

# Palliative Performance Scale: Polish validation in hospice setting - a prospective observational study

Tomasz Dzierżanowski

Warszawski Uniwersytet Medyczny

Tomasz Gradalski (✉ [tomgr@mp.pl](mailto:tomgr@mp.pl))

St Lazarus Hospice <https://orcid.org/0000-0003-3705-3846>

Michael Kozłowski

Uniwersytet Jagiellonski w Krakowie

---

## Research article

**Keywords:** performance status, scale, validation, hospice, palliative

**Posted Date:** February 18th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.23857/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Palliative Care on April 22nd, 2020.  
See the published version at <https://doi.org/10.1186/s12904-020-00563-8>.

# Abstract

**Background:** Measuring functional status in palliative care may help clinicians to assess a patient's prognosis, recommend adequate therapy, avoid futile or aggressive medical care, consider hospice referral, and evaluate provided rehabilitation outcomes. An optimized, widely used, and validated tool is preferable. The Palliative Performance Scale Version 2 (PPSv2) is currently one of the most commonly used performance scales in palliative settings. The aim of this study is the translation and validation process of a Polish translation of this tool (PPSv2-Polish).

**Methods:** Two hundred patients consecutively admitted to a free-standing hospice were evaluated twice during 2 consecutive days for test-retest reliability. In the first evaluation, two different care providers independently evaluated the same patient to establish inter-rater reliability values. PPS-Polish was compared with the Karnofsky Performance Score (KPS), Eastern Cooperative Oncology Group (ECOG) Performance Status (ECOG PS), and Barthel Activities of Daily Living (ADL) Index to determine its construct validity.

**Results:** A high level of full agreement between test and retest was seen (63%), and a good intra-class correlation coefficient of 0.85 ( $P < 0.0001$ ) was achieved. Excellent agreement between raters was observed when using PPSv2-Polish (Cohen's kappa 0.91;  $P < 0.0001$ ). Satisfactory correlations with the KPS and good correlations with ECOG PS and Barthel ADL were noticed. Persons who had shorter prognoses and were predominantly bedridden also had lower scores measured by the PPSv2-Polish, KPS and Barthel ADL. A strong correlation of 0.77 between PPSv2-Polish scores and survival time was noted ( $P < 0.0001$ ). Moderate survival correlations were seen between KPS, ECOG PS, and Barthel ADL of 0.41; -0.62; and 0.58, respectively ( $P < 0.0001$ ).

**Conclusion:** PPSv2-Polish is a valid and reliable tool measuring performance status in a hospice population and can be used in daily clinical practice in palliative care and research.

## Background

Functional decline, in addition to symptom burden, results in increased dependency on others and negatively affects a patient's quality of life. Most people want to remain symptom-free and to maintain as much independence as possible until the latest phases of life. The importance of measuring functional status in palliative care is incontestable. It may help clinicians to assess a patient's prognosis, recommend further oncological therapy, avoid futile, aggressive medical care, consider hospice referral and evaluate provided rehabilitation outcomes. All measuring tools may have disadvantages, and an optimal one, which may also be used in monitoring the outcomes of physical therapy, specifically in palliative care, is lacking [1]. The Palliative Performance Scale version 2 (PPSv2) is currently one of the most commonly used performance scales in palliative settings. It was developed as a modification of the Karnofsky Performance Scale in 1996 [2] and validated later [3]. It has also been demonstrated that PPSv2 correlates with patients' survival time [4]. Although occasionally used in Poland, it has not been

validated yet. It is reasonable to disseminate knowledge among professionals and promote its use through cultural adaptation.

## Methods

### Aim and design

The aim of this study is the translation and validation process of a Polish translation of PPSv2 (PPSv2-Polish). This is a prospective observational study of cancer patients performed by the hospice care team at St Lazarus Hospice, Krakow, Poland. All participants were evaluated twice during two consecutive days for test-retest reliability by the same palliative team member (a trained and experienced palliative care nurse, psychologist, or physiotherapist), who cared for the patient. Additionally, on the first evaluation, two different care providers (each time by different types of professionals) independently evaluated the same patient to establish inter-rater reliability values. Each patient's evaluation encompassed the PPSv2-Polish, which was compared with 3 additional performance scales to accomplish its construct validity.

