
Machine Learning-Based Assessment Indicates Patient-Reported Outcomes to Be Less Reliable than Clinical Parameters for Predicting 1-Year Survival in Cancer Patients.

[Maria Rosa Salvador Comino](#)*, Paul Youseff, Anna Heinzelmann, [Florian Bernhardt](#), Christin Seifert, [Mitra Tewes](#)

Posted Date: 1 November 2023

doi: 10.20944/preprints202311.0042.v1

Keywords: Patient Reported Outcome Measure; Artificial Intelligence; Machine Learning; Predictive Analytics; Cancer Patients; Palliative Care



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Machine Learning-Based Assessment Indicates Patient-Reported Outcomes to Be Less Reliable than Clinical Parameters for Predicting 1-Year Survival in Cancer Patients

Maria Rosa Salvador Comino ^{1,*}, Paul Youssef ^{2,3}, Anna Heinzelmann ¹, Florian Bernhardt ⁴,
Christin Seifert ³ and Mitra Tewes ¹

¹ Department of Palliative Medicine, West German Cancer Center, University Hospital Essen, University of Duisburg-Essen, Essen, Germany; maria.salvador@uk-essen.de, mitra.tewes@uk-essen.de and anna.Heinzelmann@ext.uk-essen.de

² Institute for Artificial Intelligence in Medicine (IKIM), University of Duisburg-Essen, Germany.

³ Department of Mathematics and Computer Science, University of Marburg, Germany; christin.seifert@uni-marburg.de and paul.youseff@uni-marburg.de

⁴ Department of Palliative Care, West German Cancer Center, University Hospital Muenster, University of Muenster, Muenster, Germany; florian.bernhardt@uni-muenster.de

* Correspondence: maria.salvadorl@uk-essen.de; Tel.: +49-201-723-82318; Fax: +49-201-723-4011

Simple Summary: Artificial intelligence techniques can help predict survival among cancer patients and might help distinguish who benefits the most from palliative care support. We explored the importance of different self-reported and objective variables, which were collected via electronic patient self-reported outcome measure and medical records. We used these variables to predict 1-year mortality. The results show that objective variables are much more predictive than subjective variables and indicate that subjective burden cannot be reliably used to predict survival. Further research in this area is needed to clarify the role of self-reported patient's burden and mortality prediction via artificial intelligence.

Abstract: Machine learning (ML) techniques can help predict survival among cancer patients and might help with a timely integration in palliative care. We aim to explore the importance of subjective variables self-reported and collected via electronic patient reported outcome measure (ePROM) for survival prediction. A total of 256 advanced cancer patients met the eligible criteria. We analyzed objective variables collected from electronic health records, subjective variables collected via ePROM and all clinical variables combined. We used logistic regression (LR), decision trees, and random forests to predict 1-year mortality. Receiver operating characteristic (ROC) curve - area under the curve (AUC) and the ML models feature importance were analyzed. The performance of all variables for predictions (LR reaches 0.80 [ROC AUC] and 0.72 [F1 Score]) does not improve over the performance of only clinical non-patient reported outcome (non-PRO) variables (LR reaches 0.81 [ROC AUC] and 0.72 [F1 Score]). Our study indicates that patient-reported outcome (PRO) variables, which measure subjective burden, cannot be reliably used to predict survival. Further research in this area is needed to clarify the role of self-reported patient's burden and mortality prediction via ML.

Keywords: patient reported outcome measure; artificial intelligence; machine learning; predictive analytics; cancer patients; palliative care

1. Introduction

The benefits of an early palliative care (PC) involvement during advanced malignancy are well documented [1][2] and supported by experts worldwide [3][4]. In one Delphi study, over 80% of 60

international experts agreed that an specialized palliative care (SPC) consultation would be appropriate within 3 months of advanced cancer diagnosis for patients with a median survival time of 1 year or less [5]. However, the optimal model of referral to SPC among advanced cancer patients remains unclear and referral typically occurs late [6]. Some of the reasons are that patients are often never identified or referred at any point during the course of their illness, commonly for an end-of-life care [7]. Therefore, referral models to SPC are in urgent need for optimization.

There are different approaches to promote PC integration in cancer patients [8][9]. The time-based model or early integration model relies heavily on prognosis and PC is introduced from the time of diagnosis of advanced disease. I.e., patients with a poor prognosis are considered for referral. The need-based PC model is a personalized PC based on patient' needs and delivered at the optimal time and setting. It focuses on early screening of systematic needs, where the SPC is usually reserved for patients with the greatest needs [9]. Also, the combination of different models have been pointed out to facilitate earlier PC access [10].

The need-based PC delivery model coupled with a systematic, regular, and early symptom assessment helps to prioritize patients who could benefit the most from PC. This model, known as the timely PC model, is based on routine screening. Therefore, patient reported outcome measures (PROMs) have gained weight as a routine assessment and are currently recommended as part of the clinical routine. In addition, PROMs have been shown to maximize the identification of symptoms as they arise, and thus to improve patients' quality of life [11], overall survival [12–14] as well as to reduce emergency room visits and hospitalizations [15,16].

The use of prognostic methods from the field of artificial intelligence (AI), and more specifically machine learning (ML), is rapidly increasing. It has been shown to have tremendous potential when applied to medicine [17,18]. As in other domains, ML methods have been widely used in PC. For example, Guo et al. [19] predicted lung infection during PC chemotherapy using a neural network, Van Helden et al. [20] used clustering to predict response to palliative therapy. Many other works focus mostly on survival prediction followed by data annotation and predicting morbidity or response to palliative therapy [21], leverage electronic health records [22], demographic information and other clinical features [23] for predicting survival. Furthermore, prediction of mortality via ML facilitates earlier conversations amongst clinicians and their patients about prognosis, patients goals of care and end-of-life treatment preferences [24] [25]. These ML approaches are promising given that for some models the survival prediction is key, and the attempts to estimate prognosis in PC are generally rather inaccurate [26].

To the best of our knowledge, the use of ML for analysis of self-reported symptom burden and other subjective measures regarding survival prediction has not been investigated yet. The aim of this study is to analyze the importance of subjective measures, collected via electronic patient reported outcome measure (ePROM), for survival prediction and SPC integration by using ML.

2. Materials and Methods

Study design and sample description

We conducted a retrospective, monocentric observational pilot-study including patients of a German Comprehensive Cancer Centre based on our routine use of ePROM. We collected non-patient reported outcome (non-PRO) variables such as patient demographics and characteristics by reviewing medical records, from where gender, age, Eastern Cooperative Oncology Group (ECOG), cancer diagnosis, type and stage were obtained. In addition, the answers given in the ePROM questionnaire (as described below) were collected. We obtained the date of death from the local residents' registration office. Patients were classified in two groups: those who died within 1 year after the screening and those who were alive 1 year after screening.

