UNIVERSITY *of* York

This is a repository copy of *Documentation of breakthrough pain in narrative clinical records of children with life-limiting conditions : feasibility of a retrospective review*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/135575/

Version: Accepted Version

Article:

1 **Documentation of breakthrough pain in narrative clinical records of children with life-limiting**

2 **conditions: feasibility of a retrospective review**

3

4 Linda JM Oostendorp,[1] Dilini Rajapakse,[2] Paula Kelly,[2] Joanna Crocker,[1,3] Andrew Dinsdale,[2] Lorna

5 Fraser,[5] Myra Bluebond-Langner[1]

6

7 *[1] Louis Dundas Centre for Children's Palliative Care, Institute of Child Health, University College London,*

8 *UK*

9 *[2] Louis Dundas Centre for Oncology Outreach and Children's Palliative Care, Great Ormond Street Hospital*

10 *for Children NHS Foundation Trust, UK*

11 *[3]Health Experiences Institute, University of Oxford, UK*

12 *[4] Department of Health Sciences, University of York, UK*

13

14 **Correspondence to:**

15 Prof Myra Bluebond-Langner, Louis Dundas Centre for Children's Palliative Care, Great Ormond Street

16 Hospital, UCL Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK.

17 Phone: +44 207 905 2781; Email: Bluebond@ucl.ac.uk

18

19 **Running head (50 characters):** Breakthrough pain: narrative clinical records

20

21 **Word count:** 4,292

22

22

23

24

1   **ABSTRACT**

2   This study explored the feasibility of generating reliable information on the frequency, nature and

3   management of breakthrough pain (BTP) in children with life-limiting conditions (LLCs) and life-

4   threatening illnesses (LTIs) from narrative clinical records. In the absence of standardised ways for

5   documenting BTP, we conducted a consensus exercise to develop a glossary of terms that could denote

6   BTP in the records. Thirteen clinicians who contributed to the records reached consensus on 45 terms

7   which could denote BTP, whilst emphasising the importance of contextual information. The results of this

8   approach together with guidance for improving the reliability of retrospective reviews informed a data

9   extraction instrument. A pilot test of this instrument showed poor agreement between raters. Given the

10 challenges encountered, we do not recommend a retrospective review of BTP using narrative records.

11 This study highlighted challenges of data extraction for complex symptoms such as BTP from narrative

12 clinical records. For both clinical and research purposes, the recording of complex symptoms such as BTP

13 would benefit from clear criteria for applying definitions, a more structured format and the inclusion of

14 validated assessment tools. This study also showed the value of consensus exercises in improving

15 understanding and interpretation of clinical notes within a service.

16

17 **KEYWORDS:** Retrospective review, Data Collection/instrumentation, Breakthrough Pain, Child, Life-

18 limiting conditions

19

20

21

1   **INTRODUCTION**

2

3   Retrospective reviews of clinical records are a relatively inexpensive way to generate greater insight in

4   unexplored areas of medical research and inform subsequent prospective studies (Hellings, 2004; Gearing

5   et al., 2006). Extracting useful information from clinical records, however, presents various challenges

6   including individual variation in documentation and inconsistent use of terminology (Worster and Haines,

7   2004; Samuels, 2012). In emerging areas of practice, such as paediatric palliative care, shared descriptors

8   of signs and symptoms may not even exist (Liben et al., 2008; Siden, 2012; Craig et al., 2015).

