



# The Italian Questionnaire for Cancer Breakthrough Pain Diagnosis, a Multicenter Validation Study

Boaz Gedaliahu Samolsky Dekel · Alberto Gori ·  
Marco Gunnellini · Angela Gioia · Mariacristina Di Marco ·  
Giuseppe Casale · Marzio Bevilacqua · Pietro Bersani ·  
Rita Maria Melotti

Received: April 10, 2021 / Accepted: May 17, 2021 / Published online: June 6, 2021  
© The Author(s) 2021

## ABSTRACT

**Introduction:** The literature lacks formally validated and reliable tools for the diagnosis of breakthrough cancer pain (BTcP). The Italian Questionnaire for BTcP diagnosis (IQ-BTP) is an 11-item questionnaire aimed at detecting potential-BTP and classifying it into three likelihood classes: high, intermediate, and low.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40122-021-00274-9>.

B. G. Samolsky Dekel (✉) · R. M. Melotti  
Department of Medicine and Surgery Sciences,  
University of Bologna, via Massarenti n. 9, 40138  
Bologna, Italy  
e-mail: boaz.samolskydekel@unibo.it

B. G. Samolsky Dekel · R. M. Melotti  
Anesthesia and Pain Therapy Unit, IRCCS AOSP di  
Bologna, Policlinico S. Orsola-Malpighi, Via  
Massarenti n. 9, 40138 Bologna, Italy

B. G. Samolsky Dekel · A. Gori · R. M. Melotti  
Post Graduate School of Anaesthesia, Intensive Care  
and Pain Medicine, University of Bologna, Via  
Massarenti n. 9, 40138 Bologna, Italy

M. Gunnellini  
Medical Oncology and Palliative Care Unit, USL  
Umbria 1, Presidio Ospedaliero di Gubbio, Gualdo  
Tadino, Largo Unità D'Italia, 1, 06024 Gubbio, Italy

M. Gunnellini  
Ospedale Media Valle del Tevere, Vocabolo Buda,  
Todi, 06059, Perugia, Italy

**Methods:** A multicenter, prospective, and observational study was designed to validate the IQ-BTP. In three consecutive visits with each cancer patient, the demographic and clinical details of the patient, the Brief Pain Inventory (BPI) scores, IQ-BTP outcomes, and clinicians' autonomous BTcP diagnosis (gold standard) and the agreement of this diagnosis with IQ-BTP outcomes were recorded. The assessed domains for IQ-BTP validation were: Validity, including content and face validity, construct validity (hypothesis testing, and cross-cultural

A. Gioia  
Hospice, UF Cure Palliative, Azienda USL Toscana  
NordOvest, via Cocchi, 7-9, Pisa, Italy

M. Di Marco  
Medical Oncology, IRCCS Azienda Ospedaliero-  
Universitaria di Bologna, Department of  
Experimental, Diagnostic and Specialty Medicine, S.  
Orsola-Malpighi University Hospital, Via Albertoni,  
15, Bologna, Italy

G. Casale  
Antea Foundation Research Group, piazza Santa  
Maria della pietà 5, 00135 Rome, Italy

M. Bevilacqua  
Dipartimento di Anestesia, Rianimazione e Terapia  
Intensiva U.O.C. di Medicina e Terapia del Dolore,  
ULSS2, Marca Trevigiana, Treviso, Italy

P. Bersani  
Gigi Ghirotti Association, Home Care, Corso  
Europa, 50/9, 16132 Genoa, Italy

validity\measurement invariance), and criterion validity; Reliability (internal consistency, reliability, and measurement error); Interpretability, and Responsiveness.

**Results:** Seven palliative and pain management facilities in Italy recruited 280 patients, yielding 753 evaluations. Using the IQ-BTP, the rate of potential-BTcP was 27.2%, of which its likelihood was high in 52.7% of patients, intermediate in 38.5, and low in 8.8%. The BPI item scores differed significantly between the two IQ-BTP classes (no-BTcP and potential-BTcP classes). The correlation of the latter class with BPI items was significant but low. The IQ-BTcP showed two principal components, accounting for 66.6% of the variance. Cronbach's  $\alpha$  was 0.71. The agreement rate between the gold standard and IQ-BTP outcomes was 82%. Cohen's  $\kappa$  was 0.535. The IQ-BTP showed sensitivity and specificity of 69 and 86%, respectively.

**Conclusions:** The IQ-BTP extensive formal validation showed satisfactory psychometric and validity properties. Its content, face, construct, and criterion validities and its reliability, interpretability, and responsiveness were shown. Its use enabled potential-BTcP to be identified and differentiated into three likelihood classes with direct therapeutic and epidemiological implications. The latter may be confirmed in future studies.

**Keywords:** BTcP; BTP; Construct validity; Criterion validity; Interpretability; IQ-BTP; Reliability; Responsiveness; Validation

## Key Summary Points

### Why carry out this study?

The literature lacks formally validated and reliable tools for the diagnosis of breakthrough cancer pain (BTcP).

The Italian Questionnaire for BTcP diagnosis (IQ-BTP) is an 11-item questionnaire aimed at detecting the presence of potential-BTcP and classifying it into three likelihood classes: high, intermediate, and low.

A multicenter, prospective, and observational study was designed to formally validate the IQ-BTP and to highlight its clinical usefulness.

### What was learned from the study?

The IQ-BTP showed satisfactory psychometric and validity properties.

The IQ-BTP enabled potential-BTcP to be identified and differentiated into three likelihood classes with direct congruent therapeutic and epidemiological implications.

## DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to <https://doi.org/10.6084/m9.figshare.14589066>.

## INTRODUCTION

Breakthrough Pain (BTP) in cancer (BTcP) and non-cancer patients is a challenging clinical issue at all healthcare levels. It is considered to be a predictor for poor pain outcomes in patients with chronic pain (CP) and associated with a physical, psychological, and economic

burden for both patients and caregivers [1–5]. Across studies, BTP prevalence varies mainly due to the lack of consensus on its definition, absence of formally validated and reliable tools for its diagnosis, and heterogeneity in the design and parameters of studies [6, 7].

BTcP is not mentioned in the 11th edition of the International Classification of Diseases (ICD-11); rather, with the support of the International Association for the Study of Pain (IASP), cancer pain is divided into continuous (background pain [BP]) or intermittent (episodic: related to movement or activity [incident] or unrelated [spontaneous]) pain [8]. Earlier definitions describe BTcP as “a transitory exacerbation of pain experienced by patients undergoing long-term opioid treatment for cancer-related pain whose baseline pain is adequately controlled” [2] or as “a transient exacerbation of pain that occurs either spontaneously, or in relation to a specific predictable or unpredictable trigger, despite relatively stable and adequately controlled background pain” [9]. The latter yielded “some diagnostic criteria for BTcP [9]” and later, the Diagnostic Breakthrough Pain Algorithm showing, however, limited sensitivity [10]. Over time, experts have questioned the need to include elements such as regular opioid medication and controlled BP in the BTcP definition [11, 12]. The World Health Organization (WHO) defines BTP as a “transitory flare of pain in the setting of chronic pain managed with pain medicines around the clock” [13]. However, this definition of BTP may erroneously include pain exacerbation during opioid titration or due to end-of-dose failure (EODF) of around-the-clock (ATC) opioid medication [14, 15]. Further, it does not address the issue of the daily frequency and duration of flairs.