The data analysis was in accordance with the ethical standards of the institutional and local ethical review committee and was approved by the institutional ethics board (20-9779-BO).

Data collection

A total of 1089 cancer patients were asked to fill out ePROM during the period from April 1st, 2020, to March 31st, 2021. Out of those patients, 672 filled out the ePROM questionnaire. We excluded 407 patients (of 672) because of Stadium WHO <3 tumor (n = 358), lack of solid tumor (n = 29), missing relevant medical data (n = 13) and underage (n = 7) resulting in 256 patients for our analysis.

Out of the 265 eligible patients, 70 patients died within 1 year after the screening and 195 patients lived longer than 1 year after screening (Figure 1). The screening was mainly carried out in the wards of neurosurgery (n = 413), nuclear medicine (n = 324), and dermatology (n = 203) (Figure 1) due to the ongoing implementation of ePROM in our hospital at the time of the data collection.

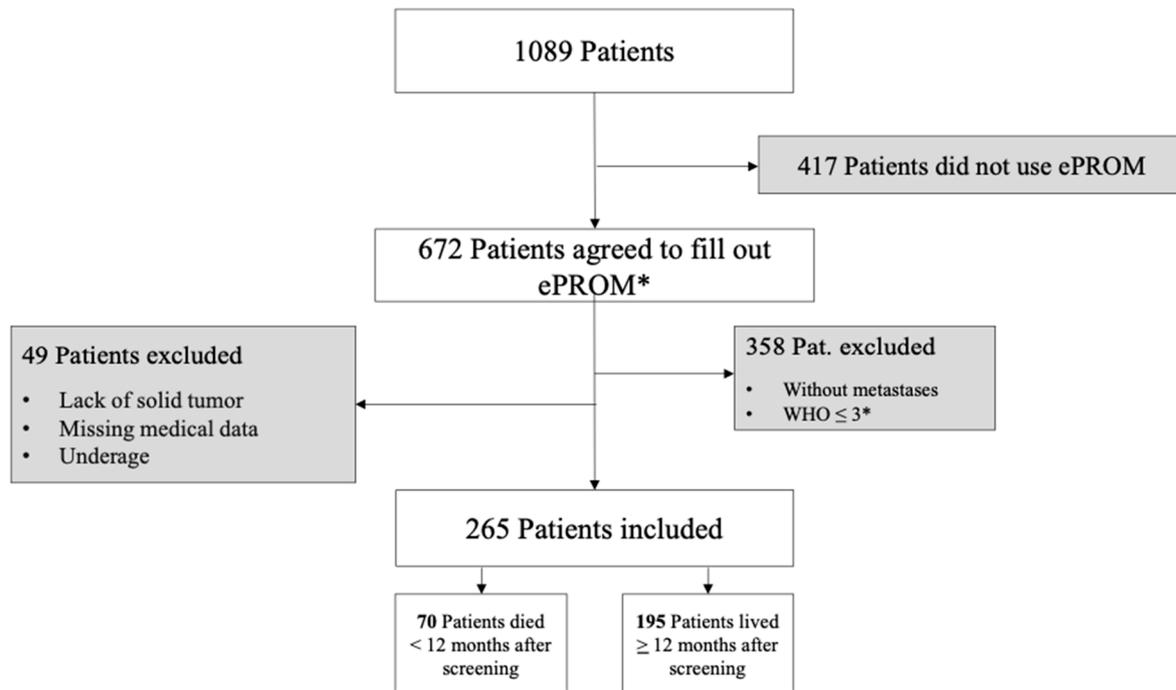


Figure 1. Data from the process of recruitment and enrollment of patients. * ePROM: electronic patient reported outcome measure, * WHO Classification.

ePROM (electronic patient reported outcome measure)

Our ePROM assessment is called ePOS (a German abbreviation for electronic psycho-oncological and palliative care screening) and assesses physical and psychological symptom burden. It contains five validated instruments to facilitate an early identification of patients with PC needs [27]. ePOS is a digital based questionnaire, distributed by medical assistants and filled out by inpatients and outpatients. After completion of the questionnaire, the provided information is transferred digitally to the hospital information system and accessible for the responsible medical provider.

ePOS consists of 51 items distributed in five validated, multidimensional symptoms and needs assessment questionnaires: the Personal Health Questionnaire 8 (PHQ8) [28] that shows depressive symptoms; the Generalized Anxiety Disorder Questionnaire (GAD7) [29] that explores anxiety and fears; the Minimal Documentation System (MIDOS 2) [30], which is a German version of the Edmonton Symptom Assessment Scale and analyzes symptom burden; the Hornheider Screening Instrument (HSI) [31] comprises seven questions about physical and mental health and the Distress Thermometer [32], which quantifies the load of burden felt within the past 10 days. In addition, it also assesses the patient's subjective need for palliative or psycho-oncological support and the patient's subjective general condition (for a detailed sample of all questions answered by the patients see our supplements).

Not all variables collected in ePROM are relevant for survival prediction, and we removed the following variables from the analysis: "On which topics/occasions may we contact you by email?" and "Intensity of other feelings during the past week".

Non-patient reported outcome variables

Non-patient reported outcome (non-PRO) variables were collected from medical records. We obtained gender, age, ECOG, cancer diagnosis, type and stage, patient demographics, date of death, medications, body measurements, therapy, level of care and number of visits to the emergency room.

Prognostic machine learning models

For this small data set we employ classical machine learning models. Specifically, we use a linear classifier, logistic regression (LR) with L1 penalty, as well as two non-linear models, a decision tree (DT), and a random forest (RF) to predict 1-year mortality. We also use a simple model that always predicts the most frequent class in our dataset for all instances, i.e., it always predicts the survival class. We refer to this model as "Majority". Adding this model helps us inspect how much can be gained from using advanced models compared to a model that makes use of a simple statistic, i.e., the most frequent class. We excluded Deep Learning models, since they require a sufficiently large number of patients for training, our case consists of 265 patients with 89 variables¹. In this regard it is important to point out that our goal was to understand the relative importance of patient reported outcome (PRO) variables for the prediction. Thus, we aimed for predictive models with high accuracy, but left potential optimization, such as extensive model selection and hyperparameter tuning for future work. Our dataset is highly imbalanced with more than 70% of the patients belonging to the survival class. To counteract this class imbalance, we used a weighted loss function that adapts the weights of the majority and minority classes based on their frequency during training. We used the standard settings of the scikit-learn library for all models². Additionally, we set the inverse of regularization strength to 0.2 in LR, and the minimum number of samples in each leaf in the DT to 20, to counteract overfitting.

We experimented with three sets of variables: PRO variables (48 subjective variables), non-PRO variables (41 objective variables) and all variables (89 variables) combined to investigate the effect that PRO variables have on survival prediction.