9   Reports from professionals and parents indicate that pain is frequently unsatisfactorily treated in

10   children with life-limiting conditions (LLCs) or life-threatening illnesses (LTIs), particularly towards the end

11   of life (Wolfe et al., 2000; Drake et al., 2003; Goldman et al., 2006; Tomlinson et al., 2011). One of the

12   most difficult types of pain to manage is breakthrough pain (BTP). Even when an active pain management

13   strategy is in place to alleviate background pain children can experience BTP. The WHO definition of BTP

14   is provided in Figure 1. In brief, BTP is characterised as "a temporary increase in the severity of pain over

15   and above the pre-existing baseline pain level" (WHO, 2012). Importantly, most evidence on pain

16   assessment and management in children relates to post-operative pain and caution is required in

17   extrapolating this evidence to children with LLCs where clinicians anticipate pain escalation as the disease

18   progresses. The single study investigating BTP in a population of children with LLCs (Friedrichsdorf et al.,

19   2007) showed that 57% out of 27 hospitalised children with cancer who had controlled background pain

20   reported to have experienced BTP in the previous 24 hours. While this study is a useful starting point,

21   further larger-scale prospective studies on the frequency, nature and management of BTP in the entire

22   population of children with LLCs or LTIs are needed to develop effective interventions.

23   Clinical records detailing childrens' pain symptoms and the strategies used to alleviate these are

24   a potential source of data for baseline information. As a first step in expanding the evidence base on the

1   frequency, nature and management of BTP in children with LLCs or LTIs we undertook a retrospective

2   review of the clinical records of children cared for by the largest specialist paediatric palliative care service

3   in the UK, giving access to a large and diverse cohort of children. While clinical records are recorded

4   electronically within a bespoke Access database and comprise structured elements, the service has not

5   implemented structured electronic health records yet. Documentation of clinical assessments carried out

6   by team members (Clinical Nurse Specialists, Medical Registrars, and Consultants) in the home, hospital,

7   or hospice, or via telephone consultations are reported in a traditional narrative format. In the absence

8   of standardised measures and shared descriptors for documenting BTP we decided to conduct a

9   consensus exercise, using features from both the Delphi consensus technique and the nominal group

10   technique, to develop a glossary of terms that could denote BTP in the records of this service. In this article

11   we report on the results of the consensus exercise as it relates to the development of a data extraction

12   instrument for a retrospective review of clinical notes. We also report on the feasibility of generating

13   reliable information about BTP when applying this instrument to clinical records. Finally, we report on

14   how such consensus exercises can inform both current practice within a clinical team and design of

15   prospective studies on a particular symptom or set of symptoms.

16

17

18

19

20

1 **METHODS**

2

3 **Design**

4 The aim of the study was to explore the feasibility of generating reliable information on the frequency,

5 nature and management of BTP in children with LLCs or LTIs from narrative clinical records to inform

6 current practice and the design of a prospective study. The study consisted of three parts, and the first

7 part was a consensus exercise among clinicians to create a glossary of terms that could denote BTP in the

8 narrative clinical records of children with LLCs or LTIs. Importantly, we did not aim to develop a definition

9 of BTP itself (we used the WHO definition as described above) rather we sought consensus on how

10 clinicians interpret and apply this definition and how they may document BTP in narrative clinical records.

11 In the second part of the study, we developed a data extraction instrument informed by the results of the

12 consensus exercise and published guidance on data extraction instruments (Eder et al., 2005; Jansen et

13 al., 2005; Engel et al., 2009; Gregory and Radovinsky, 2012). In part three of the study a pilot test of the

14 data extraction instrument was performed to assess interrater reliability (Yawn and Wollan, 2005; Lilford

15 et al., 2007).

16

17 **Participants**

18

19 *Consensus exercise*

20 Since we aimed to conduct a retrospective review of the records kept within a particular service, the most

21 appropriate participants for the consensus exercise were deemed to be the clinicians contributing to these

22 records. All six nurses and seven out of eight doctors from the service were invited to participate in the

23 consensus exercise (one doctor was a member of the study team).

24

1   *Pilot test of the data extraction instrument*

2   Two raters independently used the data extraction instrument to identify episodes of breakthrough pain

3   in the records of children with LLCs or LTIs. This included a Research Nurse and a Palliative Care Consultant

4   (DR), both experienced in caring for children with LLCs or LTIs. Data were extracted from the records of

5   52 children with a LLC or LTI, selected from the service's caseload between 1st July 2012 and 30th June

6   2014.