Effective management of BTP requires reliable identification and ongoing assessment [7]. Current guidelines for the diagnosis and management of BTcP are of low grade [16], and common pain assessment tools are inadequate for the identification of BTP [6, 17]. Further, the literature contains reports of challenges to the adequacy of currently used reporting tools to assess adult BTP. Indeed, there is no widely accepted BTP definition, classification system,

or formally validated BTP diagnostic tool with demonstrated reliability [6, 7].

In previous publications we have reported the development of a scoring system for a BTP diagnostic/prognostic tool based on an 11-item validated questionnaire (Italian Questionnaire for Breakthrough Pain [IQ-BTP]) [18, 19]. This tool enables clinicians to recognize potential-BTP presence using the information collected on BP, ATC opioids medication, EODF, and daily frequency and duration of flairs. The purpose of this tool is to classify the presence of potential-BTP into three prognostic/likelihood classes, namely, high, intermediate, and low, which may also have therapeutic implications.

Following completion of the studies on the initial validation of the IQ-BTP and the performance of its scoring system [18, 19], we sought to formally validate this tool in a multicenter cohort of cancer patients, following the Consensus-based standards to select health Measurement Instruments (COSMIN) taxonomy [20]. We hypothesized that the IQ-BTP will show satisfactory psychometric and validity properties.

## METHODS

### Settings and Patients

This was a multicenter, prospective, and observational study involving seven Italian health-care facilities that managed the palliative care and pain management of cancer patients. The leading center was the acute and CP center of Bologna’s Teaching Hospital, Italy.

Inclusion criteria were patients aged  $\geq 18$  years who had been diagnosed with any form of cancer and who had pain or were receiving ongoing pain therapy for  $> 7$  days, and who signed an informed consent form for participation in the study. The exclusion criteria were surgery within the past 48 h before enrollment, and inability to comprehend the IQ-BTP items.

As the IQ-BTP includes 11 items, the sample size was planned to be of at least ten patients per item (i.e., 110 patients and 330 evaluations at three consecutive visits).

## Procedure and Instruments

Following approval of the study by the leading center's Ethical Committee, satellite centers were invited to participate. Thus, in a preliminary panel meeting of 12 experts from the participating centers, the BTcP operational case definition, the construct of interest, and the relative IQ-BTP questionnaire items were illustrated and explained, including the aims, organization, and structure of the study. On this occasion, the panel also rated the IQ-BTP's items for their content and face validity (see below) and resolved raised issues by consensus. Only for the first item ("In the past 3–7 days, did you have continuous pain lasting  $\geq 12$  h a day") did the panel agree that 'continuous pain' also implies continuous pain treatment. The IQ-BTP was accordingly updated. Thus, each center applied for study approval from its own local Ethical Committee and communicated the approval to the leading center. Once approval had been received, meetings were held at each center to explain further the study's tools (information form and written informed consent form for patients; source document). Each center informed the leading Ethical Committee on the date of its first patient enrollment. Several months after the study started, further meetings were held at each center for peer debriefing. Each month of the study period, centers provided the accumulated number of enrolled patients. As each center started enrollment at different time points, the study's time frame was from July 2017 to December 2019. Centers stopped collecting data after enrolling a maximum of 80 patients per center or when the study's timeframe ended. Gathered data were imputed in an electronic database at the leading center.

The study included three consecutive visits (V1–V3) for each patient; physicians filled the source document at each visit. The source document included four sections. The first section included patients' demographic and clinical details; the second and third sections included the Brief Pain Inventory (BPI) and the IQ-BTP questionnaires; the fourth section included Likert-type questions for the physicians' autonomous evaluation of (1) the presence of

BTcP based on their routine practice and regardless of the IQ-BTP outcomes, and (2) the IQ-BTP contribution to a diagnosis and treatment of BTcP (see Table 4). The evaluating physicians were expert oncologists or pain physicians with expertise in BTP issues. These evaluations will be referred to as the 'gold standard.'

The IQ-BTP is a physician-administered questionnaire and is based on a previously reported BTP operational case definition [18]. Accordingly, a patient with BTcP needs to report: (1) five congruent clinical prerequisite elements (ongoing pain or pain treatment in the past 3–7 days; ATC opioid medications; controlled BP [numeric rating scale {NRS}  $\leq 4$ ]; flair [NRS  $\geq 6$ ] occurrence; and no EODF); (2) two discriminative characteristics of flares, namely, limited frequency ( $\leq 6/24$  h) and limited duration ( $\leq 30$ – $60$  min); and (3) clinical descriptive elements of the flairs (localization, predictability, cause, and physiopathology). (See Electronic Supplementary Material [ESM] Tables A and B for the English and Italian version of the IQ-BTP questionnaire, respectively).

It is assumed that patients who have fewer than the five prerequisite elements cannot have BTcP, while those who have all of the prerequisite elements potentially experience BTcP. Accordingly, in this study, the former patients are collectively referred to as the no-BTcP class and the latter as the potential-BTcP class (i.e., the 2 IQ-BTP classes). In the potential-BTcP class, the BTcP likelihood is high when both the defined flairs' discriminative elements (frequency and duration) are present; intermediate if only one is present; and low if none is present. As IQ-BTP items are not interchangeable and together form this tool's construct, the underlying model is formative [20].

Formal validation of the IQ-BTP follows the indications for measurement properties included in the COSMIN taxonomy main domains [20]. The domains assessed in this study are: (1) Validity, including content and face validity, construct validity (analysis/reduction of items; hypothesis testing; cross-cultural validity\measurement invariance), and criterion validity; (2) Reliability (internal consistency and reliability); (3) Interpretability; and (4) Responsiveness. A

**Table 1** Demographics and cancer-related features of the patient sample

Predictor	Predictor subset	Total <i>n</i> (%)
Observations		753 (100.0)
Setting	Oncology office	246 (32.7)
	Pain medicine office	197 (26.2)
	Oncology day-hospital	100 (13.3)
	Hospice	72 (9.6)
	Hospital ward	60 (8.0)
	Home care	78 (10.4)
Gender	Male	346 (45.9)
	Female	407 (54.1)
Karnofsky <sup>a</sup>		66.2 (22.3)
Age group (years)	A (18–35)	12 (1.6)
	B (36–50)	72 (9.6)
	C (51–65)	214 (28.4)
	D (66–80)	335 (44.5)
	E (> 80)	120 (15.9)
Primitive tumor site	Lungs	130 (17.1)
	Breast	114 (15.0)
	Colorectal	110 (14.4)
	Pancreas	109 (14.3)
	Urogenital (Females)	75 (9.8)
	Urogenital (Males)	65 (8.5)
	Hematologic	41 (5.4)
	Liver	25 (3.3)
	Head/throat	30 (3.9)
	Stomach	17 (2.2)
	Kidney	12 (1.6)
	Central nervous system	5 (0.7)
	Other	29 (3.8)
	> 1 site <sup>b</sup>	9 (1.2)

<sup>a</sup> Values presented as the mean (standard deviation [SD])

<sup>b</sup> Cases with more than one primitive tumor site

**Table 2** Outcomes of Italian Questionnaire for BTcP diagnosis,  $\chi^2$  analysis, and Spearman rank correlation analysis