Evaluation procedure and metrics

We trained and evaluated our prognostic models with leave-one-out cross-validation as well as report F1 score and receiver operating characteristic (ROC) curve - area under the ROC curve (AUC).

Leave-one-out cross-validation (LOOCV): Training and evaluating ML models includes having a training set, based on which the model learns the task, and a test set, on which the model is evaluated. In LOOCV, all the available data instances are used for training except for one example, which is used for evaluation. However, evaluating only on one example is not enough and would give a biased estimate regarding the model's performance. This is why, in LOOCV, the training and evaluating process is repeated for each data instance, i.e., the current data instance is excluded, and the rest of the data instances are used for training. LOOCV enables training on almost all available data (excluding only one instance at a time) and evaluating on all data points.

F1 score: The F1 score is an evaluation metric that combines both precision and recall (also known as sensitivity). Precision quantifies how often the model was correct in predicting the positive class (mortality), whereas recall refers to the proportion of the patients from the positive class that are correctly identified by the model. F1 score is the harmonic mean of the Precision and Recall, i.e., $F1 = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$.

ROC-AUC: The ROC plots the false positive rate (1-specificity) on the x-axis against recall (sensitivity) on the y-axis. Each point on the curve represents a specific threshold used to classify an instance as positive (predicting mortality) based on the model's probability for predicting the positive

¹ In line with [64], we ran initial experiments using a multi-layer perceptron classifier (MLP). This model demonstrated comparable performance to that of the Random Forest model, while generally being harder to train. We excluded the MLP from further analysis.

² <https://scikit-learn.org/>

class. The AUC quantifies the overall performance of the model. An AUC value of 1 indicates a perfect classifier, whereas an AUC value of 0.5 suggests that the model is equivalent to a classifier that randomly guesses the outcome.

Variable importance

Since LR yields the highest prediction accuracy, we investigated which variables LR models used for its predictions. In LR models, the coefficients associated with each variable indicate its importance for predicting mortality. More specifically, the magnitude of each variable is proportional to its importance, and its direction (sign) indicates whether the variable contributes to predicting mortality (positive) or survival (negative). The coefficients of all variables are shown in Table 3. Furthermore, two experienced PC physicians (M.T. and M.R.S.C) independently assessed the importance of each variable regarding its prognostic value and assigned a relevance as "important for prognosis" or "not important for prognosis". Furthermore, we researched the current literature regarding oncological prognostic factors.

3. Results

Patient characteristics

Patient characteristics are summarized in Table 1. A total of 265 patients were eligible for our study. The most common cancer types were cancer of the nervous system (n = 63, 23.8%), dermatological cancers (n = 57, 21.5%) and gastrointestinal cancer (n = 42, 15.8%). The most prevalent metastases were cerebral (n = 65, 24.5%), bone (n = 57, 21.5%) and pulmonary (n = 56, 21.1%). Seventeen (6.4%) out of all 265 patients had an inpatient SPC consultation. 70 patients died within 1 year after screening and 195 patients were still alive at time of the study. We used χ^2 -test to analyze the relation between nominal variables such as sex, primary tumor, and the number of SPC consultations. Additionally, we used Mann-Whitney-U-test to assess whether there was a significant difference in ordinal variables such as physical and mental condition between the two groups. Related p-values as seen below. The general survival rate in our study population was 73.6%.

Table 1. Patient characteristics according to survival and symptom burden.

| | Characteristics | Pat. survived < 1 year after screening (n = 70) | Pat. survived > 1 year after screening (n = 195) | p value |
|----------------------------------|------------------|---|--|---------|
| Sex n (%) * | Female | 24 (34.3) | 64 (40.5) | 0.373 |
| | Male | 46 (65.7) | 94 (59.5) | |
| Age mean (SD) ** | | 60.13 (±14.20) | 57.86 (±13.69) | |
| Age at diagnosis (SD) | | 57.23 (±14.79) | 53.15 (±14.24) | |
| Age at death (SD) *** | | 60.64 (±14.24) | | |
| Site of primary tumor n (%) * | Nervous system | 20 (28.6) | 31 (19.6) | 0.117 |
| | Gastrointestinal | 8 (11.4) | 30 (19.0) | |
| | Dermatological | 9 (12.9) | 41 (25.9) | |
| | Urogenital | 11 (15.7) | 13 (8.2) | |
| | Lung | 9 (12.9) | 15 (9.5) | |
| | Sarcom | 5 (7.1) | 7 (4.4) | |
| | Breast | 1 (1.4) | 5 (3.2) | |
| | Endocrinological | 2 (2.9) | 9 (5.7) | |
| | Others | 5 (7.1) | 7 (4.4) | |
| Site of metastases n (%) | Pulmonary | 21 (30.0) | 28 (17.7) | |
| | Hepatic | 14 (20.0) | 33 (20.9) | |
| | Ossary | 18 (25.7) | 35 (22.2) | |
| | Visceral | 10 (14.3) | 14 (8.9) | |
| | Cerebral | 22 (31.4) | 30 (19.0) | |
| SPC consultation n (%) * | | 13 (18.6) | 3 (1.9) | <0.001 |
| Physical condition n (%) **** | Very poor | 0 (0.0) | 0 (0.0) | 0.047 |
| | Rather poor | 14 (20.6) | 19 (12.0) | |
| | Moderate | 37 (54.4) | 70 (44.3) | |
| | Rather good | 17 (25.0) | 68 (43.0) | |
| | Very good | 0 (0.0) | 1 (0.6) | |

| | | | | |
|----------------------------------|-------------|-----------|-----------|-------|
| Mental condition n (%) **** | Very poor | 1 (1.4) | 1 (0.6) | 0.109 |
| | Rather poor | 12 (17.1) | 24 (15.3) | |
| | Moderate | 36 (51.4) | 65 (41.1) | |
| | Rather good | 21 (30.0) | 67 (42.7) | |
| | Very good | 0 (0.0) | 0 (0.0) | |
| General well-being n (%) **** | Very Poor | 1 (1.5) | 0 (0.0) | 0.040 |
| | Rather Poor | 10 (14.9) | 16 (10.3) | |
| | Moderate | 31 (46.3) | 64 (41.3) | |
| | Rather Good | 25 (37.3) | 63 (40.6) | |
| | Very Good | 0 (0.0) | 12 (7.7) | |

*Significant difference was analyzed via χ^2 -test

**At time of the study

***If deceased

****Significant difference was analyzed via Mann-Whitney-U-test

Predictive model performance

In our predictive ML model, we report ROC-AUC and F1 Score (Table 2). The predictive performance solely based on PRO variables worked only slightly better than the naive baseline, which has not learned anything from the data, but simply predicts the majority outcome ("survive"). The baseline achieves 0.42 F1, whereas the performance of the other models varies between 0.45 and 0.55. Non-PRO variables are much more predictive than PRO variables (LR reaches 0.72 F1, and 0.81 ROC AUC). The performance of all variables for predictions does not improve over the performance of only non-PRO variables (0.72 F1 for LR using non-PRO and all variables) (Figure 2a and 2b). (See Appendix A for the ROC curve with non-PRO variables).

