Università Campus Bio-Medico di Roma

Corso di dottorato di ricerca in Scienze Biomediche Integrate e Bioetica

XXXII ciclo a.a. 2016-2017

A machine learning approach to identify clusters of patients with different Breakthrough cancer Pain (BTcP) clinical features and specific opioids response

Grazia Armento

Coordinator

Prof. Paolo Pozzilli

Tutors
Prof. Giuseppe Tonini
Dott. Francesco Pantano

9th July 2020

To my mother,

invincible warrior

"To go fast, go alone,

to go far, go together".

-African proverb-

Table of contents

Abstract

Introduction: A large proportion of patients with cancer suffer from Breakthrough cancer pain (BTcP). Several unmet clinical needs concerning BTcP treatment, like optimal opioids dosage, are being investigated. In this analysis the hypothesis whether distinct subtypes of BTcP exist and whether they can provide new insights into clinical practice is explored with an unsupervised learning algorithm.

Methods: It was used partitioning around medoids algorithm on a large dataset of patients with BTcP previously collected by the Italian Oncologic Pain Survey (IOPS) group in order to identify possible subgroups of BTcP; the input of the algorithm consisted of different BTcP features, like its duration or its intensity. Silhouette statistics was used to pick an optimal number of clusters. Resulting clusters were analyzed in terms of BTcP therapy satisfaction, clinical features and usage of basal pain and rapid onset opioids. Opioids dosages were converted to a unique scale and BTcP-opioids-to-basal-pain-opioids ratio (OpR) was calculated for each patient. Polynomial logistic regression was used to catch non-linear relationships between therapy satisfaction and opioids usage.

Results: The cohort comprised 4016 patients with controlled basal pain and suffering from BTcP. Our algorithm identified 12 distinct BTcP clusters. Optimal OpRs differed across the clusters, ranging from 15% to 50%. In the whole cohort, OpR was more clearly associated with therapy satisfaction compared with BTcP opioids or basal pain opioids alone. The majority of the clusters were linked to peculiar association of certain drugs with therapy satisfaction or dissatisfaction. A free online tool was created for new patients cluster computation (https://mancapaolo.shinyapps.io/UCBM_BTcPclusters/) in order to validate these clusters in future studies and to provide a possible, handy indications for personalized BTcP therapy.

Discussion: This work proposes a classification for BTcP and identifies subgroups of patients with unique efficacy of different pain medications. This work supports the theory that the optimal dose of BTcP opioids depends on the dose of basal opioids and identifies novel values, possibly useful for future trials. These results will allow to target BTcP therapy based on patient characteristics and to define a "precision medicine" strategy also for supportive care.

Introduction

Breakthrough cancer pain (BTcP) is a common event that affects a considerable proportion of cancer patients[1]. A variety of definitions for BTcP have been proposed[2] [3]: according to the Italian Oncologic Pain Survey (IOPS) study group[4], BTcP should be defined as "as a relevant change in pain intensity of severe intensity in patients who receive an effective treatment with opioids". Nevertheless, despite this unique definition, BTcP encloses a wide range of manifestations which differ, among other features, for intensity, duration, frequency and triggering events. BTcP represents a clinically relevant condition with a negative impact on the patient's quality of life. In the majority is difficult to achieved an acceptable degree of relief because cancer patients have complex pain syndromes. These patients often require more intense therapeutic protocols, and therefore, more time may be required to achieve adequate pain control[5].

At present, several gaps exist in the knowledge of BTcP management. These partially unanswered questions concern, among others, the optimal drugs administration route, the pharmacokinetics, the balance between rapid onset and slow onset opioids and the eventual difference of BTcP response deriving from clinical features like stage, primary site or metastases.

In this analysis it was thus hypothesized that the unique BTcP definition might actually enclose diverse pathological entities, possibly with peculiar clinical needs and specific response to drugs. In order to explore this supposition, we used novel multiparametric artificial intelligence algorithms which can simultaneously analyse

[1] 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(3):241–249.

[2] Fallon M, Giusti R, Aielli F, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. Ann Oncol. 2018;29(Suppl 4):iv166–iv191

[3] Boceta J, De la Torre A, Samper D, Farto M, Sánchez-de la Rosa R. Consensus and controversies in the definition, assessment, treatment and monitoring of BTcP: results of a Delphi study. Clin Transl Oncol. 2016;18(11):1088–1097

[4] Mercadante S, Marchetti P, Cuomo A, Mammucari M, Caraceni A; IOPS MS study Group. Breakthrough pain and its treatment: critical review and recommendations of IOPS (Italian Oncologic Pain Survey) expert group [published correction appears in Support Care Cancer. 2017 Aug;25(8):2673-2674]. Support Care Cancer. 2016;24(2):961–968

[5] Mercadante S, Marchetti P, Cuomo A, et al. Factors Influencing the Clinical Presentation of Breakthrough Pain in Cancer Patients. Cancers (Basel). 2018;10(6):175

different clinical features and identify the existence of shared patterns. These so-called unsupervised learning algorithms have already been extensively used, for example, for the identification of breast cancer subtypes[6]. Nevertheless, to our knowledge, no authors yet tried to explore the issue of BTcP management with these techniques.

In order to fulfill this purpose, the data collected by the IOPS group in a large, multicentric national study[7] [8] [9] that enrolled 4056 patients from thirty-two centres with opioids-controlled basal pain suffering from BTcP were used. Hence, this work is a secondary analysis of the IOPS group survey that aims to identify novel subtypes of BTcP through the use of unsupervised learning algorithms.

[6] Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–10874

[7] Mercadante S, Marchetti P, Cuomo A, et al. Factors Influencing the Clinical Presentation of Breakthrough Pain in Cancer Patients. Cancers (Basel). 2018;10(6):175

[8] Mercadante S, Marchetti P, Cuomo A, et al. Breakthrough Cancer Pain: Preliminary Data of The Italian Oncologic Pain Multisetting Multicentric Survey (IOPS-MS). Adv Ther. 2017;34(1):120–135

[9] Mercadante S, Lazzari M, Reale C, et al. Italian Oncological Pain Survey (IOPS): a multicentre Italian study of breakthrough pain performed in different settings. Clin J Pain. 2015;31(3):214–221

Methods

Patients enrollment and data collection

Details concerning the enrollment of patients are extensively described in the main paper from the IOPS group[10]. Briefly, the local ethical committees approved the protocol, and written informed consent was obtained from each patient. Interviews were performed in different settings, in particular oncology, pain therapy, palliative care, and radiotherapy. Patients were ≥ 18 years old, diagnosed with cancer at any stage, stable background pain in the last week with an intensity of at most 4 on a numerical scale from 0 to 10, and episodes of BTcP with an intensity of 5 or more, clearly distinguished from background pain. The definition of BTP was: a transitory pain exacerbation of moderate to severe intensity that occurs spontaneously or predictably well distinguished from background pain (as shown in the algorithm of Figure 1). Exclusion criteria were the absence of a cancer diagnosis, uncontrolled background pain (>4 on a numerical scale of 0 to 10), or no relevant increases in pain intensity (<5) which could be interpreted as BTcP episodes. Patients unable to provide information about the data required for the study, as a result of either cognitive failure or terminal disease, were also excluded. A comprehensive list of clinical variables were collected for each patient, comprising basal pain and BTcP site, duration, frequency, intensity, relieving factors, triggers, drugs, primary cancer site and stage, concomitant symptoms, for a total of 1086 variables. Interviews were registered by collecting personnel in a closed online form and centrally stored.

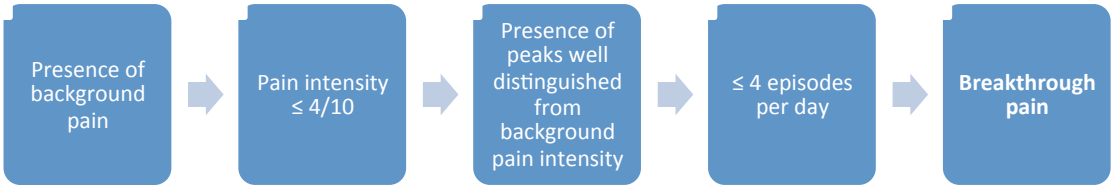


Figure 1: algorithm used for the diagnosis of BTcP during the patients' enrollment in the IOPS survey (Modified from Mercadante S, Lazzari M, Reale C et al. Clin J Pain 2015)

[10] Mercadante S, Marchetti P, Cuomo A, et al. Factors Influencing the Clinical Presentation of Breakthrough Pain in Cancer Patients. Cancers (Basel). 2018;10(6):175

Therapy satisfaction

The association of each clinical feature with satisfaction toward BTcP therapy was investigated through a simple logistic regression; therapy satisfaction was expressed as a binomial outcome; false discovery rate (FDR) method[11] was used to correct p values for multiple comparisons; features with less than 5% of missing data and associated with a corrected p value <0.1 and - for categorical features - a $|log_2(OR)|$ greater than 1 were used to build a multivariate logistic regression. In order to investigate - simultaneously for all patients and on the same scale - the correlation between the amount of opioids used and BTcP therapy satisfaction, all the doses of opioid drugs were converted[12] to the equivalent intravenous morphine (ivM) doses and expressed as a total daily dose - one for BTcP-directed opioids and for basal pain opioids. The conversion was performed to intravenous morphine and not to oral morphine because intravenous morphine has been increasingly used in different clinical situations and to get graphically more interpretable graphics in the results section. Moreover, in order to explore the interaction of fast-acting and long-acting opioids dosages, we calculated for each patient the BTcP-opioids-to-basal-pain-opioids ratio (OpR). A polynomial logistic regression was used to catch non-linear relationships between opioids doses and therapy satisfaction.