IQ-BTP		Total <i>n</i> (%)	$\chi^2$ <sup>c</sup>	Spearman rank correlation ( $\rho$ value) <sup>e</sup>
Item	IQ-BTP subset			
Sample		753 (100.0)		
1—Pain/pain therapy present			Constant <sup>d</sup>	Constant <sup>d</sup>
	Yes	753 (100.0)		
	No	0 (0.0)		
2—Strong opioid therapy			412.017	0.237 Mild
	Yes	655 (87.0)		
	No	98 (13.0)		
3—BP intensity (NRS $\leq$ 4)			226.519	0.330 Mild
	Yes	583 (77.4)		
	No	170 (22.6)		
4—Pain from flairs (NRS $\geq$ 6)			2.689	0.649 Moderate
	Yes	354 (47.0)		
	No	399 (53.0)		
5—Flairs are not related to therapy schedule?			156.240	0.420 Moderate
	Yes	512 (68.0)		
	No	241 (32.0)		
Potential-BTP <sup>a</sup>		205 (27.2)	–	–
6—Flair frequency (24 h)			534.558	0.989 Strong
	< 5 = 1	133 (17.7)		
	$\geq$ 5 = 2	72 (9.6)		
7—Flair duration			555.355	0.992 Strong
	$\leq$ 30–60 min = 1	162 (21.5)		
	> 60 min = 2	43 (5.7)		
BTP likelihood <sup>b</sup>			939.074	0.987 Strong
	Low (items 6 and 7 = 4)	18 (2.4)		
	Intermediate (items 6 and 7 = 3)	79 (10.5)		
	High (items 6 and 7 = 2)	108 (14.3)		
	Total	205 (27.2)		
8—Different sites for flair and BP?			540.462	0.266 Mild
	Yes	66 (8.8)		
	No	139 (18.5)		

**Table 2** continued

IQ-BTP		Total	$\chi^2$ <sup>c</sup>	Spearman rank correlation
Item	IQ-BTP subset	n (%)		( $\rho$ value) <sup>e</sup>
9—Flairs are predictable?			540.215	0.989 Strong
	Yes	53 (7.0)		
	No	152 (20.2)		
10—Flair cause known?			531.179	0.988 Strong
	Yes	80 (10.6)		
	No	125 (16.6)		
11—Dysesthesias/paresthesias present?			527.243	0.988 Strong
	Yes	99 (13.1)		
	No	106 (14.1)		

*BP* Background pain, *BTcp* breakthrough pain in cancer, *BTP* breakthrough pain,  $\chi^2$  Chi-square, *IQ-BTP* Italian Questionnaire for BTcP, *NRS* numeric rating scale

<sup>a</sup> Cases with ‘yes’ answers throughout items 1–5 (yielding potential-BTcP)

<sup>b</sup> Sum of items 6 and 7 outcomes (= 4, 3, or 2), yielding low, intermediate, and high BTP likelihood, respectively

<sup>c</sup>  $\chi^2$ -analysis, association of IQ-BTcP items and potential-BTcP; *P*-value = 0.000, respectively

<sup>d</sup> All cases were ‘Yes’ thus constant

<sup>e</sup> Spearman-rank correlation between IQ-BTcP items and the IQ-BTP outcomes; results are expressed as  $\rho$  absolute values

concise description of the definition, hypotheses evaluation, and statistical methods used in this study is reported in ESM Table C (Validation summary).

Briefly, for face validity, we presented, during a preliminary meeting, the BTP case definition and IQ-BTP items to a panel of 12 experts. Experts were physicians with established specialist expertise in CP and palliative care. The experts rated each item (using a five-level Likert-type scale [Strongly agree, Agree, Indifferent, Disagree, and Strongly disagree]) on its comprehensiveness, comprehensibility, and relevance to the BTP case definition as well as the grammar, wording, and randomization adequacy of each item; they added, if necessary, observations/suggestions. An item was adopted if > 75% of the experts rated it as “Strongly agree”/“Agree” for its relevance and adequacy. Content validity is the degree to which the

content of a measure is an adequate reflection of the construct to be measured [20]. Accordingly, to support the IQ-BTP content validity, we further hypothesized that: (1) there should be significant associations between IQ-BTP items and the IQ-BTP classes; (2) the content validity ratio (CVR) should be positive and > 0.5 [21]; and (3) agreement between the gold standard and IQ-BTP outcomes should be > 75% and Cohen’s  $\kappa$  should be > 0.5.

To support the IQ-BTP construct validity, we first used principal component analysis (PCA) and the scree test to identify components that are composites of this tool’s items. We further hypothesized that: (1) the scores of patients of different IQ-BTP classes for the BPI items will differ significantly; (2) the strength of the Spearman rank correlation between IQ-BTP classes and BPI items, given their dissimilar construct, will be significant but less than

**Table 3** Physicians' (gold standard) independent evaluations

Physician opinion		Total
Item	Subset	n (%)
Sample		753 (100.0)
BTcP is present (regardless the IQ-BTP outcome)? <sup>a</sup>	Yes	186 (24.7)
	No	507 (67.3)
	Don't know	60 (8.0)
IQ-BTP was helpful for BTcP diagnosis? <sup>b</sup>	Very much	276 (36.7)
	Much	373 (49.5)
	Don't know	57 (7.6)
	Slightly	35 (4.6)
	Not at all	12 (1.6)
BTcP likelihood-class may guide BTcP therapy? <sup>c</sup>	Yes	711 (94.4)
	No	6 (0.8)
	Don't know	36 (4.8)

<sup>a</sup> Physicians independent evaluations of BTcP presence

<sup>b</sup> Physicians evaluations of the IQ-BTP usefulness for BTcP diagnosis

<sup>c</sup> Physicians evaluations of the IQ-BTP usefulness for BTcP treatment

moderate (mean  $\rho < 0.4$ ) [22]; (3) the IQ-BTP will show measurement invariance (MI) for cross-cultural validity as there will be no important differences in IQ-BTP class rates between gender groups (for the latter, the null hypothesis, applying ordinal logistic regression, was that IQ-BTP class rates differ between gender groups).

For criterion validity, we assessed the correlation between IQ-BTP and the gold standard outcomes; we hypothesized that the area under the curve (AUC) would be  $\geq 0.70$  [20].

To support the internal consistency and reliability of the IQ-BTP, we hypothesized that

Cronbach's  $\alpha$  and the intraclass correlation coefficient (ICC) would both exceed 0.7. [20, 23, 24]. The reliability of the IQ-BTP was further assessed by defining its sensitivity, specificity, and the prior probability against the gold standard outcomes [25].

Interpretability was assessed by verifying whether clinicians can assign qualitative meaning (i.e., clinically understood connotations) to the IQ-BTP's outcomes [26]. Hence, we assessed whether  $> 75\%$  of the gold standard evaluations considered the IQ-BTP outcomes as a valid guide for BTcP therapeutic indications. Finally, we compared the IQ-BTP and gold standard outcomes for responsiveness analysis; we hypothesized that their agreement rate would be  $> 75\%$ , Cohen's coefficient  $> 0.5$ , and the AUC  $\geq 0.70$  [20].