Table 2. Comparison between patient self-reported outcome (PRO) variables, non-PRO (clinical) variables, and all variables together via ML. Results according to machine learning (ML) "leave-one-out cross" validation. Reporting ROC-AUC and F1 score.

| | PRO variables (subjective) | | non-PRO variables (clinical) | | all variables | |
|-----------------|-------------------------------|----------|---------------------------------|-------------|---------------|----------|
| | ROC AUC | F1 Score | ROC AUC | F1 Score | ROC AUC | F1 Score |
| Majority | 0.50 | 0.42 | 0.50 | 0.42 | 0.50 | 0.42 |
| LR | 0.55 | 0.52 | 0.81 | 0.72 | 0.80 | 0.72 |
| DT | 0.56 | 0.55 | 0.73 | 0.61 | 0.76 | 0.67 |

| | | | | | | |
|-----------|------|------|------|------|------|------|
| RF | 0.56 | 0.45 | 0.77 | 0.65 | 0.76 | 0.58 |
|-----------|------|------|------|------|------|------|

*LR: logistic regression model

*DT: decision tree model

*RF: random forest model

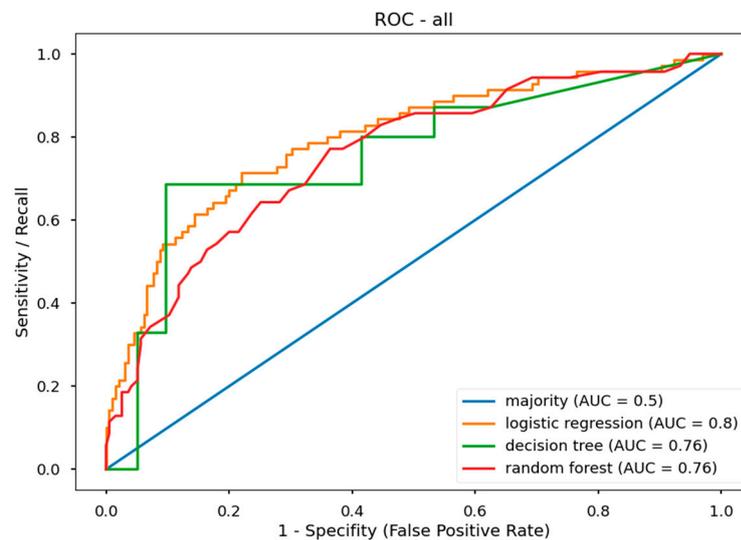


Figure 2. a.

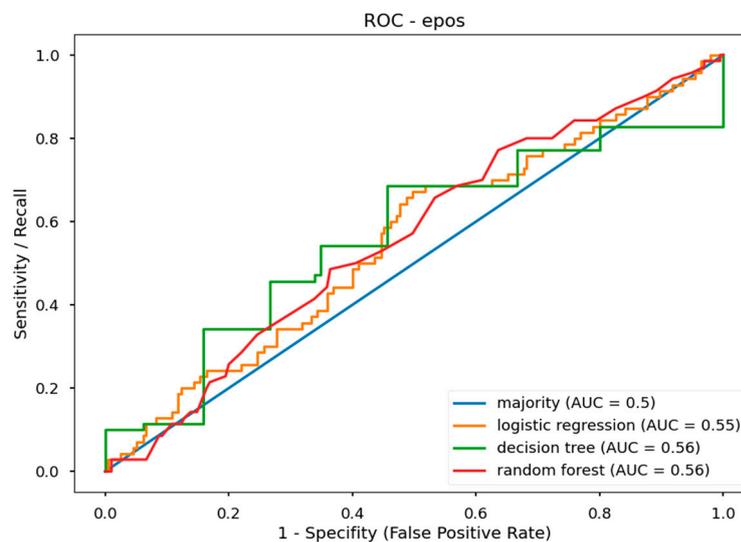


Figure 2. b.

Figure 2. 2a and 2b. Figure 2a represents ROC curves using all variables (PRO and non-PRO). Figure 2b represents ROC curves with only PRO variables.

Prognostic value

The variables are classified via ML regarding its relevance between important or unimportant, as well as a positive or negative prognostic factor for the patient' prognosis. Out of a total of 89 variables assessed via ML, 38 were included given that they had the highest absolute value (Table 3a)

and 51 were excluded as they were assessed as unimportant via ML. Regarding non-PRO variables, a total of 14 variables were included and 27 were excluded (Table 3b).

Table 3. a. Non-PRO (clinical) and PRO (subjective) variables.

| All variables | | |
|---------------|--|----------------------------|
| Weight | Variable | Expert Assessment |
| 0.564 | The patient did receive antiemetics | relevant for prognosis |
| 0.436 | Body mass index (BMI) | relevant for prognosis |
| 0.363 | The patient did not receive immunotherapy | not relevant for prognosis |
| 0.273 | Male | not relevant for prognosis |
| 0.265 | Level of care | relevant for prognosis |
| 0.26 | Number of emergency room attendances | relevant for prognosis |
| 0.12 | Initial stage | relevant for prognosis |
| 0.115 | The patient did not receive antibody therapy | relevant for prognosis |
| 0.111 | How often have you felt affected by fatigue or the feeling of having no energy during the past 2 weeks? | relevant for prognosis |
| 0.1 | ECOG/Performance status | relevant for prognosis |
| 0.076 | Intensity of constipation | not relevant for prognosis |
| 0.067 | Intensity of weakness | relevant for prognosis |
| 0.054 | The patient did receive adjuvant chemotherapy | not relevant for prognosis |
| 0.044 | How often have you felt affected by restlessness, making sitting difficult during the past 2 weeks? | not relevant for prognosis |
| 0.044 | How often did you feel affected by a decreased appetite or an excessive desire to eat during the past 2 weeks? | relevant for prognosis |
| 0.04 | The oncological therapy was terminated | relevant for prognosis |
| 0.008 | Intensity of pain | not relevant for prognosis |
| -0.016 | How many times have you felt affected by anxiety, as if something bad might happen during the past 2 weeks? | not relevant for prognosis |

Table 3b. Only non-PRO (clinical) variables.