### Participants

Two hundred adult patients (aged  $\geq 18$  years) who were consecutively admitted to an in-patient hospice, Polish native speakers, clinically stable according to the attending physician, able to communicate and complete the questionnaires, and without cognitive impairment, were recruited and enrolled in the study. The sample size for this survey was based on general guidelines for conducting qualitative research [5].

### Measures

#### Palliative Performance Scale version 2 (PPSv2) and the Polish version of the tool

The PPS provides a functional assessment of a patient's ambulation, activity level, evidence of disease, self-care, food/fluids intake, and level of consciousness. The PPS has 11 categories, from 100% (full mobile and healthy) to 0% (dead) in increments of 10%. In 2006, PPS version 2 (PPSv2) was introduced after clarification of instructions for its use [6].

A modified combined translation technique [7] of PPSv2 into Polish was applied, which consisted of 1) independent forward translations by a physician, a psychologist and a Polish native speaker, 2) team discussion on identified differences between these 3 versions until agreement, 3) independent backward translations by a physician, a psychologist and a native English speaker, and 4) second team discussion on any differences between the original and back-translated versions until all agreed that the two versions were semantically identical. This method was consistent with the Victoria Hospice translation guidelines sent to the authors. The discussed back-translated version was then preliminarily tested on patients to obtain the final version of the PPSv2 (PPSv2-Polish – see Additional file 1). The implementation process encompassed education of the medical staff participating in this study during one training session, based on the Victoria Hospice Society instructions, giving the opportunity to get feedback from the team members.

## Karnofsky Performance Score (KPS)

The KPS ranking is an 11-point scale and runs from 100% - perfect health, to 0% - dead. While first published in 1948 [8] to evaluate the ability to survive chemotherapy for cancer, it has recently undergone several evaluation adjustments [9]. The KPS provides great consistency of ratings by different oncology professionals [10]. It may also serve as a life survival predictor [11].

## Eastern Cooperative Oncology Group (ECOG) Performance Status (ECOG PS)

This scale, also called the World Health Organization (WHO) score, was published in its current form by ECOG in 1982 [12] to assess a patient's level of functioning in terms of the ability to care for himself, daily activity, and physical ability, in order to measure the impact of the disease/treatment on performance status. It has a good intra and interobserver agreement in cancer patients' performance status assessment [13]. It consists of 6 categories, from 0 - fully active, to 5 - dead, and is simpler to use, may be clinically preferable in comparison to the KPS, [14] and is widely used in the literature.

## Barthel Activities of Daily Living (ADL) Index

This "simple index of independence" was published in 1965 for measuring the improvement during rehabilitation of the chronically ill" [15]. The original version was modified in 1988 to a 20 point scale that measures in increments of 1 point: from 0 - fully dependent, to 20 - fully independent [16]. The final score can be multiplied by 5 to obtain a 100 point score, and it is proposed that scores of 0–20 indicate "total" dependency, 21–60 indicate "severe" dependency, 61–90 indicate "moderate" dependency, and 91–99 indicates "slight" dependency. It is already widely used as the measurement of daily living activities and has become a standard measure of physical disability in practice [17]. Ten categories are assessed: feeding, grooming, bathing, dressing, toilet use, presence of fecal incontinence, presence of urinary incontinence, transfers (e.g., moving from wheelchair to bed), walking on an even surface (or propelling a wheelchair if unable to walk), and ascending and descending stairs.