Cluster computation and visualization

Features defining clinical characteristics of BTcP were selected to perform clusters computation. The features with missing data accounting for more than 5% of the patients were excluded. Above-mentioned features were used to calculate a dissimilarity matrix using cluster package[13]; since features comprised also non-numeric variables, gower metric[14] was used for the dissimilarity matrix calculation. Partitioning around medoids[15] algorithm was used to compute clusters using the dissimilarity matrix as an input; the algorithm was run different times using a range of

[11] Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B. 1995; 57, 289–300. http://www.jstor.org/stable/2346101.

[12] https://www.nccn.org/professionals/physician_gls/pdf/pain.pdf

[13] Maechler M, Rousseeuw, P, Struyf, A, Hubert, M, Hornik, K. cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0. 2019

[14] Gower JC. A General Coefficient of Similarity and Some of Its Properties. Biometrics. 1971; Vol. 27, No. 4. 857-871 Published by: International Biometric Society

[15] Kaufman L, Rousseeuw PJ. Clustering by means of Medoids, in Statistical Data Analysis Based on the L1–Norm and Related Methods. Edited by Y. Dodge, North-Holland, 1987; 405–416.

clusters number spanning from 2 to 30; silhouette statistics[16] was calculated for each run; the optimal number of clusters was manually picked as being the one with the best trade-off between silhouette statistics and reasonable clinical interpretation. t-Distributed Stochastic Neighbor Embedding (tSNE)[17] algorithm was used to project dissimilarities between patients in a bidimensional space, with closer points representing patients with more similar clinical BTcP features. An online tool allows to repeat the performed classification on new sets of patients (https://mancapaolo.shinyapps.io/UCBM_BTcPclusters/).

Clusters analysis

T-test, mann-whitney and χ^2 tests were used, respectively, to assess the association of parametric, non-parametric and categorical features with each cluster. FDR method was used to correct p values for multiple comparisons. Therapy satisfaction was investigated separately for each cluster as previously described for all samples.

Data handling

Data was imported and analysed in R (v 3.5.2)[18]. Packages used for the analyses are dplyr[19], cluster[20], Rtsne[21], ggplot2[22], gmodels[23], Rmisc[24], epiR[25], mgcv[26], knitr[27], RColorBrewer[28], gmodels[29], mgcViz[30].

[16] Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 1987; 20, 53–65.
[17] Van der Maaten LJP, Hinton GE. Visualizing High-Dimensional Data Using t-SNE. 2008; Journal of Machine Learning Research, 9, pp.2579-2605.
[18] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019; URL https://www.R-project.org/.
[19] Wickham H, François R, Henry L and Müller K. dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. 2019.
[20] Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik, K. cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0. 2019
[21] Krijthe JH. Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation. URL: https://github.com/jkrijthe/Rtsne. 2015.
[22] Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.
[23] Warnes GR, Bolker B, Lumley T, Johnson RC. Contributions from Randall C. Johnson are Copyright SAIC-Frederick, Inc. Funded by the Intramural Research Program, of the NIH, National Cancer Institute and Center for Cancer Research under NCI Contract NO1-CO-12400. 2018; gmodels: Various R Programming Tools for Model Fitting. R package version 2.18.1.
[24] Hope RM. Rmisc: Rmisc: Ryan Miscellaneous. R package version 1.5. 2013; https://CRAN.R-project.org/package=Rmisc
[25] Stevenson M, Heuer C, Marshall J, Sanchez J, Thornton R, Reiczigel J et al. epiR: Tools for the Analysis of Epidemiological Data. 2018; R package version 0.9-99.
[26] Wood SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association. 2004; 99:673-686.
[27] Yihui Xie . knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.21. 2018
[28] Neuwirth E. RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. 2014. https://CRAN.R-project.org/package=RColorBrewer

[29] Warnes GR, Bolker B, Lumley T, Johnson RC. Contributions from Randall C. Johnson are Copyright SAIC-Frederick, Inc. Funded by the Intramural Research Program, of the NIH, National Cancer Institute and Center for Cancer Research under NCI Contract NO1-CO-12400. 2018; gmodels: Various R Programming Tools for Model Fitting. R package version 2.18.1.

[30] Fasiolo M, Nedellec R, Goude Y, Wood SN. Scalable visualisation methods for modern Generalized Additive Models. ArXiv preprint arXiv:1809.10632. 2018.

Results

Patients Characteristics

A total of 4016 patients were enrolled in the study during a period of 24 months. Men accounted for 54.8% of the total; mean age was 64.6 (range 18-97). The majority of visits were performed for oncologic (52.0%) and pain therapy (29.5%) purposes. Together, inpatients (37.3%) and outpatients (34.3%) settings accounted for more than half of the visits. Most common cancer primary organs were lung (24.0%), breast (11.3%), pancreas (8.3%) and colon 7.5%. The mean Karnofksy PS was 48. According to inclusion criteria, basal pain was generally controlled - mean basal Pain NRS was 3.0. Patients' characteristics are described in Table 1.

Feature	N (total=4016)
Men	2202 (54.8%)
Oncology	2087 (52%)
Pain therapy	1184 (29.5%)
Palliative care	720 (17.9%)
Radiotherapy	25 (0.6%)
Inpatients	1498 (37.3%)
Outpatients	1378 (34.3%)
Domicile	577 (14.4%)
Day hospital	462 (11.5%)
Hospice	101 (2.5%)
Lung	963 (24%)
Breast	453 (11.3%)
Pancreas	335 (8.3%)
Colon	301 (7.5%)
Prostate	197 (4.9%)
Rectum	143 (3.6%)
Stomach	141 (3.5%)
Bladder	102 (2.5%)
Multiple primary	124 (3.1%)

Age	64.6 (range 18-97)
Karnofsky	48 (range 1-100)
Baseline Pain (NRS scale)	3 (range 0-10)

Table 1: Patients' characteristics.

Cluster Computation

In order to investigate whether "subtypes" of BTcP exist, we used BTcP features to build an unsupervised clustering model. The number of BTcP episodes, the BTcP peaks duration, the BTcP type, the BTcP intensity, the number of days since the begun of BTcP episodes, the eventual benefit from pharmacotherapy, the eventual benefit from rest and whether the BTcP was enhanced by movements were the 8 BTcP-defining variables selected for the final model, which was built with k-medoids algorithm. We chose 12 as an optimal trade-off between the average width of clusters silhouette (0.45) (Suppl. Figure 1) and the interpretability of clusters themselves. The average internal dissimilarity was acceptably low, ranging between 0.05 to 0.16. Figure 2 shows a 2-dimensional tSNE projection of all patients, coloured by their clusters. An online tool is available for the classification of new patients according to our method:

(https://mancapaolo.shinyapps.io/UCBM_BTcPclusters/).

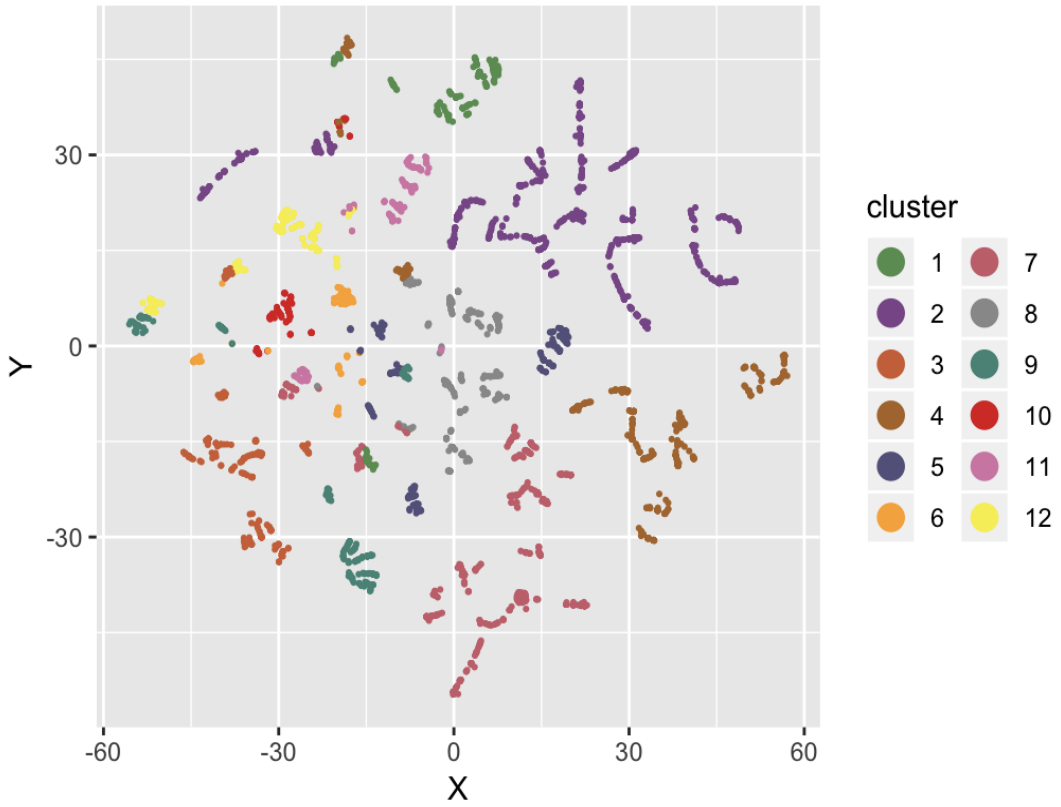


Figure 2: a 2-dimensional tSNE projection of all patients, coloured by their clusters, based on BTcP features (number of BTcP episodes, the BTcP peaks duration, the BTcP type, the BTcP intensity, the number of days since the begun of BTcP episodes, the eventual benefit from pharmacotherapy, the eventual benefit from rest and whether the BTcP was enhanced by movements). Each point represents a patient; patients dissimilarity in BTcP clinical features is represented by the points distance. Colors represent 12 cluster computed through partitioning around medoids (k-medoids) algorithm. 12 was chosen as an optimal trade-off between the average width of clusters silhouette (0.45) (look at Supplementary figures Figure 1) and the interpretability of clusters themselves.