#### Data Presentation and Statistical Analysis

Continuous data are reported as the mean  $\pm$  standard deviation (SD); when appropriate, the median and 95% upper and lower confidence intervals (CI) are reported. Categorical data and proportions are reported as absolute numbers and percentages. Analysis of variance (ANOVA) was used to assess the differences in BPI item scores between IQ-BTP classes. The latter dependence upon independent variable categories was determined using  $\chi^2$  analysis. The Spearman rank correlation analysis was used to analyze the correlation between IQ-BTP classes, BPI, and IQ-BTP items, respectively; when statistically significant, an absolute  $\rho$  value  $< 0.2$  was considered to indicate a poor correlation, 0.2–0.4 a mild correlation, 0.41–0.7 a moderate correlation, and 0.71–1.0 a strong correlation [25]. Cronbach's  $\alpha$  test was used to test internal consistency and intraclass correlation (ICC) for reliability. Contingency tables of the IQ-BTP outcomes against the gold standard were used to define the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the IQ-BTP; results are expressed in percentages and reported with their CI and margin of error (M; one-half of the width of the CI and summarizes the width of a CI relative to the whole possible range) [27]. Prior probability



**Table 4** Brief Pain Inventory item scores within the sample and the IQ-BTP classes, analysis of variance results, and Spearman rank correlation

BPI and QoL items	Sample <sup>a</sup>	IQ-BTP classes <sup>a</sup>		ANOVA <i>P</i> value*	Spearman rank correlation ( $\rho$ value) <sup>b</sup>
		Potential-BTcP	No-BTcP		
Worst pain	4.8 (4.6–5.0)	6.4 (6.1–6.6)	4.2 (4.0–4.5)	0.000	0.348 Mild
Least pain	1.5 (1.4–1.6)	1.8 (1.6–2.0)	1.4 (1.2–1.5)	0.004	0.156 Poor
Mean pain	2.9 (2.7–3.0)	3.5 (3.3–3.7)	2.6 (2.5–2.8)	0.000	0.233 Mild
Actual pain	2.1 (1.9–2.2)	2.5 (2.3–2.8)	1.9 (1.7–2.1)	0.000	0.179 Poor
Dynamic pain	3.9 (3.7–4.1)	5.3 (5.0–5.7)	3.3 (3.1–3.6)	0.000	0.324 Mild
Pain relief with therapy (%)	70.1 (68.3–71.8)	65.8 (63.2–68.5)	71.7 (69.5–73.9)	0.004	0.165 Poor
QoL items (pain Interference with...)					
General activity	3.7 (3.5–4.0)	4.7 (4.4–5.1)	3.4 (3.1–3.6)	0.000	0.214 Mild
Mood	3.8 (3.6–4.0)	4.7 (4.3–5.0)	3.5 (3.2–3.7)	0.000	0.202 Mild
Ambulation	3.3 (3.1–3.5)	4.0 (3.5–4.4)	3.1 (2.8–3.3)	0.001	0.135 Poor
Work	3.8 (3.6–4.1)	4.6 (4.2–5.0)	3.6 (3.3–3.9)	0.000	0.152 Poor
Social interactions	2.9 (2.7–3.1)	3.5 (3.2–3.9)	2.7 (2.5–2.9)	0.000	0.161 Poor
Sleep	2.6 (2.4–2.8)	2.9 (2.9–3.7)	3.3 (2.1–2.6)	0.000	0.174 Poor
Enjoyment of life	3.5 (3.3–3.7)	4.4 (4.0–4.8)	3.2 (2.9–3.4)	0.000	0.197 Poor

*ANOVA* Analysis of variance, *BPI* Brief Pain Inventory, *QoL* quality of life

\*Statistical significance  $P < 0.001$  for all items

<sup>a</sup> No-BTcP class: patients who have < 5 of the prerequisite elements and therefore do not have BTcP; Potential-BTcP class: those who have all of the prerequisite elements and potentially experience BTcPBPI. Values are presented as the mean score with the 95% confidence interval (CI) in parenthesis

<sup>b</sup> For the Spearman rank correlation between BPI items and the IQ-BTP classes, results are expressed as the  $\rho$  absolute values

was defined by calculating the PPV and PNV [28]. For agreement analysis, we first calculated the index condition rate (IQ-BTP outcome, potential-BTcP) and, thus, the rate of the overall agreement with the gold standard; Cohen’s  $\kappa$  test was hence applied [27]. When statistically significant, an absolute  $\kappa$  value (with its standard error [SE] of 0.1–0.3 was considered to indicate mild agreement, 0.31–0.5 to indicate moderate agreement, and 0.51–1.0 to indicate excellent agreement [25]. Ordinal logistic

regression analysis was conducted to investigate whether the variable gender influences the IQ-BTP outcomes. Statistical significance was defined as  $P < 0.05$ . When appropriate,  $P < 0.01$  and  $P < 0.001$  were reported.

**Ethics**

The study was authorized by the Ethics Committee of each participating center

**Table 5** Outcomes of IQ-BTP and gold standard contingency table (BTcP presence)

IQ-BTP (new test)	Physician judgment (gold standard)		
	Positive	Negative	Total
Positive	128	77	205
Negative	58	490	548
Total	186	567	753

Accuracy	%	95% CI	M
Sensitivity	69	65.5–72.1	3.3
Specificity	86	84.0–88.9	2.4
Positive predictive value	62	59.0–65.9	3.5
Negative predictive value	75	87.2–91.6	2.2

Positive and negative outcomes imply the presence or absence of BTP, respectively  
M, Margin of error (one half width of CI)

(authorization number of the leading center [Ethical Committee of AOSP di Bologna, Policlinico S. Orsola-Malpighi]: 46/2015/U/oss). Names and reference numbers of the Ethics Committees that have approved this study is provided in ESM Table D. The study was conducted according to the Helsinki Declaration of 1964 and its later amendments and IASP's guidelines for pain research in animals and humans. The investigators personally and thoroughly informed all participants on the study's aims and structure. Patients were informed that participation was voluntary and anonymous and that it would not affect their care; hence, informed consent was obtained from each patient.

## RESULTS

Of the 280 enrolled patients, 75 died before V3, yielding a total of 753 evaluations from V1 to V3, which satisfied the sample size requirements. Table 1 reports the sociodemographic and clinical features of the evaluated patients. Of the seven participating centers, 32.7% of evaluations were made at oncology offices, 26.2% at pain offices, 13.3% at oncology day-hospitals, and roughly 10% each in a hospice, hospital ward, or home care environment.

Among the evaluated patients, 54.1% were females; the mean ( $\pm$  SD) age was  $67.7 \pm 13.4$  (range 23–94) years, and 45% belonged to age group D (66–80 years). In 60.8% of patients, primitive tumor sites involved the lungs, breast, colon/rectum, and pancreas.

Table 2 reports the IQ-BTP questionnaire outcomes and the results of the  $\chi^2$  analysis for the associations between the IQ-BTP items and classes. Based on the presence of the prerequisite elements (i.e., cases with 'YES' answers for IQ-BTP items 1–5), potential-BTcP was found in 205 evaluations (27.2%). Of the potential-BTcP cases, and based on the discriminative elements (i.e., IQ-BTP items 6 and 7), BTcP likelihood was high at 52.7%, intermediate at 38.5, and low at 8.8%). Based on the answers to IQ-BTP items 8–11, BTcP and BP sites were similar in 67.8% of the potential-BTcP cases; BTcP was unpredictable in 74.1%, of unknown cause in 61%, and without neuropathic features in 51.7%.