7

8   **Questionnaire development**

9   An initial list of terms and phrases that could potentially denote BTP in the clinical records was prepared

10 for the clinicians to give feedback on. Two clinicians working within the service (DR and Liz Rasdall) and a

11 member of the study team with both clinical and research experience (PK) independently reviewed the

12 records of 20 children to extract any terms and phrases potentially used to denote BTP. These 20 children

13 were a sample of the children who had died within the last three years, from malignant and non-malignant

14 conditions. A total of 63 terms and phrases were identified and divided into five categories: (i) Symptoms

15 reported by patients, parents or healthcare professionals (comprising a list of pain descriptors) (15 items);

16 (ii) Behaviours noted in the child (18 items); (iii) Neurological symptoms (3 items); (iv) Medications (24

17 items); and (v) Other treatments (3 items) (see Supplementary File 1).

18   In the online questionnaire (1A/B) clinicians were asked to rate whether they felt the presence of

19   each of these items would indicate BTP ('yes', 'no', 'possibly') (Supplementary File 2). Clinicians could

20   also suggest additional terms/phrases. Participants were provided with a definition of BTP encompassing

21   the WHO's definition of both BTP and incident pain in children with medical illness suffering from

22   persistent pain (WHO, 2012). In including incident pain in our definition, we considered that definitions

23   vary as to whether incident pain is included (Davies et al., 2009; Friedrichsdorf and Postier, 2014) or not

24   (WHO, 2012) and decided to perform an inclusive review of BTP in children with life-limiting conditions,

1   where identifying the underlying cause of the pain can inform selection of a management strategy (see

2   Figure 1 for the full definitions).

3   In the second questionnaire (2A/B), we included contextual information alongside the items

4   selected from the first questionnaire (see Figure 2). For example, whereas questionnaires 1A/1B showed

5   terms such as 'achiness', contextual factors were added to the terms and phrases in questionnaires 2A/2B

6   to facilitate discrimination of BTP (e.g. 'sudden occurrence of achiness') (see Supplementary File 1 for an

7   overview of the contextual factors added). Participants were asked to record for each of the

8   terms/phrases whether they would indicate BTP ('yes' or 'no') in the context provided (Supplementary

9   File 3).

10

11   **Development of data extraction instrument**

12   An electronic data extracting instrument (see Supplementary File 4) and accompanying manual (see

13   Supplementary File 5) were developed. The data extraction instrument was designed to identify

14   occurrences of BTP in three steps (see Figure 3): (i) Identification of pain episodes using a glossary of 28

15   pain descriptors developed through the consensus exercise; (ii) Selection of pain episodes with evidence

16   of controlled background pain; and (iii) Confirmation of the presence of BTP through identification of

17   contextual information (derived from the consensus exercise results). The electronic instrument

18   contained prompts to aid data extraction and provide explanations, and fields were made mandatory

19   where applicable. Wherever possible, dropdown lists were generated to reduce typing and errors.

20

21   **Data collection**

22

23   *Consensus exercise*

1   The consensus process is depicted in Figure 2. Participants in the consensus exercise both completed

2   questionnaires (as also used in the Delphi consensus technique) and participated in face-to-face meetings

3   (which had features of the nominal group technique) (McMillan et al., 2016) (see Figure 2). Clinicians

4   completed four anonymous online questionnaires at their convenience (rounds 1A/1B/2A/2B). Two face-

5   to-face meetings were held to enable clinicians to discuss areas of disagreement and help the study team

6   to better understand their views (Jones and Hunter, 1995; Vakil, 2011; James and Warren-Forward, 2015).