Characteristics of BTcP clusters

It thus was analysed the enrichment of the 8 BTcP-defining variables and of other clinical features among the clusters.

Cluster 11 and 12 BTcP is more frequently a high-intensity BTcP that reaches NRS scale values ≥ 9 (ORs: 2.56 and 2.47 for cluster 11 and 12 vs other clusters, CIs: 1.83 – 3.57 and 1.79 – 3.40; Pearson χ^2: $p=9.0x10^{-10}$ and $p=4.1x10^{-8}$). The BTcP of these two clusters is often likely to respond to pharmacotherapy but not from rest. It tends to be enhanced by movements and peaks tend to have a shorter duration. Cluster 12 BTcP episodes tend to last longer than other clusters, retaining with these a median difference of 10 days more (CIs: 0.00 to 25.00, Mann Whitney: p=0.02). Expectedly,

the BTcP of these patients is more likely to heavily interfere with normal activities of daily living (ORs: 1.73 and 2.01 for cluster 11 and 12, CIs: 1.27 - 2.37 and 1.49 – 2.71; Pearson χ^2: p=0.004 and p=7.46x10^{-5}), with patients from both clusters retaining a median difference of Karnofsky performance status (PS) of 1.00 points higher compared to the rest of the patients (CIs: 0.00 - 1.00 for both clusters, Mann Whitney: p=3.5x10^{-4} for cluster 11 and 2.5 x10^{-5} for cluster 12). These two patterns of pain are more frequent in patients with lung cancer (ORs: 1.79 for cluster 11 and 1.82 for cluster 12, CIs: 1.30 – 2.46 and 1.34 – 2.48; Pearson χ^2: p=0.005 and p=0.037), with kidney cancer (OR: 2.64, CIs: 1.39 – 5.04 for cluster 12; Pearson χ^2: p=0.037), with bone metastasis (OR: 1.90, CIs: 1.39 – 2.58 for cluster 11, Pearson χ^2: p=0.001) and can be found more frequently in typically end-stage settings like palliative care visits (OR: 1.61, CIs: 1.14 – 2.27 for cluster 12; Pearson χ^2: p=0.012) and less frequently in normal outpatient regimen visits (OR: 0.42, CIs: 0.29 – 0.63 for cluster 11; Pearson χ^2: p=0.002).

Cluster 8 BTcP is the unique cluster which doesn't benefit neither from pharmacotherapy nor from rest. The BTcP of these patients is mostly unrelated to movements and the BTcP peaks tend to last less than 10 minutes. BTcP episodes of these patients is less long-lasting than the others, being a median of 5 days shorter (CIs: 0.00 to 10.00, Mann Whitney: p=0.04). Patients in this cluster have a better PS, with a median difference of 1.00 point (CIs: 1.00 to 2.00 points, Mann Whitney: p=1.2x10^{-30}). Moreover, they are more often unsatisfied of their BTcP pharmacotherapy (OR: 1.46. CIs: 1.11 – 1.93, χ^2: p=3.0x10^{-4}). A typical BTcP localization of these patients is the thorax (32.2% vs 18.8%, χ^2: p=9.4x10^{-9}). This pattern of pain is more frequent in patients with rectal cancer (OR: 2.14, CIs: 1.69 – 2.71; Pearson χ^2: p=0.01), without metastases (OR: 0.61, CIs: 0.47 - 0.79 ;χ^2: p=0.01) and particularly without lung metastases (OR: 0.54, CIs: 0.41 – 0.72; Pearson χ^2: p=6.3x10^{-4}). Expectedly, they can be found more often in the setting of pain therapy evaluation (OR: 3.71. CIs: 2.97 - 4.63; Pearson χ^2: p=4.7x10^{-30}) or in early stages setting like outpatient regimen rather than hospice or inpatient regimen (OR: 3.62 for outpatients vs non-outpatients, CIs: 2.89 - 4.53; χ^2: p=5.7x10^{-27}).

Cluster 3 BTcP is more rarely a high-intensity (≥ 9) BTcP (OR: 0.59, CIs: 0.47 - 0.75; Pearson χ^2: p=1.1x10^{-4}), it is nociceptive, it has a median difference of duration 12 minutes higher than other cluster (CIs: 5.00 to 20.00, Mann Whitney: p=0.001) and is

more likely to be a single episode (OR: 1.84, CIs: 1.45 - 2.33; Pearson χ^2: $p=6.0x10^{-7}$) with peaks longer than 10 minutes. These patients derive a benefit from pharmacotherapy and are more frequent in the home care setting (OR: 1.77, CIs: 1.33 - 2.35; Pearson χ^2: $p=3.5x10^{-5}$) and in visit performed for palliative care purposes (OR: 1.70 CIs: 1.30 - 2.21; Pearson χ^2: p=0.006). These patients are less likely to be unsatisfied of their BTcP pharmacotherapy (OR: 0.48; CIs: 0.32 - 0.73; Pearson χ^2: p=0.034) and to derive from BTcP a heavy impairment of their normal life (OR: 0.56, CIs: 0.42 - 0.75; Pearson χ^2: $p=2.6x10^{-7}$).

Patients in cluster 4 are often patients with gastric (OR: 2.01, CIs: 1.34 - 3.02; Pearson χ^2: p<0.001) or colon cancer (OR: 1.65, CIs: 1.21 - 2.24; Pearson χ^2: p<0.001). They suffer from a nociceptive pain which derives a benefit from pharmacotherapy. Expectedly, these patients have more often liver (OR: 1.71, CIs: 1.40 - 2.08; Pearson χ^2: p<0.001) and peritoneal metastases (OR: 2.36, CIs: 1.80 - 3.08; Pearson χ^2: p<0.001) and more rarely bone metastases (OR 0.45:, CIs: 0.36 - 0.55; Pearson χ^2: p<0.001). Their BTcP pain site is mostly the abdomen (OR: 2.33, CIs: 1.92 - 2.82; Pearson χ^2: p<0.001) and its triggered by eating. Intriguingly, in these patient the BTcP is not a driver of heavy life impairment (OR:0.66, CIs: 0.53 - 0.83; Pearson χ^2: p<0.001) and they are clearly very satisfied of their BTcP therapy (OR: 0.32, CIs: 0.22 - 0.47; Pearson χ^2: p<0.01). They share with cluster 3 the relatively high frequency of abdominal BTcP.

Cluster 7 BTcP is similar to cluster 3 BTcP as it is more rarely a high intensity one (OR: 0.46, CIs: 0.35 - 0.60; Pearson χ^2: $p=9.4x10^{-9}$), with single episodes (OR: 1.53, CIs: 1.27 – 1.84; Pearson χ^2: $p=8.0x10^{-6}$) with a longer overall duration compared to other patients (mean difference: 5.00 minutes, CIs: 0.00 – 10.00 minutes; Mann Whitney: p=0.042) and whose peaks last longer than 10 minutes. The sharp difference of this cluster from the third one is that they are much more likely to be unsatisfied of their pharmacotherapy (OR: 2.13, CIs: 1.73 - 2.62; $p=2.8x10^{-13}$). In fact, similarly, to unsatisfied patients from cluster 8, they are most frequent in the setting of pain therapy evaluation (OR: 2.08, CIs: 1.74 - 2.48; Pearson χ^2: $p=8.7x10^{-14}$) performed within normal outpatient visits (OR: 2.41, CIs: 2.02 - 2.87, Pearson χ^2: $p=1.8x10^{-19}$).

Patients of cluster 9 are particularly often patients with bone metastases (OR=2.59, CIs: 1.96 - 3.41; Pearson χ^2: $p=3.5x10^{-10}$) and/or with breast cancer (OR: 1.72, CIs: 1.22 - 2.43; Pearson χ^2: p=0.022). Their BTcP is typically enhanced by movements but

responsive to both rest and pharmacotherapy. Expectedly, these patients often experience BTcP in the lumbar (OR: 2.08, CIs: 1.53 - 2.73 Pearson χ^2: p<0.001;) and dorsal (OR: 1.73, CIs: 1.28 - 2.35 Pearson χ^2: p<0.001;) spine, hips (OR: 2.30, CIs: 1.66 - 3.20 Pearson χ^2: p<0.001;) and inferior limbs (OR: 2.30, CIs: 1.71 - 3.10 Pearson χ^2: p<0.001;).

Also patients in cluster 6 have a BTcP triggered by movements which benefits from rest and they are, more often, old (Mann Whitney: p<0.001) females (OR: 0.59, CIs: 0.42 - 0.84; Pearson χ^2: p=0.003) with bone metastases (OR: 2.12, CIs: 1.49 - 3.02; Pearson χ^2: p<0.001); oppositely to patients in cluster 9, their BTcP doesn't derive any benefit from pharmacotherapy. These patients have a higher PS (Mann Whitney: p=0.002) and are more frequent in end stage settings like home care visits (OR: 4.00, CIs: 2.80 - 5.73; Pearson χ^2: p<0.001) performed for palliative care purposes (OR: 3.90, CIs: 2.75 - 5.54; Pearson χ^2: p<0.001).

