Supportive Care in Cancer (2019) 27:3793–3798
https://doi.org/10.1007/s00520-018-4597-z

ORIGINAL ARTICLE

Objective assessment of WHO/ECOG performance status

Miha Sok[1,2] · Miha Zavrl[1] · Boris Greif[1] · Matevž Srpčič[1]

Received: 19 December 2017 / Accepted: 10 December 2018 / Published online: 5 February 2019

Abstract

Purpose Performance status is an important factor in determining quality of life, the choice of treatment, and prognostic tool in patients. All scoring systems currently in use measure the patient's performance subjectively. A new method of objective assessment of performance ECOG/WHO grades 2 and 3 was constructed and tested.

Methods A performance meter—an adapted USB data logger with a mercury tilt switch—was constructed. The device was tested in a feasibility study on 33 residents of a retirement home. Parallel to the objective assessment, each resident gave their own estimate of their performance, and each resident was in turn assessed by the nursing staff.

Results With the performance meter, 4 residents (12%) were assessed as PS ≥ 3 in comparison with 8 (24%) and 7 (21%) residents with an ECOG score ≥ 3 estimated by patients themselves and nursing staff respectively.

Conclusion Subjective scoring—estimated by patients themselves and by nursing staff—showed underestimation of patients' performance. In 12% of patients, a better performance score was observed with objective measurement in comparison with subjective assessment. Performance meter could be a useful tool for health care professionals for type of care decisions.

Keywords Performance status · ECOG score · Objective score · Performance meter

Introduction

Performance status (PS) is a reliable indicator for the patient's general condition. In clinical practice, it is an indispensable prognostic factor and useful tool in predictive models for survival and treatment decision [1–3]. A better PS suggests a better prognosis. PS is an important factor for determining the choice of treatment (surgery, radiation, chemotherapy) or the intensity of palliative treatment in cancer patients [4].

The Zubrod score, first published in 1982, has since been adapted into the nearly identical WHO and ECOG (Eastern Cooperative Oncology Group) scores. These scores are entirely subjective and depend on the experience and opinion of the rater. Different health care professionals score PS differently. The existing literature cites both good and bad inter-rater reliability [5].

✉ Miha Sok
miha.sok@mf.uni-lj.si

[1] Department of Thoracic Surgery, University Medical Centre Ljubljana, Zaloška cesta 7, 1000 Ljubljana, Slovenia

[2] Faculty of Medicine, University of Ljubljana, Vrazov trg 2, Ljubljana, Slovenia

Cancer patients with PS 0-1 are the best candidates for surgery, chemotherapy, and trials. It has been documented that the benefit of modern chemotherapy for advanced colorectal cancer applies not only to fit, young patients, but also to the elderly and patients with PS 2 [6]. Both the elderly and PS 2 patients are underrepresented in clinical trials and treatment strategy [7]. Thus, objective measurement is needed. Accurate measurement of patient physical activity besides other factors that are also taken into account when selecting the appropriate treatment would help oncologists select the most appropriate treatment options. Accelerometer-based monitoring and electronic activity monitoring (EAM) systems, measuring numbers of steps taken and/or the amount of time spent performing activities at different intensities, were tested in cancer patients with promising results [8–10]. Is it really possible to be objective in performance status assessment? The only statement we see that has potential to be measured is a statement in the ECOG system about staying in bed, where one of the descriptions of activity is based on the position of the body. There are distinctions between ECOG 2 and 3 on the basis of being recumbent or supine. Grade ECOG 2 marks the position of the body as being up and about more than 50% of waking hours, and grade ECOG 3 marks the state as confined to bed or chair 50% or more of waking hours (Table 1) [11].

Table 1 ECOG performance status

Grade	ECOG performance status
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours
3	Capable of only limited self-care; confined to bed or chair more than 50% of waking hours
4	Completely disabled; cannot carry on any self-care; totally confined to bed or chair
5	Dead

Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol.* 1982;5:649–655

There is a need to develop a new objective device that can numerically measure share of waking hours being up, citation that is also part of ECOG score. New objective device could replace subjective judgment of patient's activity and could descry between ECOG 2 and ECOG 3 score but not in Karnofsky score which has a higher number of choice but no numerical cut levels. The intention of the study is to test an objective tool for PS measurement on the basis of being recumbent or supine and compare these results to nursing staff and patients self -assessment of PS.