7   These meetings were facilitated by an external clinical psychologist and highly structured to minimise the

8   risk of one or more participants dominating discussions, and to make sure all relevant issues were

9   presented and discussed. To optimise attendance, these meetings were part of the clinicians' weekly

10   handover meeting. After these meetings, participants were invited to re-complete the questionnaire

11   (rounds 1B and 2B). Participants were provided with a copy of the results emphasising that they need not

12   conform to the prevailing view. An observer (JC/PK) took field notes to capture the content and tenor of

13   discussions.

14

15   *Pilot test of the data extraction instrument*

16   In the pilot test, two raters independently extracted data using the data extraction instrument to assess

17   interrater reliability and identify any issues that could impede a full retrospective review. In part 1, raters

18   extracted data from 21 clinical records that were selected randomly, stratified by year to factor in changes

19   to documentation styles, and were checked to ensure both malignant and non-malignant conditions were

20   included. The study team, including the two raters, discussed disagreements and agreed refinements of

21   the data extraction instrument to reflect the insights gained. The same two clinicians extracted data from

22   another 31 randomly selected clinical records to conduct a further test of interrater reliability. A sample

23   size calculation showed that, assuming the null hypothesis value of kappa to be 0.0, 22-30 cases would be

1   required to detect a statistically significant kappa of 0.6 or 0.7 with a power of 90% (Sim and Wright,

2   2005).

3

4   **Analysis**

5

6   *Consensus exercise*

7   After each round of the consensus process, a decision was made whether to include items in the data

8   extraction instrument (see Figure 2). Items were included in the instrument if at least 51% of participants

9   selected 'yes' in round 1B or 2B. A low threshold of 51% was considered appropriate, as we aimed to

10   develop an inclusive instrument with a high sensitivity to detect potential occurrences of BTP.

11

12   *Pilot test of the data extraction instrument*

13   In the pilot test, agreement between raters was assessed using Cohen's kappa. The analysis included

14   variables related to patient characteristics and the three steps in identifying BTP (see Figure 3). While the

15   minimum required value of kappa is arbitrary, values of 0.6 and 0.7 (moderate to substantial agreement)

16   are often reported (Landis and Koch, 1977). We used a value of 0.7 for episodes of pain identified, the

17   presence of evidence of controlled background pain, and the presence of contextual information to

18   confirm that the pain was actually BTP. A more lenient value of 0.6 was used for agreement on more

19   detailed information, such as specific pain descriptors.

20

21   **Ethical considerations**

22   The study was approved by the Hospital and University Joint Research and Development Office (13LC02;

23   12-12-2013). Written informed consent was sought from clinicians participating in the consensus exercise.

24   Data were extracted from children's records by members of the palliative care team and were

1   pseudonymised before being transferred to researchers outside of the care team for analysis. Seeking

2   parental consent was not deemed necessary for this retrospective review.

3

4

5

6

7

8

9

1   **RESULTS**

2

3   **Consensus exercise**

4

5   *Participants*

6   Out of 13 clinicians invited, 7-11 clinicians (54-85%) completed questionnaires and attended facilitated

7   discussions. These included three consultants, three specialist registrars, one locum specialist, five clinical

8   specialist nurses and one nurse consultant.

9

10   *Round 1*

11   In round 1A, participants reached consensus for 49 out of 63 items (78%) (see Supplementary File 1). Two

12   participants provided comments, including 'sudden onset of pain' and 'all words are dependent on a

13   thorough history and on who is reporting and who is taking history therefore entirely subjective'.

14   In the facilitated discussion participants commented that they needed more contextual

15   information to determine whether a term or phrase would indicate BTP. This contextual information

16   included temporal terms, underlying cause of pain, age, baseline behaviour, symptom management, and

17   presence of (other) pain descriptors.

18   In round 1B, participants reached consensus on all items; 6 items (10%) would denote BTP, 8 items

19   (13%) would not denote BTP, and 49 items (78%) would possibly denote BTP, depending on context. Two

20   comments were provided, including 'Consider temporal terms in relation to pain' and 'Using a temporal

21   relationship to pain descriptors may help to define terminology'.