## Statistical analysis

We summarized the baseline demographics using descriptive statistics and medians with interquartile ranges (IQR) in non-normally distributed data. A Wilcoxon signed-rank test was used to compare test-retest of ordinal data in one sample in test-retest and inter-rater comparison. Non-parametric data within subgroups of patients were compared using a Mann-Whitney U test. The strength of the relationship between the test-retest variables, and between the tools scores and survival time were calculated with the Spearman's rank correlation coefficient. The inter-rater reliability was estimated using Cohen's kappa statistics. Data were analyzed using STATISTICA 13.0 (TIBCO Software Inc. 2017) data analysis software. A P-value of < 0.05 was considered as the level for statistical significance. As there are no absolute rules for the sample size needed to validate a questionnaire a fair size of 200 patients was chosen [5].

# Results

According to the hospice staff involved in this study, the PPSv2-Polish was a feasible and acceptable assessment tool in their practice. At the beginning of implementation, in the training phase of this study, two out of six team members made observations that the 5-steps of PPSv2-Polish assessment was a bit prolonged, in comparison with 1-step assessment of KPS or ECOG PS. All participating staff emphasized the need for observation of a patient for a reasonable period of time to accurately assess his or her true "capable" functions based on the "observed" ones, during the day shift.

All the recruited subjects were able to complete the evaluation according to the study protocol (no missing values, a response rate of 100%). The majority of patients were aged, had advanced cancer, with short (weeks) prognosis according to Gold Standard Framework (GSF), and finally died at the hospice (Table 1).

Table 1  
Patient characteristics and description

Parameter	n = 200	%
Median age in years (IQR)	72.5	(16.0)
Females	144	72.0
Primary cancer tumor site		
Digestive	34	17.0
Respiratory	35	17.5
Genitourinary	22	11.0
Breast	21	10.5
Others	67	33.5
Nonmalignant diseases	21	10.5
Stage/prognosis (GSF)		
Stable/months (B)	32	16.0
Progressing/weeks (C)	141	70.5
Last days (D)	27	13.5
ECOG PS		
1	3	1.5
2	23	11.5
3	41	20.5
4	133	66.5
Died		
In 1–30 days	117	58.5
In > 30 days	58	29.0
Discharged	25	12.5

#### Test-retest reliability

The median PPSv2-Polish value within the first measurement was 30 (IQR 10), which correlated with the data obtained two days later (median 30; IQR 20) by the same care provider. We achieved a high level of full agreement between test and retest (63%) and a good intra-class correlation coefficient of 0.85 ( $P < 0.0001$ ).

## Inter-rater reliability

An excellent PPSv2-Polish agreement between raters was observed (Cohen's kappa 0.91;  $P < 0.0001$ ). A high level (94–99%) of full agreement between raters was observed, with the exception of ECOG PS, where this agreement was not achieved (Table 2).

Table 2  
Inter-rater agreement of scales used in the study.

Scales	Rater #1		Rater #2		% full agreement	P
	Median	IQR	Median	IQR		
PPSv2-Polish	30.0	10.0	30.0	10.0	94.0	0.13
KPS	45.0	10.0	40.0	10.0	99.0	0.18
ECOG PS	4.0	1.0	4.0	10.0	95.5	0.54
Barthel ADL	5.0	22.5	5.0	22.5	83.0	0.001
P - Wilcoxon signed-rank test						

## Criterion validity

Satisfactory correlations with the KPS and good correlation with ECOG PS and Barthel ADL were noticed (Table 3).

Table 3  
Correlations between the scales used in the study

	KPS	ECOG	Barthel ADL
PPSv2-Polish	0.69*	-0.81*	0.75*
KPS		-0.57*	0.68*
ECOG PS			-0.73*
* Spearman's Rho $P < 0.001$			

## Known-group validity

Persons who had shorter prognosis and were predominantly bedridden also had lower scores measured by the PPSv2-Polish, KPS and Barthel ADL (Table 4).

Table 4  
Responsiveness of the scales.