| Non-PRO variables | | |
|-------------------|--|----------------------------|
| Weight | Variable | Expert Assessment |
| 0.634 | The patient did receive antiemetics | relevant for prognosis |
| 0.407 | Body mass index (BMI) | relevant for prognosis |
| 0.295 | The patient did not receive antibody therapy | not relevant for prognosis |
| 0.286 | Level of care | relevant for prognosis |
| 0.283 | Number of emergency room attendances | relevant for prognosis |
| 0.262 | Male | not relevant for prognosis |
| 0.153 | The oncological therapy was terminated | relevant for prognosis |
| 0.151 | The patient did receive adjuvant Chemotherapy | not relevant for prognosis |
| 0.119 | Initial stage | relevant for prognosis |
| 0.109 | ECOG/Performance status | relevant for prognosis |
| 0.051 | The patient did not receive immunotherapy | not relevant for prognosis |
| 0.035 | The patient did receive opioids of class WHO3 | not relevant for prognosis |
| -0.252 | Number of inpatient admissions after the screening | relevant for prognosis |
| -1.351 | The oncological therapy was not terminated | relevant for prognosis |

ECOG: Eastern Cooperative Oncology Group

WHO3: World Health Organisation' analgesic ladder step 3

Expert Assessment

Two experts in the field of PC were asked to assess independently the importance of each variable regarding its prognostic value. There was disagreement in 7 variables from a total of 38. Each disagreement was discussed and compared with the current literature, where supporting references helped the final decision. A common agreement was possible in each variable. The expert consensus showed a high agreement (81%) regarding the importance of the variables used by the ML for the survival prediction (Table 3).

Table 3a and 3b. Patient self-reported outcome (PRO) and non-PRO variables and its weighting according to ML and expert assessment. The table represents the variables that ML weighted as positive prognostic factors for survival (top-left) and the variables, which were weighted as negative prognostic factors for survival (bottom-left). The importance of the variables after the expert consensus is described as well (right-side).

4. Discussion

The role of subjective measures in survival prediction among advanced cancer patients is unknown. In the present retrospective, observational study we analyzed for the first time the role of an ML method using subjective measures collected via ePROM for survival prediction.

In general, our results indicate that subjective variables, in our study collected via ePROM, cannot be reliably used to predict survival, and that adding or excluding them does not make a significant difference regarding the predictive performance (comparison of results under non-PRO and all variables, Figures 2a and 2b). Thus, ePROM might not be predictive for a time-based PC model, where prognosis is key, although it clearly has an important role in the need-based referral approach, as stated above.

Currently, in the field of PC, new AI and ML techniques are evolving and could assist in allocating PC interventions more effectively. However, these methods are indeed mainly used to predict mortality [33] [21]. Some of them focus on specific patient populations such as dementia [34] or older patients with hip fracture [23], or focus on larger samples from regional health care plans [35], Medicare & Medicaid sources [36] or include all hospitalizations episodes [37]. Its impact in clinical practice has been evaluated as well with promising results where an electronic health record mortality prediction model successfully triggered timely PC referral and thus improved the delivery of hospital-based PC [38]. These studies are of great interest given that physicians tend to overestimate prognosis and only a few prognostic scores have been shown superior to clinician predictions [39]. Interestingly, PC physicians accurately predict survival especially for short survival time periods [40]. This uncertainty regarding prognosis estimation is a relevant barrier for the integration of PC in oncology [41] as well as the limited PC personal and institutional resources, where identifying more patients with unmet PC needs does not necessarily translate into better care [42]. Thus mortality prognostication is a key factor but also one more of the many factors that should be involved in SPC-referral models [43].

In PC, the role of the self-reported subjective symptom burden in cancer patients is gaining weight. The current PC approach recommends that patients with the highest symptom burden and distress would benefit the most from SPC, regardless of diagnosis or prognosis [44]. The patient's subjective need for SPC has been till now only vaguely studied and has been shown to help prioritize which patients would benefit the most from SPC [45]. The ML used in this study shows that self-patient' burden assessment cannot be reliably used to predict survival. However, models that recommend patients for PC referral should probably not only focus on mortality given that periods of crisis and clinical deterioration are much more relevant to determine when PC is needed [21] and where unmet PC needs do not necessarily show only close to death [46]. In this regard, the combination of different referral models, in our case through ePROM (need-based) and ML for prognostic mortality (time-based) should be further investigated as they might be necessary for an accurate SPC referral, as already supported by some researchers [10].

The variables used in this study were classified via ML regarding its relevance between important and unimportant and as positive or negative prognostic factors for the patient's prognosis. The variables weighed as positive prognostic factors supported by the expert consensus were mostly accurate and already documented in the literature: need of antiemetics [47], body mass index [48] [49], level of care [50], number of emergency room attendances [51] [52], initial stage [53], decrease in appetite [50], ECOG/Performance status [49,54] male sex [54] or the oncological therapy was terminated [49] (see Table 3). The negative prognostic factors classified via ML supported by the expert consensus and already documented in the literature were: the oncological therapy was not terminated [49], number of inpatient admissions after the screening, subjective physical status and general well-being [49]. Other variables were weighted as important via ML, however, were excluded after the expert consensus and the literature research: not receiving antibody therapy, intensity of constipation, adjuvant chemotherapy, not receiving immuno- or antibody therapy, or strong opioids. The variables feeling of restlessness, anxiety or difficulties relaxing during the last 2 weeks are related to emotional distress, which has been shown to play a role in prognosis [49], however the limitation of the ePROM question "during the last 2 weeks" prevent us further assessing this variable. The

variable "not receiving immunotherapy" was assessed as a positive prognostic factor via ML but rejected from the expert panel, given that immunotherapy has been shown to contribute to prognostic improvement among glioblastoma patients [55]. However, it can be assumed that not all prognostic factors have been identified yet and ML models might focus on variables, which might warrant further exploration.

Although ML algorithms may truthfully predict mortality risk in cancer patients, there are several barriers in order to achieve a wider implementation: oncology clinicians are concerned about algorithm accuracy, biases, and ethical issues of prognostic disclosure [56]. Furthermore, a comparison of the actual literature regarding survival should be carefully approached, since the primary outcome of prognostic studies in PC and oncology has not been established yet [57]. In addition, a secure degree of digitalization and integration of health data is needed [58]. Moreover, some prognostic factors are disease-specific and not consistent across cancer types and stages [59] [48]. In this regard, it is important to note that specific guidelines for each malignancy were not used in our study. Laboratory parameters have been described as indicators of poor prognosis [60,61] [62], however the integration of such parameters are beyond the scope of our study. Thus, innovative PC delivery using different ML models might help and might serve as a promising tool to identify PC patients who are most in need.