22

23

24

1   *Round 2*

2   Contextual information was added to the remaining 49 items (see Supplementary File 1). In round 2A,

3   clinicians reached consensus that all items in one or more given contexts would denote BTP. Five

4   comments were provided, all stating that while a term could possibly denote BTP, it was difficult to say

5   this with certainty and more information about the context was required. Two contradictory comments

6   were made about the definition of BTP as used in this study, one recommending to include incident pain

7   in the definition of BTP, and the other recommending against it.

8   The facilitated discussion provided further insight into the challenges still present even with

9   additional contextual information. Clinicians indicated deciding between yes and no was challenging, with

10 some stating they had responded to uncertainty with 'yes'. In addition some distinction was made

11 between phenomena that could evoke pain (e.g. spasms) and those where pain was accepted as the

12 underlying causal mechanism for the behaviour demonstrated. Some respondents advocated more

13 contextual information (e.g. further information on pain history). As with the first facilitated discussion,

14 there was considerable discussion on symptoms associated with neuropathic pain.

15   In round 2B, consensus was achieved such that 39 items could denote BTP in at least one given

16 context (80%). Ten items (20%) did not reach consensus on 'yes' in any given context and were excluded.

17

18 **Pilot test of data extraction instrument**

19 In the first part of the pilot test, data were extracted from the clinical records of 21 children with a

20 malignant (24%) or non-malignant (76%) condition. The two extractors identified 82 episodes of pain; only

21 28 were identified by both (34% agreement; κ= -0.475; p<0.001). Following discussions, the data

22 extraction instrument was refined by explicating the first pain descriptor 'pain of any description',

23 amalgamating similar pain descriptors, and adding a list of relevant medications and interventions to the

24 database (see Supplementary File 4).

1   In the second part of the pilot test, data were extracted from the clinical records of 31 children

2   with a malignant (23%) or non-malignant (77%) condition. The two extractors identified 121 episodes of

3   pain, of which 58 were identified by both (48% agreement; κ= -0.346; p=0.001).

4

5

6

7

8

9

10

11

12

13

14

15

16

17

1 **DISCUSSION**

2

3 **Main findings**

4 In this study, we investigated the feasibility of a retrospective review of the rich information in the records

5 of children with LLCs or LTIs to obtain more insight into BTP. Using a structured consensus approach, we

6 found that clinicians recognised a wide range of terms and phrases that could be used to document BTP

7 in narrative records, although single word descriptors were usually considered insufficient without further

8 contextual elements. A pilot test of a structured data extraction instrument based on the terms and

9 phrases identified showed that even with detailed guidance in place, interrater reliability in identifying

10 episodes of BTP from the narrative clinical records was unsatisfactory. Data extracted under these

11 conditions are unlikely to provide reliable information about the frequency, nature and management of

12 BTP (Landis and Koch, 1977), therefore we concluded that a retrospective review of narrative clinical

13 records was not feasible..

14

15 **Findings in context of literature**

16 Assessing interrater reliability should be a routine component of any data extraction study but is often

17 overlooked (Allison et al., 2000). Recent retrospective reviews in the population of children with LLCs and

18 LTIs include reviews of symptoms (Jagt-van Kampen C., 2015), pain medication (Orsey et al., 2009; Schiessl

19 et al., 2008), and care provided (Feudtner et al., 2003; Ho and Straatman, 2013). While some of these

20 studies described measures taken to improve data quality, none described a reliability assessment.