Scales	GSF A-B n = 32		GSF C-D n = 168		P	ECOG 1-2 n = 26		ECOG 3-4 n = 174		P
	Median	IQR	Median	IQR		Median	IQR	Median	IQR	
PPSv2-Polish	50.0	15.0	30.0	10.0	*	60.0	20.0	30.0	10.0	*
KPS	60.0	15.0	40.0	10.0	*	60.0	10.0	40.0	10.0	*
Barthel ADL	42.5	40.0	5.0	7.5	*	45.0	60.0	5.0	10.0	*
Median scores of each scale obtained within the known subgroups of longer (GSF A-B) and shorter (GSF C-D) prognosis and also within higher (ECOG PS 1–2) and lower (ECOG PS 3–4) performance status * P < 0,00001, Mann-Whitney U test										

### Survival estimation

A strong correlation between PPSv2-Polish scores and survival time was noted (0.77; P < 0.0001). Moderate survival correlations were observed between KPS, ECOG PS, and Barthel ADL Index scores (0.41, -0.62, and 0.58, respectively; P < 0.0001).

## Discussion

The PPSv2-Polish which was created with a combined translation technique appears to be a valuable clinical assessment tool in the hospice setting within the Polish population of cancer patients. The implementation process and training of experienced palliative care medical staff proceeded without any particular difficulties. The team members reported that the PPSv2-Polish tool was clear and easy, although a bit time consuming to use in daily practice, and it required observation of a patient for a significant period of time (e.g. through a whole shift) to assess the potential capability of evaluated functions. When comparing various performance tools, it appears that no one is statistically superior to others in terms of inter-rater reliability [18]. The ECOG PS or KPS are often used in determining eligibility for clinical trials. However, there could be a substantial disagreement in the assessment of performance status between oncologists, even when using as simple a tool as the ECOG PS [19]. Patients' age, preferences and socio-economic background may also influence the assessment. Numerous studies confirmed good correlations of KPS, ECOG PS, and PPS tools, and meta-analysis favors KPS as descriptively better [18]. Authors of the original version of the PPSv2 noticed that in contrary to the KPS, this scale does not focus on the need for hospitalization (which is of poor definability and does not help in defining performance) but instead assesses food/fluid intake and level of consciousness [2]. The ambiguity of the KPS assessment when patients were bedridden (KPS ≤ 40%) led to scale modification in Australia [20]. PPSv2 usage assumes a 5-step assessment, which can be problematic at the beginning, but after training may be more comprehensive and accurate. Both KPS and PPSv2 scales need standardized, appropriate instructions regarding performance evaluation. Compared with the 5-point

ECOG PS, the 11-point PPSv2 seems to be a more precise tool, especially for lower performance statuses. This phenomenon was also affirmed in our observations, where the ECOG PS did not achieve significant reliability between the scoring of different types of professionals.

The high level of agreement with a very good correlation in serial scoring by one rater in a two days interval appeared better than in another study with two weeks between consecutive assessments [3]. The chosen interval of 2 days between assessments allowed for an optimal period not to recall the first scoring, yet not too long to allow changes in performance in most cases.

The excellent inter-rater concordance observed in this study was higher than in an updated meta-analysis recently published [18]. The explanation of this phenomenon partially could be explained by the careful inclusion of rating staff, who attended to patients for several hours daily and had a great understanding of their mobility and functionality [10].

The strength of this study was the high inter-rater agreement each time between different professionals (nurse, psychologist, or physiotherapist) took part in the assessment. There are inconsistencies in the evaluation of a patient's performance status using the same tool by different types of professionals (doctors rated patients as healthier than nurses using the PPSv2 scale), and it may be explained by different amounts of time spent with the evaluated person [10]. However, even research assistants rated patients similar to physicians (oncologists or radiation therapists) in one study [21]. Optimally, it should be advised to score patients within interdisciplinary team meetings to gain a more accurate assessment [22].