Patients in cluster 5 are often patients with breast cancer (OR: 1.80, CIs: 1.28 - 2.51; Pearson χ^2: p<0.001) seen in the context of early stage settings like outpatient visits (OR: 1.44, CIs: 1.11 - 1.88; Pearson χ^2: p=0.006) or day hospital (OR: 1.72, CIs: 1.22 - 2.43; Pearson χ^2: p=0.002) and for oncologic purposes rather than for palliative care or pain therapy (OR: 2.25, CIs: 1.70 - 2.99; Pearson χ^2: p<0.001); coherently, this group is characterized by an average lower PS. Their BTcP does derive a benefit from rest but not from pharmacotherapy. Their typical site of BTcP is the cervical spine (OR: 2.26, CIs: 1.42-3.61; Pearson χ^2: p<0.001) and it's shorter compared to rest of the cohort. These patients are deeply unsatisfied of their pharmacotherapy (OR: 3.97, CIs: 2.96 - 5.32; Pearson χ^2: p<0.001) and, actually, very often they are not even receiving any BTcP pharmacotherapy (OR: 0.10, CIs: 0.08-0.13; Pearson χ^2: p<0.001). Conversely they often seem to derive from BTcP only small impairment of their normal life (OR: 1.52, CIs: 1.16-2.00; Pearson χ^2: p=0.002).

Patients in the cluster 2, similarly to patients in cluster 5, are mostly found in the day hospital setting (OR: 1.94, CIs: 1.54 - 2.39; Pearson χ^2: p<0.001), where they go to receive antitumoral treatments. Also these patients do not derive any serious life impairment (OR: 0.54, CIs: 0.42-0.69; Pearson χ^2: p<0.001) from their BTcP but, differently from the fifth cluster, these patients are often satisfied of their BTcP therapy (OR: 0.65, CIs: 0.52-0.81; Pearson χ^2: p<0.001).

Patients in the cluster 10 are the ones who more often suffer from a neuropathic BTcP. They experience BTcP peaks longer than 10 minutes but they do derive a benefit from pharmacotherapy and they rarely experience high-intensity BTcP (OR: 0.57, CIs: 0.38 - 0.83; Pearson χ^2: p=0.003) . Patients who underwent therapy with taxanes are more likely to suffer from this type of BTcP (OR: 2.24, CIs: 1.34 - 3.76; Pearson χ^2: p=0.024). Patients with pleural mesothelioma seem particularly at risk of developing this kind of BTcP (OR=4.75. CIs: 1.83 - 12.31; Pearson χ^2: p=0.019)

Patients of cluster 1 are often patient who ever underwent any kind of radiotherapy (OR: 5.55, CIs: 2.19-14.03; Pearson χ^2: p<0.001) who are under active oncologic (OR: 1.57, CIs: 1.19-2.08; Pearson χ^2: p<0.001) treatment. The BTcP of these patients derives a benefit from both pharmacotherapy or rest but determines severe life impairment to the patients (OR: 1.87, CIs: 1.42-2.46; Pearson χ^2: p<0.001).

A brief description of each cluster is available in table 2. A summary of BTcP features according to the cluster is represented in Figure 3.

Brief description of clusters			
1	Patients undergone RT deriving a QOF impairment from BTcP	7	BTcP similar to C3; satisfaction to C8
2	Controlled BTcP	8	Early stage patients with uncontrolled BTcP
3	End stage patients with controlled BTcP	9	Patients with bone metastasis and controlled BTcP
4	Patients with GI cancers and excellent BTcP control	10	Patients with neuropathic BTcP
5	Early stage patients with untreated BTcP	11	End stage patients with severe BTcP
6	Old woman with bone metastasis and end stage disease with BTcP unresponsive to drugs	12	End stage patients with severe BTcP and bone mets

Table 2: brief description of clusters.

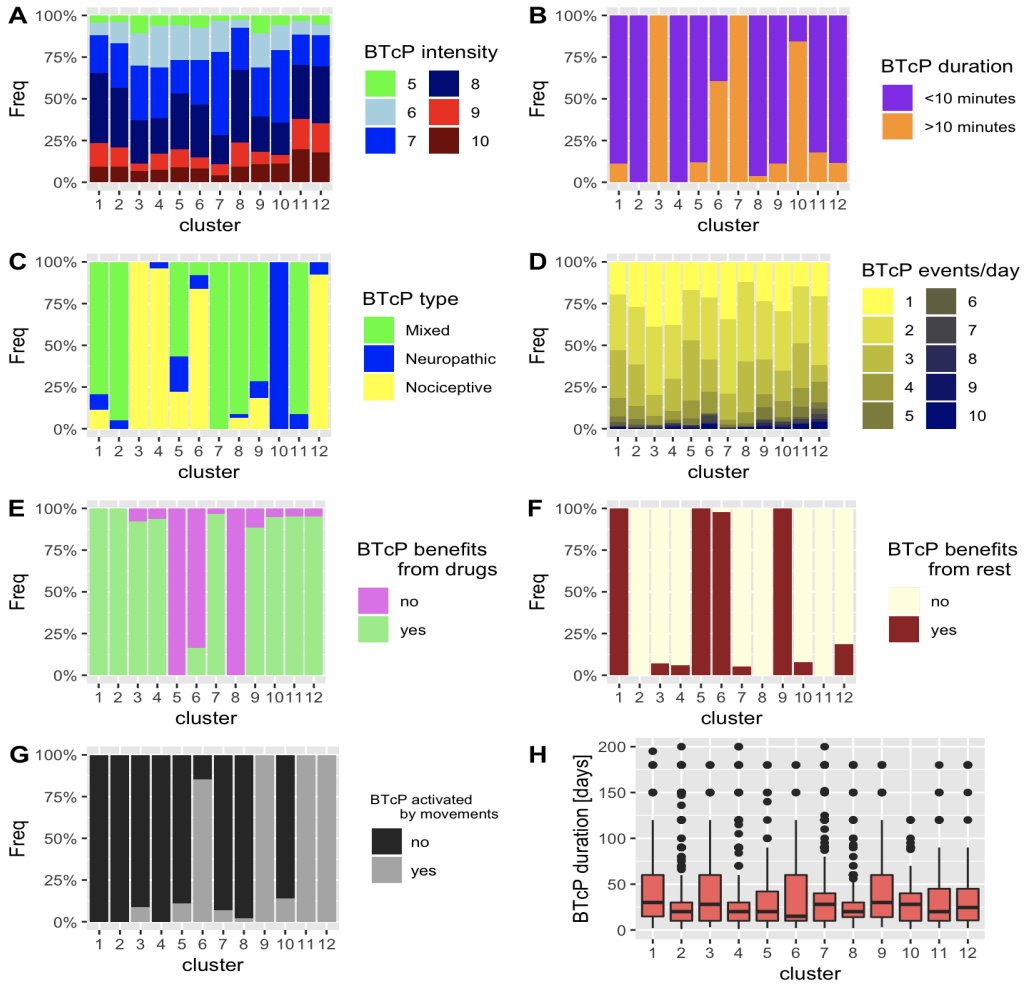


Figure 3: defining features of the 12 BTcP clusters. Plots from A to H represent, in order: BTcP intensity using NRS scale; BTcP peak duration; BTcP type; number of BTcP events per day; the presence of benefit in BTcP management with pharmacotherapy; the presence of benefit in BTcP management with rest; the presence of BTcP activation with movements and the days since BTcP episodes started.

BTcP Therapy satisfaction

Finally, we tried to assess what influenced the BTcP therapy satisfaction reported by the patients. After converting the opioids dose to a unique scale - corresponding to the equivalent dose of ivM - we investigated with a non-linear model the influence of basal opioids dose, BTcP opioids dose and OpR in patients reported therapy satisfaction. Very intriguingly, while basal opioids dose didn't show any big impact on therapy satisfaction and BTcP opioids dose showed some irregular peaks of satisfaction whose confidence intervals often reached the indifference line, the OpR seemed to depict a clear, optimal peak between 0.4 and 0.45 (Figure 4); this roughly corresponds to a daily dose of 100 mcg of sublingual fentanyl for BTcP and a daily dose of 30 mg of oral morphine of basal opioids dose.

We separately performed the same analyses on previously defined clusters (Figure 5). Of interest, not all the clusters showed the same relationship between OpR and satisfaction: for clusters 1, 6 and 10 the satisfaction seemed to grow indefinitely with the increase of the OpR opioids while clusters 7, 8 and 11 seemed to have clear, optimal peaks of OpR. Despite the interpretation being challenged by some large confidence intervals, we can say from this data that optimal OpR ranges - depending on the cluster - from 15% to 50%.

We then investigated which features are more associated with therapy satisfaction in the whole cohort. Among statistically significant features in the univariate regression (Figure 3), we highlight how the use of immediate release (IR) morphine for basal pain - which represents a wrong indication - is associated with higher therapy dissatisfaction and that, conversely, different fentanyl formulations are associated with a higher BTcP therapy satisfaction. Multivariate regression (supplementary results) confirmed, among others, the positive effect of fentanyl and that, although augmenting the dose of BTcP opioids is associated with a higher therapy satisfaction, doing the same for basal pain opioids doesn't provide the same beneficial effect.

After repeating univariate linear regressions separately for each cluster (fully available as supplementary tables) and adjusting for multiple comparisons, we could identify some peculiar associations. Among others, it is worth citing the positive association of therapy satisfaction with fentanyl buccal tablet (FBT) and fentanyl pectin nasal spray (FPNS) for cluster 2, basal pain acetaminophen for cluster 5, BTcP IR morphine, BTcP acetaminophen, FPNS for cluster 7, FBST and FPNS for cluster 8 and the negative association of therapy satisfaction with basal pain acetaminophen, OTFC, basal pain SC morphine, basal pain oxycodone/naloxone and basal pain tapentadol for cluster 2, BTcP acetaminophen for cluster 4, basal pain acetaminophen, basal pain oxycodone/naloxone and OTFC for cluster 7, BTcP acetaminophen and basal pain codein/acetaminophen for cluster 8, basal pain tapentadol for cluster 10, basal pain codein/acetaminophen for cluster 11 and BTcP acetaminophen for cluster 12 (ORs, confidence intervals and adjusted p values are available as supplementary results).