21 Several retrospective reviews in adult patient populations that did include an assessment of interrater

22 reliability have demonstrated problems (Thomas et al., 2002; Meschia et al., 2006; Goulet et al., 2007;

23 Reeves et al., 2008). 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; Jansen et al., 2005; Engel et al., 2009; Gregory and Radovinsky, 2012). Terms and phrases to be used in round 1 were derived from a review of children's clinical records. A Canadian study among hospitalised children provided a much more extensive list of terms and phrases to describe any pain (e.g. including physiological indicators) (Rashotte et al., 2013), but there was considerable overlap with the terms and phrases identified in our study. 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 C., 2015), or no information on who extracted data was provided (Feudtner et al., 2003; Schiessl et al., 2008; Orsey et al., 2009; Chang et al., 2015). 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 which cares for children with LLCs or LTIs both in hospital and in the community. Clinical notes may be

1   completed when a team member has assessed a child in hospital, at home or in a hospice. They may also

2   be completed after a telephone consultation with a parent or healthcare professional. These proxy

3   assessments result in children's pain experiences being further filtered through other reporters (Twycross

4   et al., 2015), prior to documentation in the clinical records. Similar issues have been noted in other

5   retrospective reviews in this population, including discontinuity in terms of location of care, and lack of

6   documentation who assessed the child using what methods, and whether consultations were performed

7   by phone or face to face (Jagt-van Kampen C., 2015; Kelly et al., 2018; Thrane et al., 2017). Suboptimal

8   documentation was also discussed in retrospective reviews in adult populations, and it was suggested that

9   this could be improved by using additional data sources (e.g. meeting the clinician who recorded the

10   notes) (Goulet et al., 2007), but this may not be feasible.

11   An additional challenge is the complexity of BTP. It is by definition intermittent and often

12   unpredictable such that reporting and recording will vary. As highlighted in the consensus exercise,

13   documentation of contextual factors was regarded as essential in determining the presence of BTP.

14   Another factor possibly contributing to poor interrater reliability are the inconsistencies in the narrative

15   nature of the notes and the lack of a routinely used structured pain assessment rating scale (Hunt, 2012;

16   Chang et al., 2015). Correspondingly, a lack of structure has been highlighted in previous studies on

17   documentation in palliative care (Gunhardsson et al., 2008; McEvoy, 2000; Stewart et al., 2017; Curtis et

18   al., 2018; Furuno et al., 2017; Yawn and Wollan, 2005). In one study, pain was commonly noted but

19   different terminology was used, with pain characteristics often lacking (Gunhardsson et al., 2008). The

20   authors emphasised the impact of accurate documentation on quality of care, symptom control and

21   effective hand-over (Gunhardsson et al., 2008; McEvoy, 2000). The use of guidelines, checklists and more

22   structured forms have been suggested to improve documentation (Gunhardsson et al., 2008; McEvoy,

23   2000; Stewart et al., 2017; de la Cruz et al., 2016). A study among children's nurses showed that the

1 majority of these nurses agreed that the introduction of pain assessment tools across the hospital would

2 improve documentation (Simons and Macdonald, 2004).

3

4 **Strengths and limitations**

5 While recognising the potential value of the information contained in clinical records for research

6 purposes, we were aware of the challenges in extracting reliable data and decided to use a structured

7 process in developing a data extraction instrument, using clinician feedback and published guidance, and

8 assess the reliability of data extracted. While extracting reliable data on BTP from childrens' narrative

9 clinical records remains elusive, this study is one of the first to investigate the description of pain in

10 childrens' clinical records. A Canadian study has investigated the description of pain in childrens' clinical

11 records but this included any pain, not specifically BTP, and the study was limited to a 24-hour period for

12 inpatients in Canadian paediatric hospitals (Rashotte et al., 2013). Our study specifically focused on BTP

13 in children with malignant and non-malignant life-limiting illnesses, cared for by a paediatric palliative

14 care team, adding the complexity of multiple settings. Admittedly, this study was confined to a single

15 tertiary palliative care service and therefore terms and phrases identified in this service may not be

16 recognised or utilised in narrative records by other healthcare professionals in this field. Moreover, other

17 services may have adopted structured electronic health records, which would have implications for the

18 structure of notes and consequently the design of retrospective reviews. However, there are indications

19 that even when structured electronic records are used, clinicians may still resort to narratives to

20 document pain episodes in detail (Rashotte et al., 2013). Finally, one could argue that to develop a truly

21 inclusive data extraction instrument, a threshold higher than 51% should have been used for excluding

22 terms/phrases. However, for most of the 18 items that were excluded, at least 70% of participants did not

23 think that this term/phrase in the given context would indicate BTP.