The  $\chi^2$  analysis showed significant associations between all IQ-BTP items and the outcome potential-BTP ( $P$  value = 0.000), except for item 1 (Pain/ongoing pain-therapy), which was a constant.

Table 3 reports the physicians' opinion on the presence of BTcP and the usefulness of the

**Table 6** Principal component analysis explaining total variance

Component	Initial eigenvalues			Factor weight					
	Total	Variance	Cumulative	Not-rotated			Rotated		
				Total	Variance	Cumulative	Total	Variance	Cumulative
1	5.431	54.31	54.31	5.431	54.31	54.31	5.165	51.65	51.65
2	1.224	12.24	66.55	1.224	12.24	66.55	1.490	14.90	66.55

For each component, with eigenvalue greater than unity, total values, variance and cumulative proportions of initial eigenvalues along with rotated and not-rotated factors weight are reported. Variance and cumulative values are expressed in percentages

IQ-BTP for its diagnosis and treatment (i.e., the gold standard). Without consulting the IQ-BTP outcomes, physicians diagnosed BTcP in 24.7% of all evaluations. Physicians considered IQ-BTP much/very much helpful for BTcP diagnosis in 86.2% of the evaluations. They considered the BTcP likelihood classification useful as a guide for BTcP treatment in 94.4% of the cases. These data will be used to define the IQ-BTP reliability and content validity (see below).

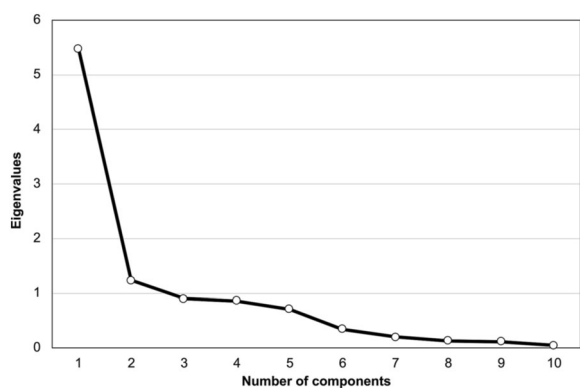
Table 4 reports the scores for the BPI items within the sample and the IQ-BTP classes and details the results of the ANOVA and the Spearman-Rank Correlation analyses. For all BPI items, the mean scores of the potential-BTcP class were higher than those of the No-BTcP class except for the percentage of pain relief with the therapy, which was lower. Indeed, ANOVA analysis showed significant differences between the two classes (ANOVA,  $P < 0.01$ ). In

particular, for the BPI items ‘Worst pain in the past 24 h’ and ‘Dynamic pain,’ the difference between the two IQ-BTP classes was  $\geq 2$  points. For correlations between BPI items and the IQ-BTP classes analysis, see section [Construct Validity](#).

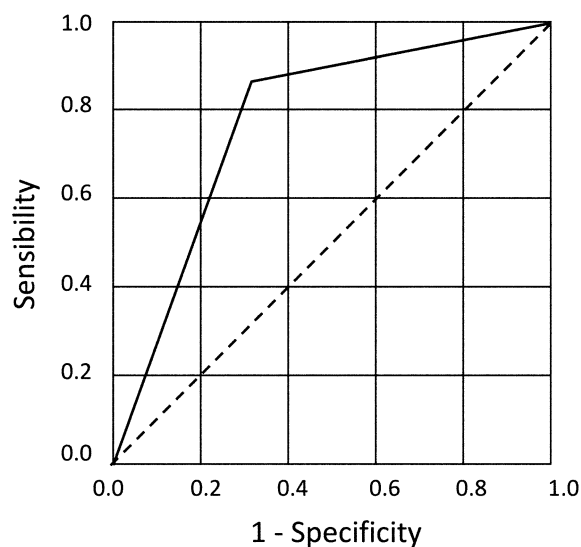
**Validation**

**Reliability: Internal Consistency and ICC**

Internal consistency and ICC of the IQ-BTP computations yielded Cronbach’s  $\alpha$  and an ICC of 0.766, respectively (ICC,  $P = 0.000$ , 95% CI 0.741–0.790).



**Fig. 1** Scree test. Two components show eigenvalues larger than unity



**Fig. 2** The area under the curve of the correlation between the gold standard and the IQ-BTP classes. The area is 0.776

Table 5 is a contingency table of the frequency distribution of the gold standard and IQ-BTP outcomes in terms of BTcP presence. Positive and negative outcomes imply the presence or absence of BTcP, respectively. It also reports the sensitivity, specificity, and prior probability analyses along with their lower and upper CIs and its margin of error (M). The rate of the index condition (potential-BTcP) was 27.2%, and the overall agreement between the outcomes of the IQ-BTP and the gold standard for the presence of BTcP was 82%. Cohen's  $\kappa$  was 0.535 (SE 0.035), showing excellent agreement. The IQ-BTP test showed satisfactory values of reliability.; in particular, sensitivity was 69%, specificity was 86%, PPV was 62%, and NPV was 75% with particularly low M values.

### Validity

**Content Validity (associations, CVR, and agreement)** Experts' opinion on the comprehensiveness, comprehensibility, and relevance of the IQ-BTP items to the construct of interest and on the adequacy of the items yielded an overall approval for all items' face/content validity, with 98.5% of the experts responding 'Strongly agree'/'Agree' on their relevance and adequacy, respectively.

Significant associations were found between all IQ-BTP items and IQ-BTP classes (Table 3;  $\chi^2$  analysis,  $P < 0.001$ , respectively). The CVR for the gold standard was calculated for the number of raters who considered the IQ-BTP as 'Much'/'Very much' helpful for BTcP diagnosis ( $n = 649$ , CVR 0.7) and for those who judged the BTcP likelihood classification as a useful guide for BTcP treatment ( $n = 711$ , CVR 0.9). Thus, in both analyses, CVR was extensively above the 0.5 cutoff. As mentioned earlier, the overall agreement between the gold standard and IQ-BTP outcomes was  $> 75\%$  and Cohen's  $\kappa > 0.5$ , showing excellent agreement.

### Construct Validity

**Correlations Between IQ-BTP Classes and BPI Items** As shown in Table 4, correlations between BPI items and the IQ-BTP classes were significant (Spearman rank correlation,  $P < 0.01$ ). Absolute  $\rho$  values, however, ranged

from 0.135 to 0.348 (mean  $0.200 \pm 0.07$ ). In particular,  $\rho$  values yielded mild correlations of IQ-BTP classes with the three pain items (worst, mean, and dynamic pain) and two pain interference-with-QoL items (general activity and mood). Correlations were poor with the remaining eight BPI items. Conversely, as shown in Table 2, correlations between IQ-BTP items and the IQ-BTP classes were significant (Spearman rank correlation,  $P < 0.001$ ), and absolute  $\rho$  values ranged from 0.237 to 0.989 (mean  $0.712 \pm 0.33$ ), showing high strength of the correlations. The different strength of correlations between the IQ-BTP classes and the two questionnaires (BPI and the IQ-BTP) imply distinct constructs and support the IQ-BTP autonomous construct validity.