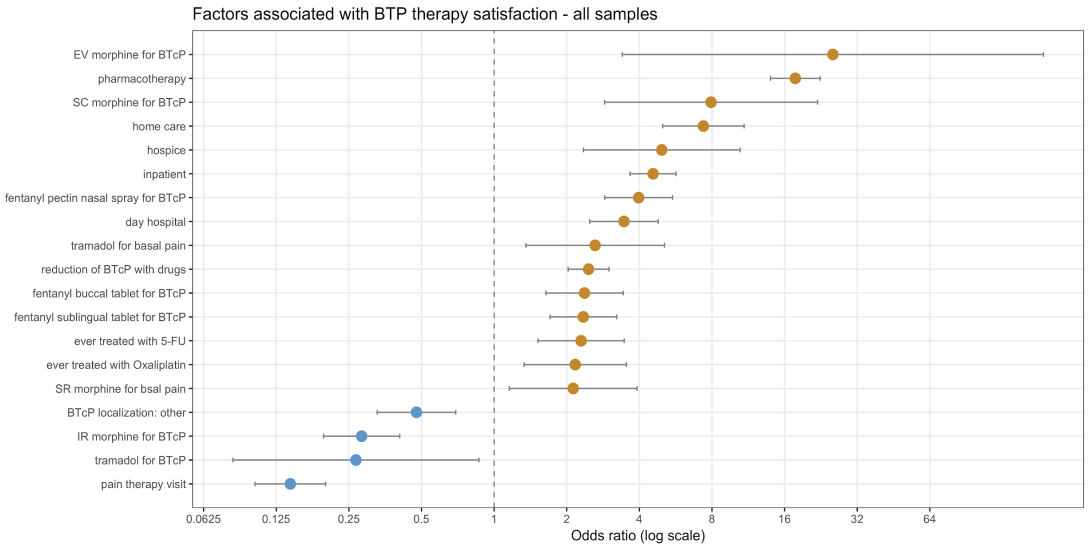


Figure 3: factors associated with BTcP therapy satisfaction. We excluded from this plot non-significant features (after p-value correction) and features with an OR comprised between 0.5 and 2. Orange and blue identify, respectively, features associated with therapy satisfaction and therapy dissatisfaction. Black lines depict 95% confidence intervals. Among statistically significant features in the univariate regression we highlight how the use of immediate release (IR) morphine for basal pain - which represents a wrong indication - is associated with higher therapy dissatisfaction and that, conversely, different fentanyl formulations are associated with a higher BTcP therapy satisfaction.

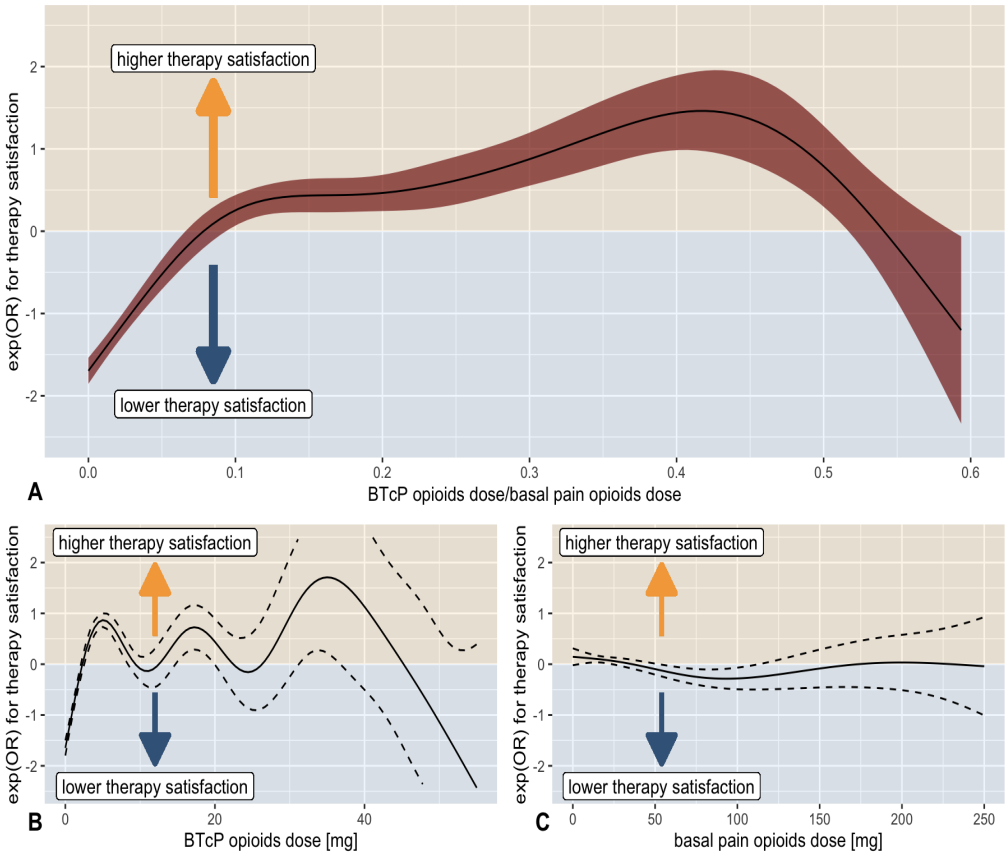


Figure 4: figures A, B and C show the correlation of BTcP therapy satisfaction with, respectively, the ratio of BTcP opioid drugs dose and basal pain opioid drugs dose (A), BTcP opioid drugs dose alone (B) and basal pain opioid drugs dose (C). The ratio has the clearest correlation with satisfaction, indicating an optimal range of opioids used for BTcP and opioids used in basal pain. While basal opioids dose didn't show any big impact on therapy satisfaction, and BTcP opioids dose showed some irregular peaks of satisfaction whose confidence intervals often reached the indifference line, the ratio has the clearest correlation with satisfaction, indicating an optimal range of dose ratio of opioids for BTcP and opioids used in basal pain: optimal peak between 0.4 and 0.45, which roughly corresponds to a daily dose of 100 mcg of sublingual fentanyl for BTcP and a daily dose of 30 mg of oral morphine. Solid lines represent logistic regressions calculated with more than 1 degree of freedom, dashed lines represent 95% confidence intervals.

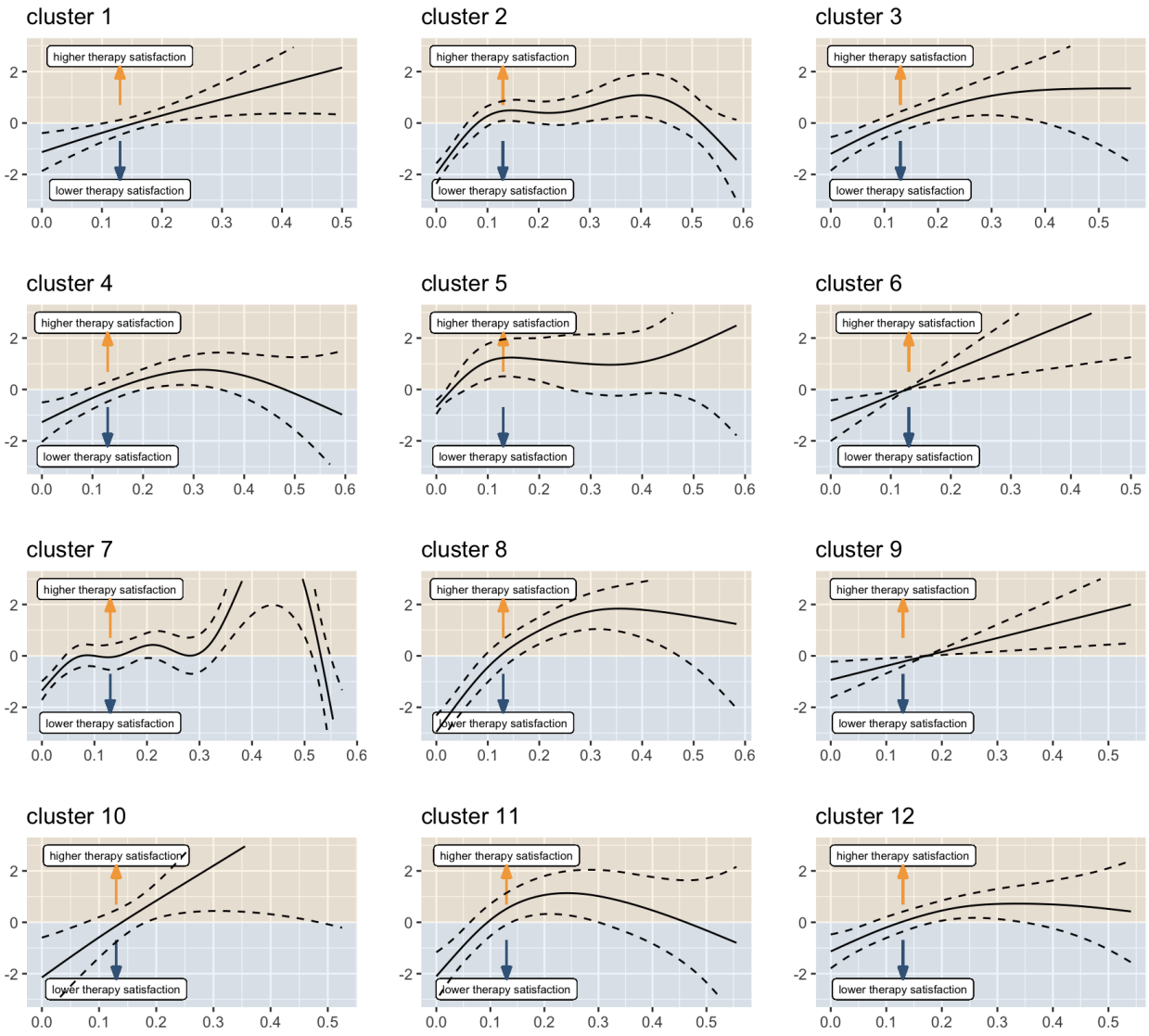


Figure 5: correlation between fast to basal opioids ratio (OpR) and therapy satisfaction for each cluster. Of interest, not all the clusters showed the same relationship between OpR and satisfaction: for clusters 1, 6 and 10 the satisfaction seemed to grow indefinitely with the increase of the OpR, while clusters 7, 8 and 11 seemed to have clear, optimal peaks of OpR in a definite range of it.