Besides objective measurement of patient physical activity, other factors are also taken into account when selecting the appropriate treatment that would help oncologists select the most appropriate treatment options.

Methods

Non-cancer patients were tested for practical reason. The idea was to find representative group of older subjects for testing feasibility, usability, and practicability of a new device for PS measurement. Retirement institution seemed to be appropriate, since residents and nursing staff measurement could be compared. Residents with cognitive impairment were excluded. This was a prospective, observational, feasibility study.

Performance status device

A Logger data performance status device (LDPS) prototype was developed. Electrical engineers were asked to construct a simple and cheap device for this purpose. They used USB data logger found on the market (Voltcraft DL-111K USB Temperature Data Logger), which normally measures temperature at regular intervals. A tilt switch mounted over the temperature sensor provided constant output regarding the horizontal and vertical positions of the instrument. Due to its simple technology in production process, there was no need for calibration between measurements. A pencil-like data logger was used, as fixation of the instrument is crucial. It was fixed to an underlying base with upper and lower clips. In general, a vertical position indicates the upright position of the individual, while a horizontal position of the instrument indicates supine position of the individual. After experimenting with the position of the device on various parts of the body (the leg, shirt pocket, taped to the torso), it was decided that the best part of the body was the presternal region, tie region, and mode of attachment was with a crocodile clip to a shirt at two points, top and bottom; thus, the device was held quite firmly in place with no shift because of torso movement (Fig. 1). The movements of the residents were constantly measured with sampling interval of 10 min, thus eliminating the problem of overabundance of data to process, and unnecessary short records of temporary excessive movements of the torso. This seemed reasonable as some brief activity, for example going to the toilet, cannot be judged as being up.

Positions between < 60 and $\geq 60°$ were tested as horizontal and vertical respectively, i.e., the supine and upright position of the individual torso. It is an arbitrary definition based upon observation and some preliminary measurement with LDPS in presternal region that normal daily body movement, walking, and sitting are characterized with the inclination of the device $\geq 60°$ and lying in a bed, head or back elevation, or anti-Trendelenburg position were registered as inclination of the device $< 60\%$.

Fig. 1 Active device in place. Note upper and bottom fixation

Participants

The study group of this feasibility study was chosen from a retirement institution. The reason for this decision was the possibility to have a uniform cohort of individuals with chronic illness and other age-related sickness which could best resemble patients affected by cancer and cancer treatment. At the beginning of the study, a closed group of permanent residents in one department with nursing staff was instructed at a group meeting about the idea of the study, PS characteristics, protocol of the study, and LDPS performances. ECOG PS 0–5 scale was used. The ECOG scale was developed by the Eastern Cooperative Oncology Group (ECOG) and published in 1982. It circulates in the public domain and is therefore available for public use (Table 1) [11]. ECOG PS 0 represents full activity, 1 represents restricted physically activity, 2 is up more than 50% of waking hours, 3 represents confined to bed or chair more than 50% of waking hours, 4 is totally confined to bed or chair, and 5 means dead. Residents and nursing staff were precisely informed about ECOG PS characteristics, and every score was discussed. They were asked to make self-judgment at the end of the monitoring time.

Residents were basically accepted at retirement home as ECOG ≤ 2. Although mutual assistance was observed between the residents, and eventually, some could be in worse condition. No one was hospitalized for at least 6 months prior to data of acquisition. Participation of the residents was voluntary, and informed consent was obtained at the beginning. All residents were ambulatory, no wheelchair users, and no resident in the study had the diagnosis of past or recent cancer or acute inflection. Alongside electronic measurement, their performance status was independently assessed by themselves, and by the nursing staff.

Residents were monitored only for 1 day. The matron of the department actively participated in the study with an obligation to permanently control protocol implementation.