Our study legitimizes the usage of the PPSv2-Polish in prognostication of patients with advanced cancer, which is in line with previous studies [4][6]. The strongest correlation between the PPSv2-Polish and survival time among the analysed tools was remarkable, and as to the best of our knowledge it was not published earlier. In another recent study of advanced cancer patients with prognosis in terms of weeks, the PPSv2 assessment was as accurate as subjective clinical survival predictions [23]. The similarity between the subjective prognosis assessment of the attending physician, expressed by GSF staging, and the PPSv2 scoring in our observations legitimizes the good responsiveness of this newly validated tool to patients' changing prognosis and physical condition. This finding could be explained by the strong relationship between the hospice staff and the patient, as this factor was described to have an impact on accuracy [24].

This study was not without limitations. First, the majority of patients recruited presented a low performance status, were mostly sitting or bedridden, and were not representative for the whole palliative population. Secondly, only in-patient subjects were recruited and most of the unstable patients were excluded due to the test-retest requirements.

## Conclusion

We believe that the Polish version of PPSv2 is a valid and reliable tool measuring performance status in a hospice population, which can be used in daily clinical practice in palliative care, research, and prognostication.

## Abbreviations

PPS

Palliative Performance Scale

PPSv2

Palliative Performance Scale Version 2

PPSv2-Polish

Polish translation of the Palliative Performance Scale Version 2

KPS

Karnofsky Performance Score

ECOG

Eastern Cooperative Oncology Group

ECOG PS

Eastern Cooperative Oncology Group Performance Status

Barthel ADL

Barthel Activities of Daily Living Index

IQR

interquartile range

GSF

Gold Standard Framework

## Declarations

Availability of data and materials

The dataset used and analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The study was approved by the Bioethics Committee of the Medical University of Warsaw (AKBE/309/2019). At the start of each patient evaluation, all hospice care team members confirmed verbally their consent to the researcher.

Consent for publication

Not applicable

## Competing interests

The authors declare that they have no conflicts of interest.

## Authors' contribution

The authors: T. G. and T. D. contributed equally to this manuscript. Conceptualization: TG, TD; Methodology TG, TD; Tool translation TG, TD, MK; Data collection: TG; Formal analysis: TD, TG; Writing - review and editing: TG, TD, MK

## References

1. Looney F, Cobbe S, Ryan A, Barriscale I, McMahon A, Real S. The Search for a Functional Outcome Measure for Physical Therapy in Specialist Palliative Care: An Ongoing Journey. *Rehabil. Oncol.* 2020. doi:10.1097/01.REO.00000000000000194.
2. Anderson F, Downing GM, Hill J, Casorso L, Lerch N. Palliative Performance Scale (PPS): A New Tool. *J. Palliat. Care* 1996; 12:5–11.
3. Ho F, Lau F, Downing MG, Lesperance M. A reliability and validity study of the Palliative Performance Scale. *BMC Palliat. Care* 2008; 7:10.
4. Myers J, Kim A, Flanagan J, Selby D. Palliative performance scale and survival among outpatients with advanced cancer. *Support. Care Cancer* 2015; 23:913–918.
5. Shamin T. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J. Anesth.* 2019; 13:281.
6. Lau F, Downing GM, Lesperance M, Shaw J, Kuziemy C. Use of Palliative Performance Scale in End-of-Life Prognostication. *J. Palliat. Med.* 2006; 9:1066–1075.
7. Cha ES, Kim KH, Erlen JA. Translation of scales in cross-cultural research: Issues and techniques. *J. Adv. Nurs.* 2007; 58:386–395.
8. Karnofsky D a, Abelmann WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma. *Cancer* 1948; 1:634–656.
9. Péus D, Newcomb N, Hofer S. Appraisal of the Karnofsky Performance Status and proposal of a simple algorithmic system for its evaluation. *BMC Med. Inform. Decis. Mak.* 2013; 13:1.
10. Zimmermann C, Burman D, Bandukwala S, Seccareccia D, Kaya E, Bryson J, Rodin G, Lo C. Nurse and physician inter-rater agreement of three performance status measures in palliative care outpatients. *Support. Care Cancer* 2010; 18:609–616.
11. Jang RW, Caraiscos VB, Swami N, Banerjee S, Mak E, Kaya E, Rodin G, Bryson J, Ridley JZ, Le LW, Zimmermann C. Simple Prognostic Model for Patients With Advanced Cancer Based on Performance Status. *J. Oncol. Pract.* 2014; 10:e335–e341.
12. Oken, Martin; Creech, Richard, Tormey, Douglass; Horton, John; Davis, Thomas; McFadden, Eleanor; Carbone P. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin*