Discussion

This work shows a novel approach for the investigation of BTcP. According to these findings, 12 subtypes of BTcP with peculiar response to drugs and clinical presentation were identified. It was acknowledged that this study was not designed to perform this analysis and, moreover, the large number of clusters might interfere with their interpretability and clinical utility. Nevertheless this study represents a proof-of-concept for this investigational approach.

Some of these findings might already provide some indication for future clinical practice. First, it seems that an optimal ratio between BTcP opioids and basal pain opioids exists. Another group proposed, using our same data, 0.20 (one fifth) as the optimal ratio[31]; nevertheless they used a frequentistic approach, being 0.20 simply the most common ratio among the cohort. This analysis, instead, modelled the ratio toward an outcome (BTcP therapy satisfaction) and highlighted a peak of satisfaction within a ratio range of 0.40 - 0.45. What seems clear, though, is that such an optimal level exists: this possibly suggests that, other than starting BTcP opioids titration with the lowest possible dosage as proposed previously[32], titration could start immediately with an optimal opioids dosage. Moreover, cluster analysis reveals that this ratio might not be the same for all patients: some patients might benefit from a higher BTcP opioids dosage (cluster 2 and 7) while other from a lower one (cluster 11). Finally, for some patients we didn't observe an upper threshold for this ratio (cluster 1, 6 and 9), perhaps pointing out cases in which a strong BTcP opioids dosage increase is required.

Some potentially interesting clues, other than opioids dosage, derive from our clusters analysis. For example, patients with neuropathic BtcP tend to enrich in a cluster - the number 10 - which has a strong association of BTcP therapy dissatisfaction with the use of tapentadol for basal pain and with a higher rate of patients with mesothelioma; patients in cluster 7 have a long lasting, low intensity BTcP which causes deep dissatisfaction towards therapy and is associated with oral transmucosal fentanyl therapy failure; gastrointestinal cancers seem to enrich in a cluster - the number 4 - whose BTcP is relatively well controlled and located in the abdomen, possibly

[31] Mercadante S, Marchetti P, Cuomo A, et al. Factors Influencing the Clinical Presentation of Breakthrough Pain in Cancer Patients. Cancers (Basel). 2018;10(6):175

[32] Davies AN, 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; 13:331-8.

identifying a movement-unrelated incident but not predictable BTcP[33]. The interpretation is made very difficult by the multiple associations and, at this stage, is not mature to suggest any immediate change in clinical practice. However, we made available an online free tool (https://mancapaolo.shinyapps.io/UCBM_BTcPclusters/) which allows the classification of new patients according to our algorithm and returns a proposed BTcP therapy which depends on the patient cluster optimal OpR and his basal opioids dose. We suggest that this tool might be used in the future to prospectively validate the clinical importance of our clusters in the clinical practice and to compare our proposed opioids dosage in settings different from ours (Suppl. Figure 12 is an example taken from the free online tool interface).

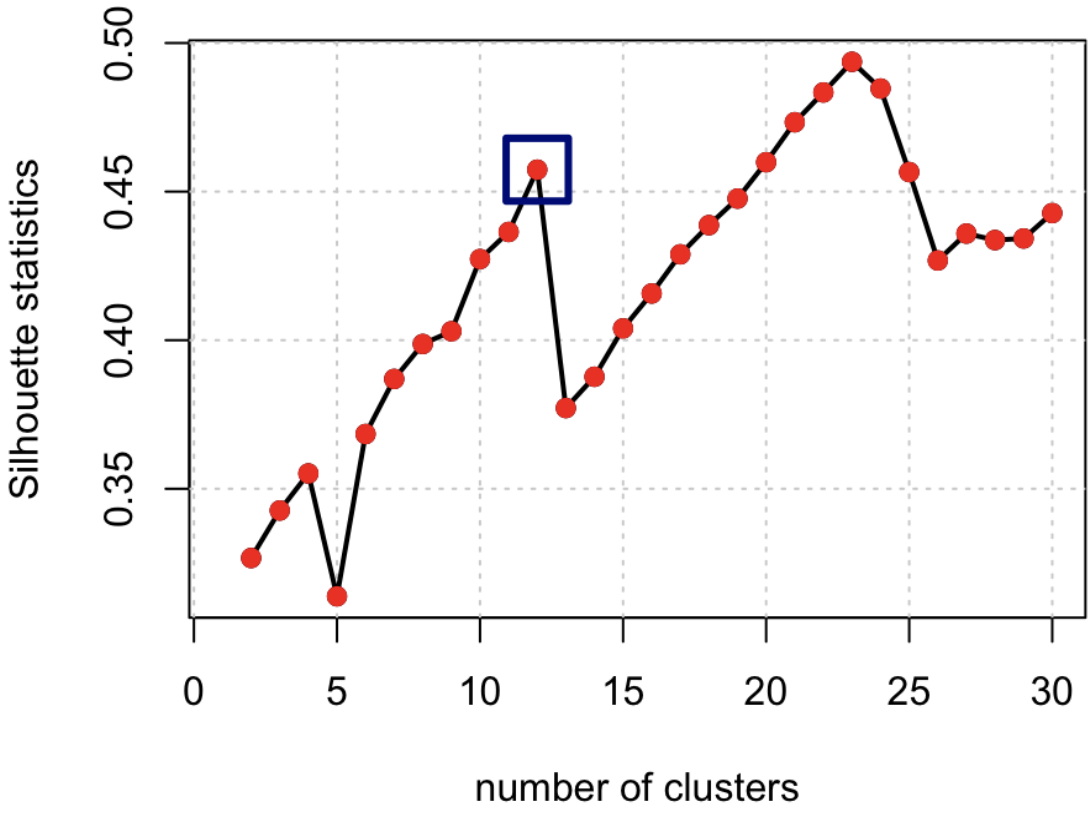
The presence of distinct BTcP phenotypes, each one associated with specific clinical features, could also be the reflection of diverse underlying pathophysiological mechanisms: our work suggests that preclinical research might gain insights on these possible differences and help the development of a tailored therapy also for BTcP.

The main limitation of this study is the appropriateness of collected data toward the scope of our work. We believe that a prospective study specifically designed for the investigation of BTcP clusters - possibly with long term follow-up and therapy success outcomes and not limited to a single timepoint evaluation - might enable a clearer identification of distinct clusters. Also, our approach lacks an external validation of cluster consistencies and reproducibility. Nevertheless all these limitations do not interfere with the main scope of this work, which was to offer a proof-of-concept for an innovative approach for BTcP management.

In conclusion, this work identifies criteria for optimal BTcP opioids therapy personalization and offers a reproducible classification for the enrollment and stratification of patients in future BTcP trials. This study lays the foundation for future trials in order to target BTcP therapy based on patient characteristics and to define a "precision medicine" strategy also for supportive care.

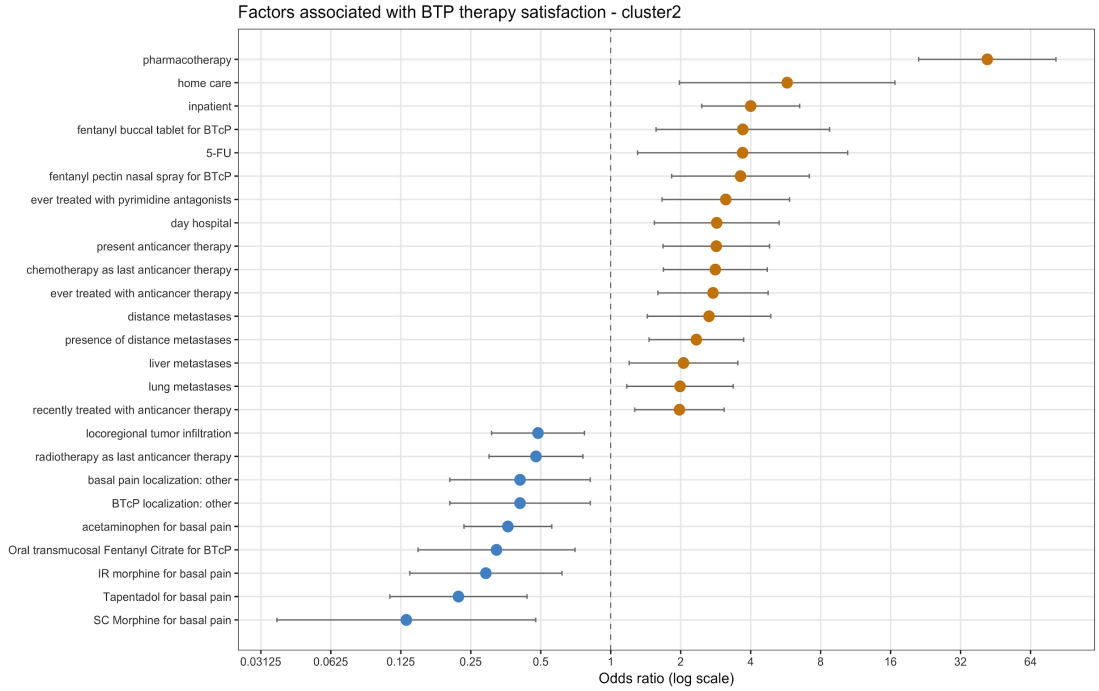
[33] A. Davies, A. Buchanan, G. Zeppetella, et al.Breakthrough cancer pain: an observational study of 1000 European oncology patients J Pain Symptom Manage. 2013; 46, 619-628

Supplementary figures and tables

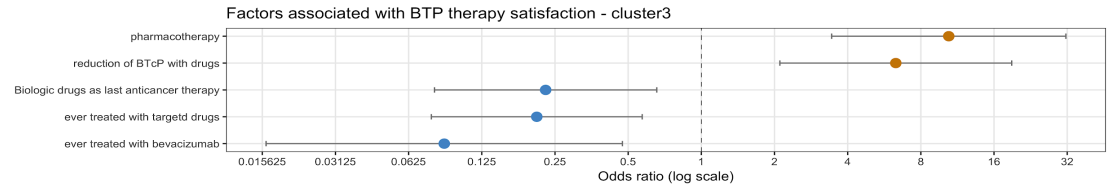


Suppl. Figure 1: silhouette statistics for cluster from 2 to 30. 12 was picked as the appropriate number of clusters for further analyses, an optimal trade-off between the average width of clusters silhouette (0.45) and the interpretability of clusters themselves.