Our investigation has several limitations. First, our analysis was conducted retrospectively in a single center among patients with advanced cancer. Thus, our results may differ from studies with non-oncological participants or in different settings. Second, our final sample size is limited and is neither representative nor balanced, thus further analysis with a larger sample size is required. Third, using ePROM might exclude patients who are not able to complete this self-questionnaire (high symptom burden, language barrier, somnolence, very poor general condition, etc.). Fourth, important variables not included in our ePROM or electronic health records might impact these results and should be further studied. Fifth, the general survival rate in our study population was relatively high, 73.6%, given that all patients had an advanced disease and as an example, the relative survival after 1 year for lung cancer in both sexes is 28.7% [63]. This unexpected lower mortality rate could be a study population' bias.

The use of self-reported symptom burden and its implication regarding SPC referral among cancer patients together with digitalization is promising where ML has a growing role in improving diagnostic, treatment, and prognostic in PC. The integration of ePROM predicting short-term relevant clinical deterioration and ML predicting overall mortality is promising and should be further explored. Moreover, clinical studies are required to improve how AI and ML models are integrated in our healthcare system and how they can nurture the field of PC.

5. Conclusions

The optimal moment for a SPC referral among cancer patients remains unclear. Although the use of ePROM has numerous benefits in PC, our data and ML techniques used in this study indicate that PRO variables, which measure the patient's subjective burden, cannot be reliably used alone to predict survival in advanced cancer patients and that AI-based tools may be more suitable for this task. Further research in this area is needed to clarify the role of self-reported patient's burden and mortality prediction via ML to achieve better care for patients with unmet PC needs.

Author Contributions: All authors contributed to the study conception and design. Material preparation and data collection were performed by A.H. with the assistance of M.T. All authors contributed to analysis. The first draft of the manuscript was written by M.R.S.C. and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The Open Access Publication Fund of the University of Duisburg-Essen supported our study.

Institutional Review Board Statement: All methods and data analysis were in accordance with the ethical standards of the local ethical review committee of the University Duisburg-Essen (20-9779-BO, date of approval January the 14th 2021).

Informed Consent Statement: Informed consent was not required from our ethic committee since the assessment was carried out anonymously and retrospectively from the existing medical records and questionnaires and was in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Data Availability Statement: Data is available from the corresponding author on reasonable request.

Acknowledgment: We acknowledge support by Professor Ulrich Sure (Department of Neurosurgery), Professor Ken Hermann (Department of Nuclear Medicine) and Professor Dirk Schadendorf (Department of Dermatology).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vanbutsele, G; Pardon, K; Van Belle, S et al. Effect of Early and Systematic Integration of Palliative Care in Patients with Advanced Cancer: A Randomised Controlled Trial. *Lancet Oncology* **2018**, *19*(3), 394–404, doi:10.1016/s1470-2045(18)30060-3.
2. Haun, M.W.; Estel, S.; Rücker, G. et al. Early Palliative Care for Adults with Advanced Cancer. *Cochrane Db Syst Rev* **2017**, *2017*, CD011129, doi:10.1002/14651858.cd011129.pub2.
3. Kaasa, S.; Loge, J.H.; Aapro, M. et al. Integration of Oncology and Palliative Care: A Lancet Oncology Commission. *Lancet Oncol* **2018**, *19*, e588–e653, doi:10.1016/s1470-2045(18)30415-7.
4. German Guideline Program in Oncology (German Cancer Society, G.C.A., AWMF): Palliative Care for Patients with Incurable Cancer, Extended Version—Short Version 2.2, 2020 AWMF-Registration Number 128/001OL. Available online: <https://leitlinienprogramm-onkologie.de/Leitlinien.7.0.html> (accessed on 11 August 2023).
5. Hui, D.; Mori, M.; Watanabe, S.M. et al. Referral Criteria for Outpatient Specialty Palliative Cancer Care: An International Consensus. *Lancet Oncol* **2016**, *17*, e552–e559, doi:10.1016/s1470-2045(16)30577-0.
6. Hui, D.; Hannon, B.L.; Zimmermann, C. et al. Improving Patient and Caregiver Outcomes in Oncology: Team-based, Timely, and Targeted Palliative Care. *CA: A Cancer J. Clin.* **2018**, *68*, 356–376, doi:10.3322/caac.21490.
7. Coventry, PA; Grande, GE; Rucgards, DA et al. Prediction of Appropriate Timing of Palliative Care for Older Adults with Non-Malignant Life-Threatening Disease: A Systematic Review. *Age Ageing* **2005**, *34*(3), 218–27., doi:10.1093/ageing/afi054.
8. Hui, D.; Hannon, B.L.; Zimmermann, C. et al. Improving Patient and Caregiver Outcomes in Oncology: Team-based, Timely, and Targeted Palliative Care. *Ca Cancer J Clin* **2018**, *68*, 356–376, doi:10.3322/caac.21490.
9. Hui, D.; Heung, Y.; Bruera, E. Timely Palliative Care: Personalizing the Process of Referral. *Cancers* **2022**, *14*, 1047, doi:10.3390/cancers14041047.
10. Hui, D; Mori, M; Meng, YC et al. Automatic Referral to Standardize Palliative Care Access: An International Delphi Survey. *Supportive Care in Cancer* **2017**, doi:10.1007/s00520-017-3830-5.
11. Basch, E.; Mody, G.N.; Dueck, A.C. Electronic Patient-Reported Outcomes as Digital Therapeutics to Improve Cancer Outcomes. *Jco Oncol Pract* **2020**, *16*, 541–542, doi:10.1200/op.20.00264.
12. Basch, E.; Deal, A.M.; Dueck, A.C. et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. *JAMA* **2017**, *318*, 197, doi:10.1001/jama.2017.7156.
13. Denis, F.; Basch, E.; Septans, A.L. et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. *JAMA* **2019**, *321*.
14. Basch, E.; Deal, A.M.; Kris, M.G. et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* **2016**, *34*, 557–565, doi:10.1200/jco.2015.63.0830.
15. Barbera, L.; Sutradhar, R.; Seow, H. et al. Impact of Standardized Edmonton Symptom Assessment System Use on Emergency Department Visits and Hospitalization: Results of a Population-Based Retrospective Matched Cohort Analysis. *JCO Oncol. Pract.* **2020**, *16*, e958–e965, doi:10.1200/jop.19.00660.
16. Basch, E.; Wilfong, L.; Schrag, D. et al. Adding Patient-Reported Outcomes to Medicare’s Oncology Value-Based Payment Model. *JAMA* **2020**, doi:10.1001/jama.2019.19970.
17. Killock, D. AI Outperforms Radiologists in Mammographic Screening. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 134–134, doi:10.1038/s41571-020-0329-7.
18. Esteva, A.; Kuprel, B.; Novoa, R. et al. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature*, doi:10.1038/nature21056.
19. Guo, W.; Gao, G.; Dai, J. et al Prediction of Lung Infection during Palliative Chemotherapy of Lung Cancer Based on Artificial Neural Network. *Comput. Math. Methods Medicine* **2022**, *2022*, 4312117, doi:10.1155/2022/4312117.