**Principal Component Analysis** The hypothesis that the IQ-BTP items were uncorrelated was ruled out as Bartlett's test of sphericity was statistically significant ( $P < 0.0001$ ) and the Kaiser–Meyer–Olkin (KMO) statistic was  $> 0.6$  (KMO 0.914). The PCA (and scree test (Fig. 1) of the IQ-BTP identified two components with eigenvalues larger than unity. Table 6 shows, for each identified component, the total value, variance, and cumulative proportions of the initial eigenvalues together with rotated and not-rotated factor weights. These factors accounted for 66.6% of the variance. According to the Varimax rotation method with Kaiser normalization, the rotated matrix of these components reached convergence criteria with three iterations. The first factor, accounting for 54.3% of the variance, included items relative to the presence, duration, frequency, predictability, and causes of flares. The second factor, accounting for 12.2% of the variance, included items relative to ATC opioid use, BP, and EODF.

**Criterion Validity** Figure 2 describes the AUC of the correlation between the gold standard (independent physician judgment for the presence of BTcP) and the IQ-BTP classes. In particular, the AUC was found to be 0.776 (95% CI 0.734–0.819) and thus  $> 0.70$ .

**Measurement Invariance** We conducted an ordinal logistic regression analysis to investigate

whether gender variables influence the IQ-BTP outcomes. The predictor variables were tested a priori to verify there was no violation of the assumption of no multicollinearity. The null hypothesis, i.e., IQ-BTP class rates differ between gender groups, was rejected as the predictor variable gender (females) was not found to contribute to the model (estimate  $-0.196$ , SE  $0.165$ , Wald  $1.396$ ,  $P = 0.237$ ). Indeed, potential-BTcP rates within the male and female subgroups were similar (29.0 and 24.4%, respectively).

## DISCUSSION

In a multicenter cohort of cancer patients and using the IQ-BTP for 753 evaluations, the rate of potential-BTcP was 27.2%. Of the latter cases, BTcP likelihood was high in 52.7% of patients, intermediate in 38.5, and low in 8.8%.

Correct identification is imperative for BTcP appropriate pain management. Currently, standard pain assessment tools are inadequate for BTcP identification [6, 17], and the adequacy of BTcP diagnostic tools is hampered due to the lack of a widely accepted definition of BTcP and formally validated BTcP diagnostic tools with demonstrated reliability [6, 7]. Among existing BTP assessment tools, only the breakthrough pain assessment tool (BAT) [29] meets some of the COSMIN criteria. However, the BAT is not a diagnostic tool; it only assesses the characteristics of otherwise empirically diagnosed BTP [7]. In comparison, the IQ-BTP may be used for the diagnosis and epidemiological evaluation of BTcP.

The originality of the IQ-BTP questionnaire and its scoring system lies in its prognostic ability [18, 19]. It identifies potential-BTcP presence (based on prerequisite elements) and then its risk (based on discriminative elements) as high, intermediate, or low likelihood.

The five IQ-BTP prerequisite elements to identify potential-BTcP are ongoing pain/pain-treatment in the past 3–7 days; ATC opioid medications; controlled BP; flair occurrence; and no EODF. These elements complete the WHO's definition of BTP, which otherwise may erroneously identify as BTcP flairs during opioid

titration or due to opioid medications' EODF [14, 15]. Moreover, for BTcP identification, the inclusion of items relating to ATC opioid medications and controlled BP is supported by the US Food and Drug Administration (FDA) recommendations. Indeed, as the appropriate treatment for BTcP is currently rapid-onset opioids (ROOs; e.g., transmucosal fentanyl citrate) [17, 30], the FDA recommends that safe use of this medication requires that patients be opioid tolerant based on concurrent regular use of opioid medication [31]. A controlled BP condition is essential as uncontrolled BP implies that opioid titration is not completed and excludes the advent of BTP [9, 14]. In our BTcP case definition, the expert panel adopted the cutoff score of  $NRS \leq 4$  for controlled BP (i.e., none/mild pain using the verbal rating scale). In the literature, Webber et al. also used the same cutoff for their BTcP diagnostic algorithm reliability study [10]. In the latter study, the algorithm's sensitivity was limited using the "mild" cutoff to define controlled BP and increased (while specificity strongly decreased) when the cutoff level was moderate BP. Our study showed that the IQ-BTP has satisfactory values for both specificity and sensitivity with the adopted controlled BP cutoff (see below). We believe that higher cutoff scores for controlled BP preclude the efficacy of the therapy for the BP and thus are incompatible with BTcP presence.

The presence of all prerequisite elements is suggestive to the caregiver that the patient can be considered as having potential-BTcP presence; however, this is not a sufficient basis for therapeutic or clinical decisions. The frequency and duration of flairs are crucial aspects for BTcP recognition and have a strong relevance for therapeutic decisions. Thus, any evaluation of the likelihood of BTcP based on the quantitative features of flairs may help caregivers in their clinical decision-making. ROOs, notably with a limited number of allowed daily administrations and brief action, are the recommended medications for BTcP. However, the treatment of potential-BTcP with flairs of high frequency ( $> 6$ ), long duration ( $> 60$  min), or both will be incompatible with the known pharmacological features of the mentioned ROOs. Discriminative items used by the IQ-BTP to classify the BTcP

likelihood are the daily frequency of flairs (< 7/24 h) and the duration of flairs ( $\leq$  30–60 min). Accordingly, among patients who potentially experience BTcP, the latter's likelihood is high when both the defined discriminative items are present; intermediate if only one is present, and low if none of them are present [18, 19].

Clinical prognosis infers the risk of outcomes in people with a given health condition and provides health stakeholders with reliable evidence for decision-making and cost-effectiveness of care [32]. It is reasonable to speculate that for BTcP of high likelihood, the congruent treatment would be with ROOs. An intermediate or low likelihood of BTcP imposes careful evaluation of the opportunity to use short-acting opioids (SAO; e.g., oral morphine sulfate) or to ameliorate the ATC opioid regimen, respectively [18, 19]. These observations support the interpretability of the IQ-BTP as clinicians can assign clinically understood connotations to the IQ-BTP's outcomes [26]. Indeed, our findings show that practicing physicians considered the IQ-BTP to be 'Much'/'Very much' helpful for BTcP diagnosis in 86.2% of the evaluations and the BTcP likelihood classification to be a useful guide for its treatment in 94.4% of the cases.

Support for the IQ-BTP face and content validity comes from several sources. First, the panel of 12 experts found the questionnaire's items to be comprehensive, comprehensible, and relevant to the underlying construct and to be adequate. By consensus, the panel preliminarily adjusted the first item on the IQ-BTP to include pain treatment in the past 3–7 days with the intention to classify a pain patient. Indeed, a patient can report limited or no BP because of an ongoing efficacious pain treatment; however, this condition is still susceptible for BTcP to occur. Secondly, we found that the IQ-BTP adequately reflected the construct to be measured. Indeed, IQ-BTP items were significantly associated with IQ-BTP classes. The CVR was positive and  $> 0.5$ , and the agreement (rate and Cohen's  $\kappa$  between gold standard and IQ-BTP outcomes was high.

Evidence for the construct validity hypotheses of the IQ-BTP comes from the PCA, the strength of correlation between its classes and the BPI items, meaningful changes between

relevant subgroups, cross-cultural validity\measurement invariance, and responsiveness. Following our operational case definition and the construct hypothesis, IQ-BTP items loaded on two factors in the PCA. The first factor, which can be named 'flairs' features,' included items relative to the presence, duration, frequency, predictability, and causes of flares. The second factor, called 'BP features,' included items relative to opioid use, BP, and EODF presence.