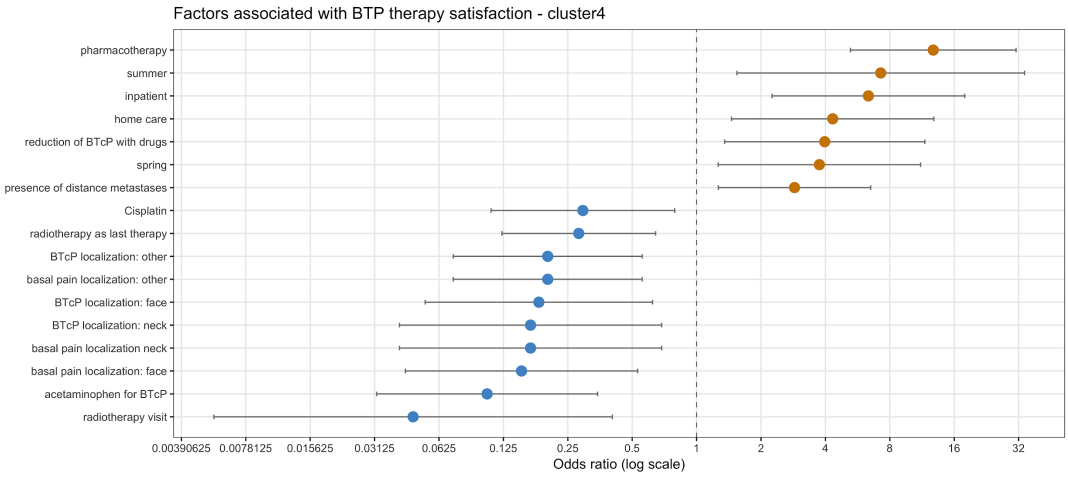
Suppl. Figure 2: features associated with therapy satisfaction for cluster 1.



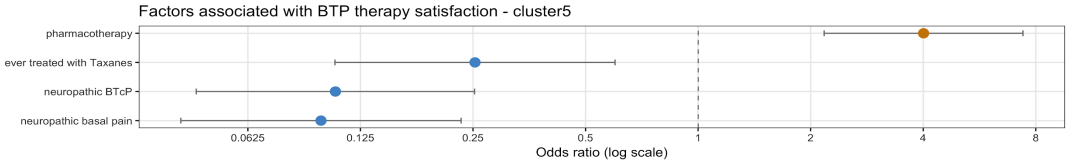
Suppl. Figure 3: features associated with therapy satisfaction for cluster 2.



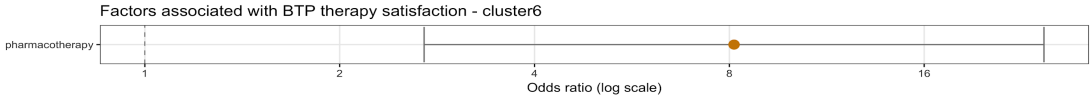
Suppl. Figure 4: features associated with therapy satisfaction for cluster 3.



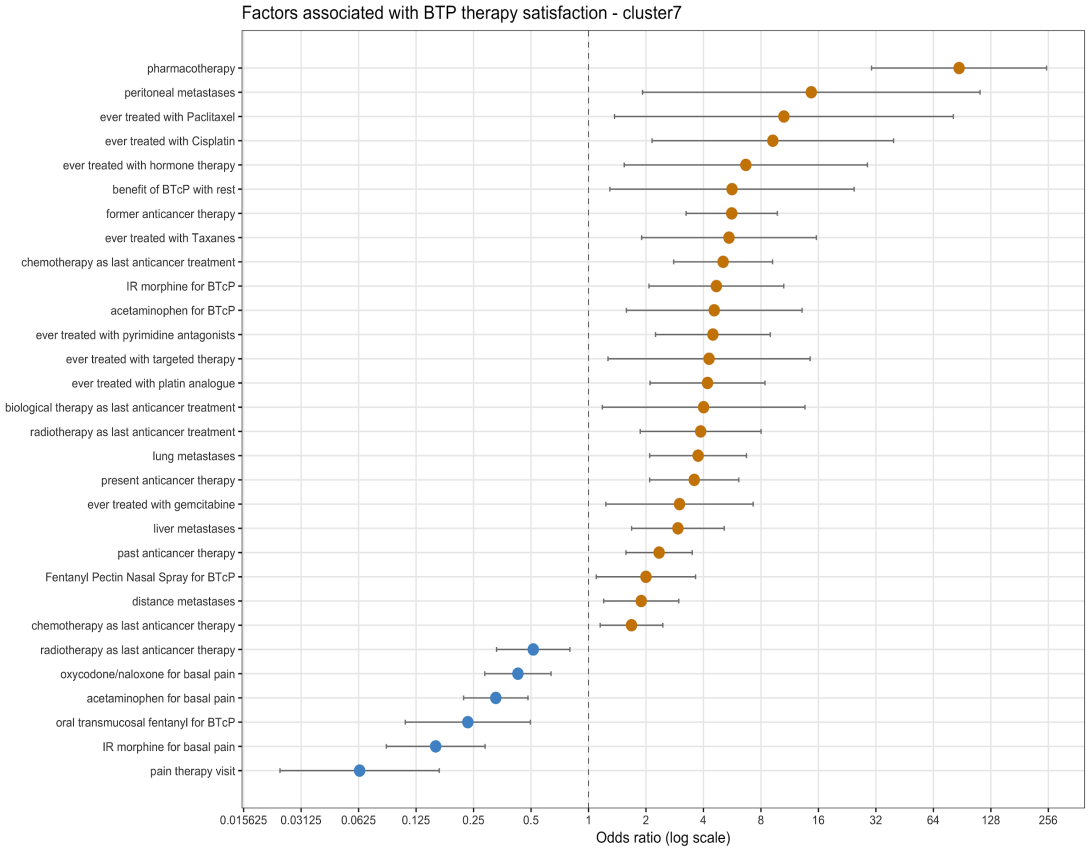
Suppl. Figure 5: features associated with therapy satisfaction for cluster 4.



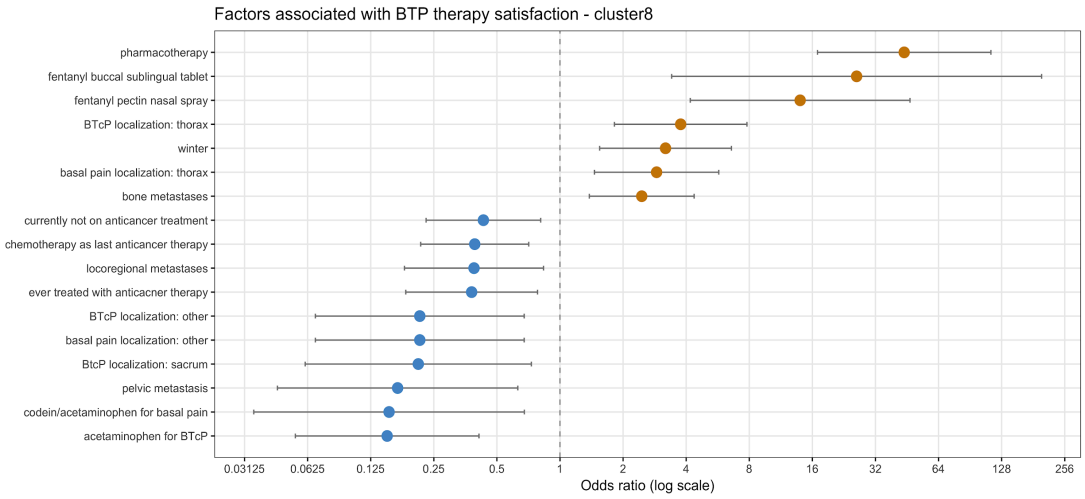
Suppl. Figure 6: features associated with therapy satisfaction for cluster 5.



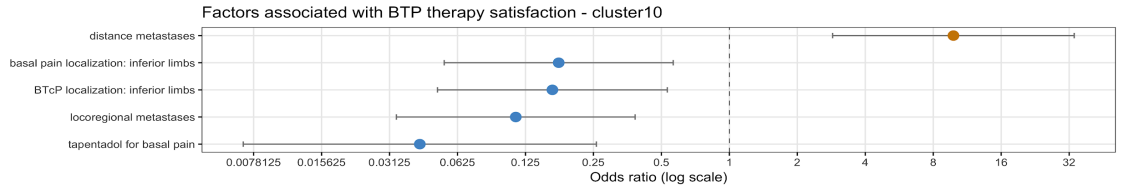
Suppl. Figure 7: features associated with therapy satisfaction for cluster 6.



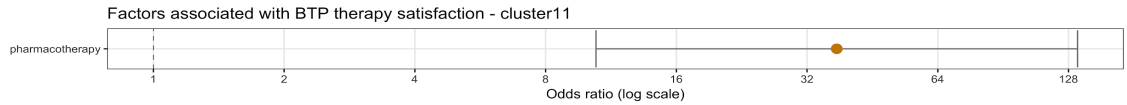
Suppl. Figure 8: features associated with therapy satisfaction for cluster 7.



