Rashotte J, Coburn G, Harrison D, et al. (2013) Health care professionals' pain narratives in hospitalized children's medical records. Part 1: pain descriptors. *Pain Research and Management* 18: e75–e83.
- Reeves MJ, Mullard AJ and Wehner S (2008) Inter-rater reliability of data elements from a prototype of the Paul Coverdell National Acute Stroke Registry. *BMC Neurology* 8: 19.
- Samuels JG (2012) Abstracting pain management documentation from the electronic medical record: comparison of three hospitals. *Applied Nursing Research* 25: 89–94.
- Schiessl C, Gravou C, Zernikow B, et al. (2008) Use of patient-controlled analgesia for pain control in dying children. *Supportive Care in Cancer* 16: 531–536.
- Siden H, Steele R, Brant R, et al. (2012) Longitudinal assessment of symptoms in the charting the territory study: when a child has a progressive, metabolic, neurological, or chromosomal condition. *Journal of Palliative Care* 28: 230–230.
- Sim J and Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* 85: 257–268.
- Simons JM and Macdonald LM (2004) Pain assessment tools: children's nurses' views. *Journal of Child Health Care* 8: 264–278.
- Stewart K, Doody O, Bailey M, et al. (2017) Improving the quality of nursing documentation in a palliative care setting: a quality improvement initiative. *International Journal of Palliative Nursing* 23: 577–585.
- Thomas EJ, Lipsitz SR, Studdert DM, et al. (2002) The reliability of medical record review for estimating adverse event rates. *Annals of Internal Medicine* 136: 812–816.
- Thrane SE, Maurer SH, Cohen SM, et al. (2017) Pediatric palliative care: a five-year retrospective chart review study. *Journal of Palliative Medicine* 20: 1104–1111.
- Tomlinson D, Hinds PS, Bartels U, et al. (2011) Parent reports of quality of life for pediatric patients with cancer with no realistic chance of cure. *Journal of Clinical Oncology* 29: 639–645.
- Twycross A, Voepel-Lewis T, Vincent C, et al. (2015) A debate on the proposition that self-report is the gold standard in assessment of pediatric pain intensity. *The Clinical Journal of Pain* 31: 707–712.
- Vakil N (2011) Editorial: consensus guidelines: method or madness? *The American Journal of Gastroenterology* 106: 225–227; quiz 228.
- World Health Organization (WHO; 2012) *WHO Guidelines on the Pharmacological Treatment of Persisting Pain in Children with Medical Illnesses*. Geneva: World Health Organization.
- Wolfe J, Grier HE, Klar N, et al. (2000) Symptoms and suffering at the end of life in children with cancer. *The New England Journal of Medicine* 342: 326–333.
- Worster A and Haines T (2004) Advanced statistics: understanding medical record review (MRR) studies. *Academic Emergency Medicine* 11: 187–192.
- Yawn BP and Wollan P (2005) Interrater reliability: completing the methods description in medical records review studies. *American Journal of Epidemiology* 161: 974–977.