Correlations between BPI items and the BTcP classes were significant but of mild or poor strength (mean  $\rho$  values 0.200). Such a low strength of correlation is evidence of the two measures being of a different construct [20, 22]. Both measures are a two-factor model; however, they differ by the items these factors load and thus they measure unrelated constructs. The BPI factors load pain intensity and pain interference with quality of life (QoL) items, respectively; BPI does not include items essential for the diagnosis of BTP [6, 18] (e.g., BP treatment, EODF, or quantitative features of flairs). The IQ-BTP's two factors, unlike the BPI, include these items and objectively assess them. This differentiation between the two measures explains their poor correlation and confirms the autonomous construct of the IQ-BTP. Interestingly, correlations between IQ-BTP items and BTcP classes were significant and of high strength (mean  $\rho$  value 0.712). The different strength of correlations between the IQ-BTP classes and the two measures (BPI and the IQ-BTP) further support their distinct constructs and hence the validity of the IQ-BTP construct.

Meaningful differences between the IQ-BTP classes (i.e., relevant subgroups) were shown for all scores of the BPI items. In CP literature, the concept of minimal clinically important change (MCIC) is used to describe clinically significant improvements in pain scores (e.g., using the NRS) following congruent treatment. Aside from significant differences, MCIC requires a change of  $\geq 2$  points in comparisons of NRS scores [33, 34]. Indeed, for all BPI items, the mean scores of the potential-BTcP class were significantly higher than those of the no-BTcP class. In particular, for the BPI item 'worst pain in the past 24 h,' relevant for BTcP patients, and

‘dynamic pain,’ an often cause of BTcP, the difference between the two IQ-BTP classes was  $\geq 2$  points. These results confirm meaningful changes between relevant IQ-BTP subgroups—i.e., patients with expected high (potential-BTcP) versus low (no-BTcP) levels of the construct of interest.

Cross-cultural validity is evaluated in the IQ-BTP by assessing whether the scale is measurement invariant. MI refers to whether respondents from different groups with the same latent trait level respond similarly to a particular item [20]. We evaluated whether the rates of the IQ-BTP classes are similar within the genders subgroups (male vs. female). The null hypothesis was that ordinal logistic regression would show that the class rates of the IQ-BTP differ between gender groups (i.e., the predictor variable gender [females] significantly contributes to the model). As the null hypothesis was rejected, potential-BTcP rates within the male and female subgroups were similar, supporting the IQ-BTP’s MI and its cross-cultural validity.

A measure’s criterion validity refers to the degree to which its scores are an adequate reflection of a gold standard [20]. We assumed as gold standard the independent physician’s judgment for the presence of BTcP. The AUC of the correlation between the gold standard and the IQ-BTP classes was  $> 0.70$  (AUC 0.776), thus confirming the IQ-BTP’s criterion validity and responsiveness (see below). The sensitivity and specificity of the IQ-BTP may further support its criterion validity. The IQ-BTP showed satisfactory values of sensitivity (69%), specificity (86%), PPV (62%), and NPV (75%), with shallow M values. Sensitivity refers to the test’s true positive rate, while specificity refers to its true negative rate. Our results confirm the ability of the IQ-BTP to detect (or rule out) BTcP presence. Prior probability estimates the congruency of the IQ-BTP with the clinical context in which it was assessed. PPV and NPV describe the likelihood of the condition of interest given the positive or negative test result, respectively. Our results show that the population in which the IQ-BTP was tested represents the clinical context in which the test is to be applied, and the studied cohort (i.e., patients with cancer pain) was congruent with the test. To our knowledge,

this is the first time that the sensitivity and specificity features of a BTP diagnostic tool has been reported.

Evidence for the reliability of the IQ-BTP comes from its internal consistency and the ICC analyses. We found that both Cronbach’s  $\alpha$  and ICC are 0.766. Reliability is considered to be acceptable when both Cronbach’s  $\alpha$  and ICC are  $> 0.7$  [20, 23, 24]. As both values are  $> 0.70$ , the IQ-BTP reliability is supported.

Measure responsiveness refers to the ability of the construct to be measured to detect change over time [20] and is less relevant when an instrument, such as the IQ-BTP, is used as a diagnostic instrument. This measure can be assessed in situations in which a gold standard is available and tests hypotheses on expected differences in changes between subgroups. Comparing a measure to a gold standard is considered as evidence for the criterion approach for responsiveness. In this study, we compared the IQ-BTP outcomes (IQ-BTP classes) with the gold standard.

An outcome measure can be described by its clinical capacity to correctly identify individuals who present an important clinical change against an external standard for change [35]. Thus, the test’s sensitivity and specificity may describe its responsiveness to change. Sensitivity and specificity evaluate the measure’s capacity to reflect differences in the change between groups regarding the external gold standard (presence/absence of BTcP). Further, the AUC expresses the instrument’s discriminative ability or the probability of correctly classifying both those patients who show BTcP and those who do not. AUC provides a broad vision of the relationship between a measure and an external standard for change. The sensitivity and specificity of the IQ-BTP, meaningful differences in the scores of BPI items between subgroups (IQ-BTP patient classes), and AUC outcomes confirm the IQ-BTP’s criterion validity and responsiveness.

### Study’s Limitations

Construct validity in this study was based, among others, on the strength of the

correlation between IQ-BTP classes and the BPI, i.e., tools with a dissimilar construct. In the literature, correlation with tools measuring similar constructs may provide more evidence than correlation with tools measuring dissimilar constructs [20]. We chose the BPI, a worldwide used and well-validated pain assessment tool, although one with a dissimilar construct, as the literature on well-validated diagnostic tools for BTcP is limited and may hinder quantitative comparisons [7].

The study does not include a structural analysis. Indeed, structural validity is only relevant for tools that are based on a reflective model. The IQ-BTP items are not interchangeable nor highly correlated. As the IQ-BTP items together form the IQ-BTP construct, the underlying model is a formative one. In the latter model, structural validity is not relevant [20]. We also did not perform the IQ-BTP's measurement error analysis. The latter is the error in a participant's score that is not attributable to the construct being measured; it is analyzed by assessing outcomes over time in patients with stable health status [20]. As in this study, the latter could not be guaranteed over time, we omitted this analysis.

## CONCLUSIONS

The IQ-BTP has undergone extensive formal validation and shown satisfactory psychometric and validation properties. We have demonstrated its content, face, construct, and criterion validities and its reliability, interpretability, and responsiveness. Using the IQ-BTP in cancer patients enables potential-BTcP to be identified and differentiated into three likelihood classes with direct therapeutic and epidemiological implications. The latter may be confirmed in future studies.

## ACKNOWLEDGEMENTS

**Funding.** Kyowa Kirin unconditionally supported the study. The journal's Rapid Service Fee is funded by the corresponding author's

personal budget at the Department of Medical and Surgical Sciences of Bologna's university, Italy.