Suppl. Figure 9: features associated with therapy satisfaction for cluster 8.



Suppl. Figure 10: features associated with therapy satisfaction for cluster 10.



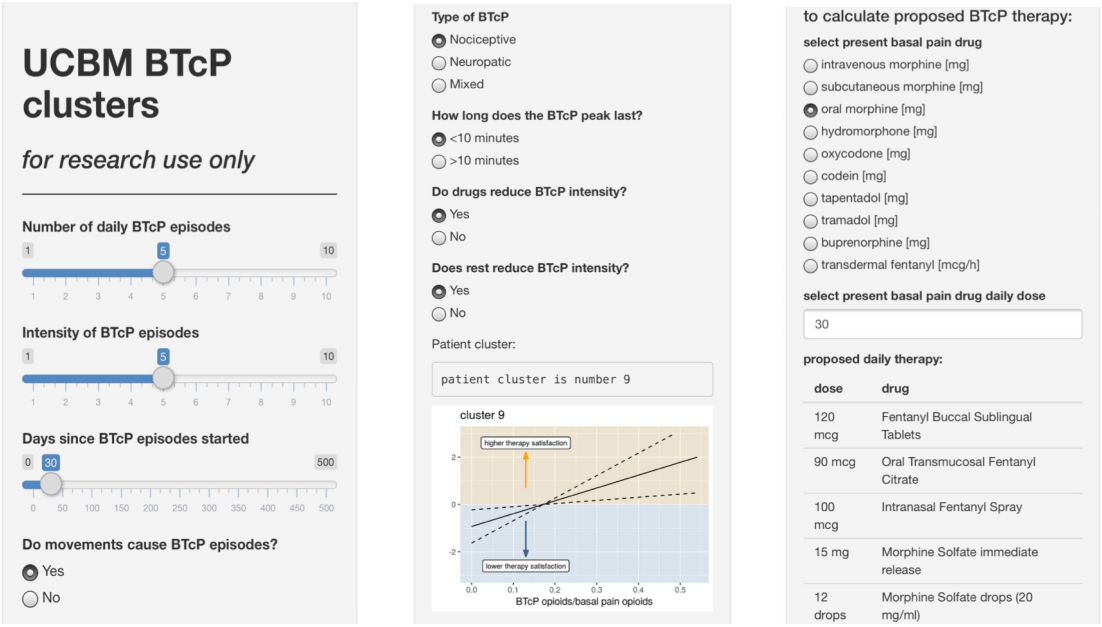
Suppl. Figure 11: features associated with therapy satisfaction for cluster 11.

	OR	lower	upper	pval	sign
(Intercept)	1.400387691	0.92074591	2.12988802	0.10823731	ns
day hospital	1.7342325	1.10021102	2.7336232	0.015532	*
home care	1.225251153	0.38968661	3.85242997	0.72283909	ns
hospice	0.616661136	0.16074775	2.36563775	0.47205487	ns
inpatient	1.538478959	1.10678007	2.13856173	0.00889373	*
palliative care visit	0.891372746	0.29533061	2.69035906	0.83507743	ns
radiotherapy visit	1.317049917	0.31996444	5.42129153	0.69707901	ns
pain therapy visit	0.265299651	0.19055651	0.36935974	1.06E-15	*
ever treated with 5-FU	1.190430722	0.70473688	2.01085731	0.50604217	ns
ever treated with Oxaliplatin	1.002256212	0.54726521	1.83552232	0.99405644	ns
tramadol for BTcP	0.282979189	0.06451491	1.24122043	0.08769596	ns
SR morphine for bsal pain	1.452976108	0.7079841	2.9819025	0.29864864	ns
tramadol for basal pain	2.953357796	0.76142036	11.455331	0.11007966	ns
IR morphine for basal pain	0.766119164	0.35931225	1.63350559	0.48159369	ns
BTcP localization: other	0.527971214	0.32275597	0.86366676	0.00944176	*
BTcP benefits from pharmacotherapy	1.86041304	1.43021101	2.4200182	2.34E-06	*

pharmacotherapy	8.254174518	5.98345482	11.3866318	2.48E-39	*
fentanyl buccal tablet for BTcP	1.834895702	0.9317812	3.61333995	0.07321914	ns
fentanyl sublingual tablet for BTcP	4.327023625	2.37312729	7.88964567	1.07E-06	*
fentanyl pectin nasal spray for BTcP	4.760433866	2.61475737	8.66685789	1.91E-07	*
SC morphine for BTcP	5.782172658	0.4524421	73.8956895	0.16837455	ns
EV morphine for BTcP	22.0337862	1.19124975	407.544877	0.03400902	*
acetaminophen for basal pain (n)	0.999747063	0.99964172	0.99985241	1.58E-06	*
oxycodone/naloxone for basal pain (n)	0.993302828	0.98967855	0.99694038	0.00023636	*
tapentadol for basal pain (n)	0.998670901	0.99700623	1.00033835	0.11083731	ns
tramadol/acetaminophen for basal pain (n)	0.980039442	0.94469703	1.01670406	0.27224035	ns
tramadol for basal pain (n)	1.00105382	0.99252868	1.00965218	0.80544856	ns
IR morphine for basal pain (n)	0.99446093	0.98320224	1.00584855	0.32922892	ns
basal pain opioids dose (n)	0.990335208	0.98716456	0.99351604	1.39E-09	*
number of BTcP episodes (n)	0.798278236	0.74390473	0.85662601	1.69E-10	*
fentanyl buccal tablet dose (n)	1.000840077	0.99827377	1.00341298	0.51302696	ns
fentanyl buccal sublingual tablet dose (n)	0.999359179	0.99725286	1.00146994	0.54342622	ns
fentanyl pectin nasal spray (n)	0.999817651	0.9970022	1.00264105	0.89708817	ns
SC morphine for BTcP (n)	0.980884795	0.79520784	1.20991637	0.85405888	ns

EV morphine for BTcP (n)	0.989039475	0.79973826	1.22314903	0.9173662	ns
BTcP opioids dose (n)	1.034532627	1.01014773	1.05950617	0.0044194	*

suppl. Table 1: multivariate regression for BTcP therapy satisfaction, all patients



Suppl. Figure 12: free online tool interface

(https://mancapaolo.shinyapps.io/UCBM_BTcPclusters/)

Acknowledgments

I would like to express my deepest respect and most sincere gratitude to all those who made this dissertation possible by their support specially my advisors, Prof. Giuseppe Tonini, Prof. Daniele Santini, Dr. Francesco Pantano and Dr. Paolo Manca, for the source of knowledge, guidance, and support that they have been to me over the past years.

I also give my special thanks to the president of the doctorate school, Prof. Paolo Pozzilli and his collaborators, for their kind support.

I would also like to give my special thank to all the members of the Department of Medical Oncology of Campus Bio-Medico for their support and patience in the moments in which the study took my time from clinical and assistance activities. Progress depends on collaboration and our group always shows it.

A big thank you to my colleagues-friends, my friends-consultants, listeners and supporters: Annalisa, Antonella, Bruno, Cecilia, Claudia, Emanuela called Coc, Loretta, Luciano called Lucio, Marco, Marianna, Michele, Tea, Tina, Vladimir, Carmela, Iacopo, Marco "Stelino", Mariella, Ale Ma, Alessandro and Ale Mi, Angelo, Claudia, Cristina, Fabrizio, Flavia, Giusy, Marco D., Matteo, Roberta, Simone and Valentina .

An endless thanks to a special person, able to read and support all my weaknesses, able to listen to me without reservations and a guide for work and life. The choice to persevere during this difficult path of PhD school is also thanks to him. Thank you Andrea.

Special thanks to a person who can teach me something even when she seems to be invisible: Martina. Your reflections are for me a manual of life and happiness hoped for.

A big thank you full of gratitude for his patience, spirit of endurance and unwavering love to Roberto: I started this path when we planned a life together and I end it with a family and a child. Life always surprises me.

Thanks to my little child Francesco: he was able to understand immediately how medical research needs perseverance and silence! He taught me how wonderful life is with its small and wonderful eyes.

I would give my special thanks full of gratitude to my second family: Patrizia, Nunzio, Aurora, Franco, Angelo, Carmelina and Alberto. Your daily gestures and your words make life a more harmonious path.

I also give my sincere gratitude to my old friends: Silvia, Claudia & Luca (my creatures and my personal translators), Antonello & Francesco, Giada (my simultaneous translator), Fabrizio, Giusy, Giulio & Ileana, Antonella and Domenico.

Thanks to all the members of my wonderful family, close and affectionate even in the darkest moment of my life ("zia Pina", "zio Ninuccio", Mariateresa & Giovanni, "zia Caterina", "zio Enzo", "zio Michele", "zia Anna", Angelo sr, Iris sr, Angelo jr, Iris jr, Tania & Giovanni).

An atypical thank you to my brother Giuseppe: he has practically disappeared from my life in the last three years, but in the darkest or happiest moments he has reappeared to remind me how important I am for him and what I can aspire to. I hope that these words will be an invitation for him to re-enter into my life with love and enthusiasm.

Last but not least, the greatest example of my life: my father. He unconsciously guides my every choice, resolves all my doubts, strengthens my every feeling, supports me in every difficulties. Thanks for what you have been, for what you are and for what you will be for me.

Finally an immense thank you. A "thank you" that the word cannot explain. This dissertation and all my life as an oncologist were conceived just thinking about you. It is for you and for all the people who fight and will fight against cancer that I will make my work a masterpiece and an endless mission. I know that you are here, with your imperceptible presence. You move everything without noise, as the celestial spheres turn. Thank you "mamma", fount of my certainty.