24

1 **What this study adds**

2 Several lessons can be learned from this feasibility study. First, the consensus exercise yielded important

3 insights into BTP that can benefit future studies. Even with a WHO definition of BTP available, clinicians

4 within a single service differed in their interpretation and application of this definition, for example in

5 whether incident pain should be regarded as a subtype of BTP. As a result, clinicians identified a large

6 number and wide variety of terms and phrases that could be used to document BTP in childrens' narrative

7 notes. Clinicians found it particularly challenging to decide whether neurological symptoms (e.g. dystonia)

8 and neuropathic pain and its management strategies could be regarded as BTP.

9   Second, this study illustrates several methodological challenges related to using narrative clinical

10 records for research purposes. While narrative clinical records provide a richer description of pain

11 experiences than pain scores or records of analgesia, the potential to identify a complex pain symptom

12 through the presence of specific words or phrases was challenged by palliative care clinicians who

13 emphasised the subjective and contextual aspects of assessments. In addition our study also

14 demonstrated that clinical records serve several purposes including an assessment of the child, a report

15 of actual and planned interventions and are often written to illustrate the rationale for clinical decisions.

16 They are not neutral reports, but rather a constructed account of complex events.

17

18 **Implications for practice**

19 Given the challenges encountered, we do not recommend a further retrospective review of BTP using

20 narrative records. While it was not possible to generate reliable information from a retrospective review

21 of narrative clinical records, challenges identified and lessons learned can give direction to future

22 documentation in clinical records and research. Clear and consistent recording of notes is crucial for the

23 clinical care of children with LLCs or LTIs, especially because data from a wide variety of sources and care

24 settings are used in care and treatment. Narrative clinical records could benefit from a more structured

1   format, the inclusion of validated assessment tools and consistent recording of contextual factors.

2   Opportunities to debate and develop a consensus view among clinicians are also critical in improving the

3   consistency of notes and as such should be a feature of practice review. Better consistency of

4   documentation will enable researchers to make use of the information that is already available to obtain

5   a better understanding of complex symptoms and, ultimately, develop effective interventions. Currently

6   available guidance for management of BTP in children is based on clinical experience rather than research

7   evidence (Friedrichsdorf, 2014). The best way forward would be a prospective study recording episodes

8   of BTP and pain management interventions, based on a clear definition of all components and utilising

9   structured pain assessment tools.

REFERENCES

Allison JJ, Wall TC, Spettell CM, et al. (2000) The art and science of chart review. *Jt Comm J Qual Improv* 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. *Clin J Pain* 31: 189-197.

Craig F, Henderson EM and Bluebond-Langner M. (2015) Management of respiratory symptoms in paediatric palliative care. *Curr Opin Support Palliat Care* 9: 217-226.

Curtis JR, Sathitratanacheewin S, Starks H, et al. (2018) Using Electronic Health Records for Quality Measurement and Accountability in Care of the Seriously Ill: Opportunities and Challenges. *J Palliat Med* 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. *Eur J Pain* 13: 331-338.

de la Cruz M, Reddy A, Vidal M, et al. (2016) Impact of a Palliative Care Checklist on Clinical Documentation. *J Oncol Pract* 12: e241-247.

Drake R, Frost J and Collins JJ. (2003) The symptoms of dying children. *J Pain Symptom Manage* 26: 594-603.

Eder C, Fullerton J, Benroth R, et al. (2005) Pragmatic strategies that enhance the reliability of data abstracted from medical records. *Appl Nurs Res* 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. *Eval Health Prof* 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 Med* 1: 3.