**Authorship.** All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

**Authors' Contributions.** All authors contributed to the study's conception and design. Material preparation and data collection were performed by Boaz Gedaliahu Samolsky Dekel, Alberto Gori, Marco Gunnellini, Angela Gioia, Mariacristina Di Marco, Giuseppe Casale, Marzio Bevilacqua, Pietro Bersani, and Rita Maria Melotti. Boaz Gedaliahu Samolsky Dekel performed the analyses and wrote the first draft of the manuscript; all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**List of Investigators.** Guido Biasco, Stefano Bravi, Alice Calvieri, Lucia Casagrande, Laura Cavalli, Roberta Cherubini, Michele De Roivo, Valentina Fanti, Eva Freier, Caterina Magnani, Sabrina Marchi, Loredana Minelli, Federica Morteo, Riccardo Rossetti, Cristina Sorella, and Alessio Vasarri.

**Disclosures.** The authors declare that Kyowa Kirin unconditionally supported the study. All authors (Boaz Gedaliahu Samolsky Dekel, Alberto Gori, Marco Gunnellini, Angela Gioia, Mariacristina Di Marco, Giuseppe Casale, Marzio Bevilacqua, Pietro Bersani and Rita Maria Melotti) declare no personal, financial, commercial, or academic conflicts of interest.

**Compliance with Ethics Guidelines.** The study was authorized by each participating centers' Ethical Committee (the leading center Ethical Committee—AOSP di Bologna, Policlinico S. Orsola-Malpighi, Bologna, Italy—authorization number is 46/2015/U/oss). Names and reference numbers of the all ethics committees that has approved this study is provided in a table within the supplementary material file



(Table D). The study was conducted according to the Helsinki Declaration of 1964 and its later amendments and IASP's guidelines for pain research in animals and humans. The investigators personally and thoroughly informed all participants on the study's aims and structure. Patients were informed that participation was voluntary, anonymous, and would not affect their care; hence, informed consent for participating in the study was obtained from each patient.

**Data Availability.** The datasets generated and/or analyzed during the current study are not publicly available as data belongs to the single participating institutions.

**Open Access.** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Fortner BV, Okon TA, Portenoy RK. A survey of pain-related hospitalizations, emergency department visits, and physician office visits reported by cancer patients with and without history of breakthrough pain. *J Pain*. 2002;3:38–44.
- Portenoy RK, Hagen NA. Breakthrough pain: definition, prevalence and characteristics. *Pain*. 1990;41:273–81.
- Bruera E. The challenges of prognosis in cancer pain. *Pain*. 2012;153:513–4.
- Burton B, Zeppetella G. Assessing the impact of breakthrough cancer pain. *Br J Nurs*. 2011;20(S14):S16–9.
- Mercadante S, Maddaloni S, Roccella S, Salvaggio L. Predictive factors in advanced cancer pain treated only by analgesics. *Pain*. 1992;50:151–5.
- Haugen DF, Hjermstad MJ, Hagen N, Caraceni A, Kaasa S, European Palliative Care Research Collaborative (EPCRC). Assessment and classification of cancer breakthrough pain: a systematic literature review. *Pain*. 2010;149:476–82.
- Greenfield K, Holley S, Schoth DE, Bayliss J, Anderson A-K, Jassal S, 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.
- Scholz J, Finnerup NB, Attal N, Aziz Q, Baron R, Bennett MI, et al. The IASP classification of chronic pain for ICD-11. *Pain*. 2019;160:53–9.
- Davies AN, Dickman A, Reid C, Stevens AM, Zeppetella G. 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(4):331–8.
- Webber K, Davies AN, Cowie MR. Accuracy of a diagnostic algorithm to diagnose breakthrough cancer pain as compared with clinical assessment. *J Pain Symptom Manag*. 2015;50:495–500.
- Mercadante S, Lazzari M, Reale C, et al. Italian Oncological Pain Survey (IOPS). *Clin J Pain*. 2015;31:214–21.
- Løhre 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 Manag*. 2016;51:1013–9.
- World Health Organization. WHO guidelines for the pharmacological and radiotherapeutic management of cancer pain in adults and adolescents. Geneva: WHO; 2018. Accessed 31 Mar 2021.
- Daeninck P, Gagnon B, Gallagher R, et al. Canadian Recommendations for the Management of Breakthrough Cancer Pain. *Curr Oncol*. 2016;23:96–108.

15. Simmonds MA. Management of breakthrough pain due to cancer. *Oncology (Williston Park)*. 1999;13: 1103–8 (discussion 1110, 1113–4).
16. 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.
17. Mercadante S, Marchetti P, Cuomo A, Mammucari M, Caraceni A. Breakthrough pain and its treatment: critical review and recommendations of IOPS (Italian Oncologic Pain Survey) expert group. *Support Care Cancer*. 2016;24:961–8.
18. Samolsky Dekel BG, Remondini F, Gori A, Vasarri A, Di Nino G, Melotti RM. 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.
19. Samolsky Dekel B, Palma M, Sorella MC, Gori A, Vasarri A, Melotti RM. Development and performance of a diagnostic/prognostic scoring system for breakthrough pain. *J Pain Res*. 2017;10:1327–35.
20. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN manual for systematic reviews of PROMs, user manual. 2018. [https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual\\_version-1\\_feb-2018.pdf](https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf). Accessed 29 Aug 2020.
21. Lawshe CH. A quantitative approach to content validity. *Pers Psychol*. 1975;28:563–75.
22. Jensen MP. Questionnaire validation: a brief guide for readers of the research literature. *Clin J Pain*. 2003;19:345–52.
23. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
24. Thorndike RM. Book review: psychometric theory (3rd ed.) by Jum Nunnally and Ira Bernstein New York: McGraw-Hill, 1994, xxiv + 752 pp. *Appl Psychol Meas*. 1995;19:303–5.
25. Myles P, Gin T. Predicting outcome: diagnostic tests or predictive equations. In: Myles P, Gin T, editors. *Statistical methods for anaesthesia and intensive care*. Oxford: Reed Educational and Professional Publishing Ltd; 2000. p. 94–104.
26. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;3(7): 737–45.
27. Hess AS, Shardell M, Johnson JK, et al. Methods and recommendations for evaluating and reporting a new diagnostic test. *Eur J Clin Microbiol Infect Dis*. 2012;31:2111–6.
28. Myles P, Gin T. 8—Predicting outcome—diagnostic tests or predictive equations. In: Myles P, Gin T, editors. *Statistical methods for anaesthesia and intensive care*. Oxford: Reed Educational and Professional Publishing Ltd; 2008. p. 1–11.
29. Webber K, Davies AN, Zeppetella G, Cowie MR. Development and validation of the breakthrough pain assessment tool (BAT) in cancer patients. *J Pain Symptom Manag*. 2014;48:619–31.
30. Mercadante S. The use of rapid onset opioids for breakthrough cancer pain: the challenge of its dosing. *Crit Rev Oncol Hematol*. 2011;80:460–5.
31. U.S. Food & Drug Administration. Transmucosal immediate-release fentanyl (TIRF) medicines; 2020. <https://www.fda.gov/drugs/information-drug-class/transmucosal-immediate-release-fentanyl-tirf-medicines>. Accessed 31 Mar 2021.
32. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ*. 2013;346: e5595.
33. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *J Pain*. 2014;15:569–85.
34. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94:149–58.
35. de Yébenes Prous MJ, Rodríguez Salvanés FCOL. Responsiveness of outcome measures. *Rheumatol Clin*. 2008;4:240–7.