- Oncol 1982:649–656.
13. Roila F, Lupattelli M, Sassi M, Basurto C, Bracarda S, Picciafuoco M, Boschetti E, Milella G, Ballatori E, Tonato M, Favero A Del. Intra and interobserver variability in cancer patients' performance status assessed according to Karnofsky and ECOG scales. *Ann. Oncol.* 1991; 2:437–439.
  14. Taylor AE, Olver IN, Sivanthan T, Chi M, Purnell C. Observer error in grading performance status in cancer patients. *Support. Care Cancer* 1999; 7:332–335.
  15. Mahoney , Florence; W Barthel D. Functional evaluation: the Barthel Index. *Md. State Med. J.* 1965; 14:56–61.
  16. Collin C, Wade DT, Davies S, Horne V. The barthel ADL index: A reliability study. *Disabil. Rehabil.* 1988; 10:61–63.
  17. Wade, DT; Collin C. The Barthel ADL Index: A standard measure of physical disability. *Int Disabil Stud.* 1988; 10:64–67.
  18. Chow R, Zimmermann C, Bruera E, Temel J, Im J, Lock M. Inter-rater reliability in performance status assessment between clinicians and patients: A systematic review and meta-analysis. *BMJ Support. Palliat. Care* 2019. doi:10.1136/bmjspcare-2019-002080.
  19. Datta SS, Ghosal N, Daruvala R, Chakraborty S, Shrimali RK, Van Zanten C, Parry J, Agrawal S, Atreya S, Sinha S, Chatterjee S, Gollins S. How do clinicians rate patient's performance status using the ECOG performance scale? A mixed-methods exploration of variability in decision-making in oncology. *Ecancermedicalscience* 2019; 13:1–23.
  20. Abernethy AP, Shelby-James T, Fazekas BS, Woods D, Currow DC. The Australia-modified Karnofsky Performance Status (AKPS) scale: A revised scale for contemporary palliative care clinical practice [ISRCTN81117481]. *BMC Palliat. Care* 2005; 4:1–12.
  21. Campos S, Zhang L, Sinclair E, Tsao M, Barnes EA, Danjoux C, Sahgal A, Goh P, Culleton S, Mitera G, Chow E. The palliative performance scale: Examining its inter-rater reliability in an outpatient palliative radiation oncology clinic. *Support. Care Cancer* 2009; 17:685–690.
  22. Kutner JS, Bryant LL, Beaty BL, Fairclough DL. Symptom Distress and Quality-of-Life Assessment at the End of Life: The Role of Proxy Response. *J. Pain Symptom Manage.* 2006; 32:300–310.
  23. Hui D, Ross J, Park M, Dev R, Vidal M, Liu D, Paiva CE, Bruera E. Predicting survival in patients with advanced cancer in the last weeks of life: How accurate are prognostic models compared to clinicians' estimates? *Palliat. Med.* 2019:026921631987326.
  24. Selby D, Chakraborty A, Lilien T, Stacey E, Zhang L, Myers J. Clinician accuracy when estimating survival duration: The role of the patient's performance status and time-based prognostic categories. *J. Pain Symptom Manage.* 2011; 42:578–588.

## Additional Files

Format: .pdf

Title/description: Polish translation of the Palliative Performance Scale Version 2

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [instrukcjaPPSv230.01.20.pdf](#)