20. Van Helden, E.J.; Vacher, Y.J.L.; Wieringen, W.N. et al. Radiomics Analysis of Pre-Treatment [18F]FDG PET/CT for Patients with Metastatic Colorectal Cancer Undergoing Palliative Systemic Treatment. *Eur. J. Nucl. Medicine Mol. Imaging* **2018**, *45*, 2307–2317, doi:10.1007/s00259-018-4100-6.
21. Vu, E.; Steinmann, N.; Schröder, C. et al. Applications of Machine Learning in Palliative Care: A Systematic Review. *Cancers* **2023**, *15*, 1596, doi:10.3390/cancers15051596.
22. Avati, A.; Jung, K.; Harman, S. et al. Improving Palliative Care with Deep Learning. *BMC Méd. Informatics Decis. Mak.* **2018**, *18*, 122, doi:10.1186/s12911-018-0677-8.
23. Cary, M.P.; Zhuang, F.; Draelos, R.L. et al. Machine Learning Algorithms to Predict Mortality and Allocate Palliative Care for Older Patients With Hip Fracture. *Journal of the American Medical Directors Association* **2021**, Pages 291–296, doi:10.1016/j.jamda.2020.09.025.
24. Manz, C.R.; Parikh, R.B.; Small, D.S. et al. Effect of Integrating Machine Learning Mortality Estimates With Behavioral Nudges to Clinicians on Serious Illness Conversations Among Patients With Cancer. *Jama Oncol* **2020**, *6*, e204759, doi:10.1001/jamaoncol.2020.4759.
25. Parikh, R.B.; Manz, C.; Chivers, C. et al. Machine Learning Approaches to Predict 6-Month Mortality Among Patients With Cancer. *JAMA Netw. Open* **2019**, *2*, e1915997, doi:10.1001/jamanetworkopen.2019.15997.
26. Hui, D.; Paiva, C.E.; del Fabbro, E.G. et al. Prognostication in Advanced Cancer: Update and Directions for Future Research. *Supportive Care in Cancer* **2019**, *27*, 1973–1984.
27. Mitra, T.; Rettler, T.M.; Beckmann, M. et al. Patient-Reported-Outcome-Messung (PROM) Psychosozialer Belastung Und Symptome Für Ambulante Patienten Unter Kurativer Oder Palliativer Tumorthherapie – Eine Retrospektive Analyse Eines Comprehensive Cancer Center. *Der Onkologe* **2018**, *24*(1), 69–75, doi:10.1007/s00761-017-0324-5.
28. Kroenke, K.; Löwe, B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A Systematic Review. *General Hospital Psychiatry* **2010**, *32*, 345–359, doi:10.1016/j.genhosppsych.2010.03.006.
29. Plummer, F.; Manea, L.; Trepel, D. Screening for Anxiety Disorders with the GAD-7 and GAD-2: A Systematic Review and Diagnostic Metaanalysis. *General Hospital Psychiatry*, doi:10.1016/j.genhosppsych.2015.11.005.
30. Stiel, S.; Bertram, M.E.M.L.; Ostgathe, C. et al. Validation of the New Version of the Minimal Documentation System (MIDOS) for Patients in Palliative Care: The German Version of the Edmonton Symptom Assessment Scale (ESAS)]. *Schmerz*, doi:10.1007/s00482-010-0972-5.
31. Buchhold, B.; Lutze, S.; Freyer-Adam, J. et al. Validation of the Psychometric Properties of a “Modified Version of the Hornheider Screening Instrument” (HSI-MV) Using a Sample of Outpatient and Inpatient Skin Tumor Patients. *JDDG: J. Dtsch. Dermatol. Ges.* **2022**, *20*, 597–609, doi:10.1111/ddg.14692.
32. Ownby, K.K. Use of the Distress Thermometer in Clinical Practice. *J. Adv. Pr. Oncol.* **2019**, *10*, doi:10.6004/jadpro.2019.10.2.7.
33. Storick, V.; O’Herlihy, A.; Abdelhafeez, S. et al. Improving Palliative and End-of-Life Care with Machine Learning and Routine Data: A Rapid Review. *HRB Open Res.* **2019**, *2*, 13, doi:10.12688/hrbopenres.12923.1.
34. Wang, L.; Sha, L.; Lakin, J.R. et al. Development and Validation of a Deep Learning Algorithm for Mortality Prediction in Selecting Patients With Dementia for Earlier Palliative Care Interventions. *JAMA Netw. Open* **2019**, *2*, e196972, doi:10.1001/jamanetworkopen.2019.6972.
35. Zhang, H.; Li, Y.; McConnell, W. Predicting Potential Palliative Care Beneficiaries for Health Plans: A Generalized Machine Learning Pipeline. *J. Biomed. Inform.* **2021**, *123*, 103922, doi:10.1016/j.jbi.2021.103922.
36. Berg, G.D.; Gurley, V.F. Development and Validation of 15-Month Mortality Prediction Models: A Retrospective Observational Comparison of Machine-Learning Techniques in a National Sample of Medicare Recipients. *BMJ Open* **2019**, *9*, e022935, doi:10.1136/bmjopen-2018-022935.
37. Blanes-Selva, V.; Ruiz-García, V.; Tortajada, S. et al. Design of 1-Year Mortality Forecast at Hospital Admission: A Machine Learning Approach. *Heal. Inform. J.* **2021**, *27*, 1460458220987580, doi:10.1177/1460458220987580.
38. Courtright, K.R.; Chivers, C.; Becker, M. et al. Electronic Health Record Mortality Prediction Model for Targeted Palliative Care Among Hospitalized Medical Patients: A Pilot Quasi-Experimental Study. *J. Gen. Intern. Med.* **2019**, *34*, 1841–1847, doi:10.1007/s11606-019-05169-2.
39. Chu, C.; White, N.; Stone, P. Prognostication in Palliative Care. *Clin. Med.* **2019**, *19*, 306–310, doi:10.7861/clinmedicine.19-4-306.
40. Lee, E.S.; Hiratsuk, Y.; Suh, S.Y. Clinicians’ Prediction of Survival and Prognostic Confidence in Patients with Advanced Cancer in Three East Asian Countries. *Journal of Palliative Medicine* **2023**, doi:10.1089/jpm.2022.0380.
41. Parajuli, J.; Hupcey, J.E. A Systematic Review on Barriers to Palliative Care in Oncology. *American Journal of Hospice and Palliative Medicine* **2021**, 1361–1377, doi:10.1177/1049909120983283.
42. Gothe, H.; Brinkmann, C.; Schmedt, N. Is There an Unmet Medical Need for Palliative Care Services in Germany? Incidence, Prevalence, and 1-Year All-Cause Mortality of Palliative Care Sensitive Conditions:

- Real-World Evidence Based on German Claims Data. *J. Public Heal.* **2022**, *30*, 711–720, doi:10.1007/s10389-020-01319-7.
43. Windisch, P.; Hertler, C.; Blum, D. Leveraging Advances in Artificial Intelligence to Improve the Quality and Timing of Palliative Care. *Cancers* **2020**, *12*, 1149, doi:10.3390/cancers12051149.
 44. Carrasco-Zafra, M.I.; Gómez-García, R.; Ocaña-Riola, R. et al. Level of Palliative Care Complexity in Advanced Cancer Patients: A Multinomial Logistic Analysis. *J Clin Medicine* **2020**, *9*, 1960, doi:10.3390/jcm9061960.
 45. Tewes, M.; Rettler, T.; Wolf, N. et al. Predictors of Outpatients' Request for Palliative Care Service at a Medical Oncology Clinic of a German Comprehensive Cancer Center. *Support Care Cancer* **2018**, *26*, 3641–3647, doi:10.1007/s00520-018-4245-7.
 46. Vogt, J.; Beyer, F.; Sistermanns, J. et al. Symptom Burden and Palliative Care Needs of Patients with Incurable Cancer at Diagnosis and During the Disease Course. *Oncol* **2021**, *26*, e1058–e1065, doi:10.1002/onco.13751.
 47. Lyu, X.; Luo, B. Prognostic Factors and Survival Prediction in HER2-positive Breast Cancer with Bone Metastases: A Retrospective Cohort Study. *Cancer Med.* **2021**, *10*, 8114–8126, doi:10.1002/cam4.4326.
 48. Greenlee, H.; Unger, J.M.; LeBlanc, M. et al. Association between Body Mass Index and Cancer Survival in a Pooled Analysis of 22 Clinical Trials. *Cancer Epidemiology Prev. Biomark.* **2017**, *26*, 21–29, doi:10.1158/1055-9965.epi-15-1336.
 49. Tripodoro, V.A.; Llanos, V.; Daud, M.L. et al. Palliative and Prognostic Approach in Cancer Patients Identified in the Multicentre NECesidades PALiativas 2 Study in Argentina. *ecancermedicalscience* **2021**, *15*, 1316, doi:10.3332/ecancer.2021.1316.
 50. Ediebah, D.E.; Quinten, C.; Coens, C. et al. Quality of Life as a Prognostic Indicator of Survival: A Pooled Analysis of Individual Patient Data from Canadian Cancer Trials Group Clinical Trials. *Cancer* **2018**, *124*, 3409–3416, doi:10.1002/cncr.31556.
 51. Qian, A.S.; Qiao, E.M.; Nalawade, V. et al. Impact of Underlying Malignancy on Emergency Department Utilization and Outcomes. *Cancer Med.* **2021**, *10*, 9129–9138, doi:10.1002/cam4.4414.
 52. Mushtaq, R.R.; Yekkaluri, S.; Meyer, E. et al. Emergency Room Presentation of Lung Cancer as a Highly Practical Prognostic Marker for Consideration in Clinical Trials: A SEER Database Analysis.; **2020**.
 53. Owusuaa, C.; Dijkland, S.A.; Nieboer, D. et al. Predictors of Mortality in Patients with Advanced Cancer—A Systematic Review and Meta-Analysis. *Cancers* **2022**, *14*, 328, doi:10.3390/cancers14020328.
 54. Garinet, S.; Wang, P.; Mansuet-Lupo, A. et al. Updated Prognostic Factors in Localized NSCLC. *Cancers* **2022**, *14*, 1400, doi:10.3390/cancers14061400.
 55. Ishikawa, E.; Sugii, N.; Matsuda, M. et al. Maximum Resection and Immunotherapy Improve Glioblastoma Patient Survival: A Retrospective Single-Institution Prognostic Analysis. *BMC Neurol.* **2021**, *21*, 282, doi:10.1186/s12883-021-02318-1.
 56. Parikh, R.B.; Manz, C.R.; Nelson, M.N. et al. Clinician Perspectives on Machine Learning Prognostic Algorithms in the Routine Care of Patients with Cancer: A Qualitative Study. *Support Care Cancer* **2022**, *30*, 4363–4372, doi:10.1007/s00520-021-06774-w.
 57. Spooner, C.; Vivat, B.; White, N. et al. What Outcomes Do Studies Use to Measure the Impact of Prognostication on People with Advanced Cancer? Findings from a Systematic Review of Quantitative and Qualitative Studies. *Palliat. Med.* **2023**, 2692163231191148, doi:10.1177/02692163231191148.
 58. Semler, S.C.; Wissing, F.; Heyder, R. German Medical Informatics Initiative. *Methods Inf. Med.* **2018**, *57*, e50–e56, doi:10.3414/me18-03-0003.
 59. Gaertner, J.; Hallek, J.W.; Hallek, M. et al. Standardizing Integration of Palliative Care into Comprehensive Cancer Therapy—a Disease Specific Approach. *Supportive Care in Cancer volume* **2011**, *19*, 1037–1043, doi:10.1007/s00520-011-1131-y.
 60. Chen, S.; Li, R.; Zhang, Z. et al. Prognostic Value of Baseline and Change in Neutrophil-to-Lymphocyte Ratio for Survival in Advanced Non-Small Cell Lung Cancer Patients with Poor Performance Status Receiving PD-1 Inhibitors. *Transl. Lung Cancer Res.* **2021**, *10*, 1397–1407, doi:10.21037/tlcr-21-43.
 61. Berlanga, M.; Cupp, I.; Tyoulaki, E. et al. Neutrophil to Lymphocyte Ratio and Cancer Prognosis: An Umbrella Review of Systematic Reviews and Meta-Analyses of Observational Studies. *BMC Medicine* **2020**, *18*, doi:10.1186/s12916-020-01817-1.
 62. Amano, K.; Maeda, I.; Morita, T. et al. Clinical Implications of C-Reactive Protein as a Prognostic Marker in Advanced Cancer Patients in Palliative Care Settings. *J Pain Symptom Manag* **2016**, *51*, 860–867, doi:10.1016/j.jpainsymman.2015.11.025.
 63. SEER*Explorer: An Interactive Website for SEER Cancer Statistics [Internet]. Surveillance Research Program, National Cancer Institute. Available from: https://seer.cancer.gov/report_to_nation/ (accessed 10 Sept 2023).
 64. Sandham, M.H.; Hedgecock, E.A.; Siegert, R.J. et al. Intelligent Palliative Care Based on Patient-Reported Outcome Measures. *J. Pain Symptom Manag.* **2022**, *63*, 747–757, doi:10.1016/j.jpainsymman.2021.11.008.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.