Friedrichsdorf SJ, Finney D, Bergin M, et al. (2007) Breakthrough pain in children with cancer. *J Pain Symptom Manage* 34: 209-216.

Friedrichsdorf SJ and Postier A. (2014) Management of breakthrough pain in children with cancer. *J Pain Res* 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. *J Palliat Med* 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. *J Can Acad Child Adolesc 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: e1179-1186.

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. *Eval Health Prof* 30: 376-392.

Gregory KE and Radovinsky L. (2012) Research strategies that result in optimal data collection from the patient medical record. *Appl Nurs Res* 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. *Am J Hosp Palliat Care* 25: 45-51.

Hellings P. (2004) A rich source of clinical research data: medical records and telephone logs. *J Pediatr 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. *J Child Neurol* 28: 40-44.

Hunt A. (2012) Pain Assessment. In: Goldman A. HR, Liben S. (ed) *Oxford Textbook of Palliative Care for Children.* 2nd edition ed. Oxford: Oxford University Press, 204-217.

Jagt-van Kampen C. WvdMD, Schouten-van Meeteren A.Y.N. (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.

James D and Warren-Forward H. (2015) Research methods for formal consensus development. *Nurse Res* 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. *J Clin Epidemiol* 58: 269-274.

Jones J and Hunter D. (1995) Consensus methods for medical and health services research. *BMJ* 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.

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. *J Health Serv Res Policy* 12: 173-180.

McEvoy M. (2000) Development of a new approach to palliative care documentation. *Int J Palliat Nurs* 6: 288-297.

McMillan SS, King M and Tully MP. (2016) How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 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. *J Stroke Cerebrovasc Dis* 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. *Pediatr 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 Res Manag* 18: e75-83.

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 Neurol* 8: 19.

Samuels JG. (2012) Abstracting pain management documentation from the electronic medical record: comparison of three hospitals. *Appl Nurs Res* 25: 89-94.

Schiessl C, Gravou C, Zernikow B, et al. (2008) Use of patient-controlled analgesia for pain control in dying children. *Support Care Cancer* 16: 531-536.

Siden H, Steele, R., Brant, R., Cadell, S., Davies, B., Straatman, L., Andrews, G. (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. *Phys Ther* 85: 257-268.

Simons JM and Macdonald LM. (2004) Pain assessment tools: children's nurses' views. *J 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. *Int J Palliat Nurs* 23: 577-585.

Thomas EJ, Lipsitz SR, Studdert DM, et al. (2002) The reliability of medical record review for estimating adverse event rates. *Ann Intern Med* 136: 812-816.

Thrane SE, Maurer SH, Cohen SM, et al. (2017) Pediatric Palliative Care: A Five-Year Retrospective Chart Review Study. *J Palliat Med* 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. *J Clin Oncol* 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. *Clin J Pain* 31: 707-712.

Vakil N. (2011) Editorial: consensus guidelines: method or madness? *Am J Gastroenterol* 106: 225-227; quiz 228.

WHO. (2012) WHO guidelines on the pharmacological treatment of persisting pain in children with medical illnesses.

Wolfe J, Grier HE, Klar N, et al. (2000) Symptoms and suffering at the end of life in children with cancer. *N Engl J Med* 342: 326-333.

Worster A and Haines T. (2004) Advanced statistics: understanding medical record review (MRR) studies. *Acad Emerg Med* 11: 187-192.

Yawn BP and Wollan P. (2005) Interrater reliability: completing the methods description in medical records review studies. *Am J Epidemiol* 161: 974-977.

Supplementary File 1: Participants' rating of whether the proposed terms and phrases denote breakthrough pain in a child's clinical records

Supplementary File 2: Questionnaire 1A

Supplementary File 3: Questionnaire 2A

Supplementary File 4: Screenshots of data extraction instrument

Supplementary File 5: Contents of manual accompanying the electronic data extraction instrument