Protocol

The matron of the department was responsible for the study and administration. In the morning, she fixed the LDPS to one resident with information to ensure that the device was fixed at the base and that turning of the device had to be reported to the nursing staff. Nursing staff assessed ECOG rate on the basis of the resident's self-care ability and activities during the day. Three ECOG scores were gathered for one resident in 1 day, one per 8-h nurse shift. The LDPS was picked up by the matron the next morning; it was restarted and fixed to another resident. No calibration was needed for new measurement. At the end of the study, the LDPS was delivered to the conductor of the study with the PS values of the resident and nursing staff recorded in the protocol note.

One LDPS was used. The recording interval was between 9 a.m. and 10 p.m. "Waking hours" was defined as the interval between 9 a.m. (waking of residents, distribution of breakfast and medication etc.) and 10 p.m. (the average bedtime of the residents), and the results were calculated accordingly. Data were uploaded to a computer via USB. The data for each resident are shown either graphically, indicating the time on the X-axis and body position of the instrument on the Y-axis, or as a set of numeric values in a table (Table 2).

Data were analyzed by hand, values TEMP = 0.0 were recognized as upright, active position and values TEMP > 0.0 as laying, resting position.

Table 2 Text data output (example). Different daily time is shown and differences in numbers under TEMP. TEMP number indicates position of the body (e.g., TEMP value 0.0 represent upright position, values > 0 represent supine position)

	DATE DP		TIME		TEMP	RH
1	01 Oct. 2008	09:54:00	−40.0	0.0	0.0	
2	01 Oct. 2008	10:04:00	−40.0	0.0	0.0	
3	01 Oct. 2008	10:14:00	−40.0	0.0	0.0	
4	01 Oct. 2008	10:24:00	−40.0	0.0	0.0	
5	01 Oct. 2008	10:34:00	−40.0	0.0	0.0	
6	01 Oct. 2008	10:44:00	−40.0	0.0	0.0	
7	01 Oct. 2008	10:54:00	−40.0	0.0	0.0	
8	01 Oct. 2008	11:04:00	−40.0	0.0	0.0	
31	01 Oct. 2008	14:54:00	−40.0	0.0	0.0	
32	01 Oct. 2008	15:04:00	−40.0	0.0	0.0	
33	01 Oct. 2008	15:14:00	−40.0	0.0	0.0	
34	01 Oct. 2008	15:24:00	31.6	40.2	16.4	
35	01 Oct. 2008	15:34:00	31.9	40.6	16.8	
36	01 Oct. 2008	15:44:00	32.0	40.3	16.8	
37	01 Oct. 2008	15:54:00	32.1	40.7	17.0	
38	01 Oct. 2008	16:04:00	32.2	39.7	16.7	
39	01 Oct. 2008	16:14:00	−40.0	0.0	0.0	
40	01 Oct. 2008	16:24:00	−40.0	0.0	0.0	
41	01 Oct. 2008	16:34:00	−40.0	0.0	0.0	
119	01 Oct. 2008	05:34:00	34.4	53.0	23.4	
120	01 Oct. 2008	05:44:00	34.3	53.4	23.4	
121	01 Oct. 2008	05:54:00	34.3	52.5	23.1	
122	01 Oct. 2008	06:04:00	−40.0	0.0	0.0	
123	01 Oct. 2008	06:14:00	−40.0	0.0	0.0	
124	01 Oct. 2008	06:24:00	−40.0	0.0	0.0	
125	01 Oct. 2008	06:34:00	−40.0	0.0	0.0	
126	01 Oct. 2008	06:44:00	−40.0	0.0	0.0	
127	01 Oct. 2008	06:54:00	−40.0	0.0	0.0	
128	01 Oct. 2008	07:04:00	−40.0	0.0	0.0	

Statistics

Data gathered were categorical variables. Agreement in PS assessment between three methods was calculated. Chi square test and Kappa test were used. Statistical data analysis was performed using SPSS for Windows version 19.0 (SPSS INC/IBM Corporation, Somers, NY, USA). P value of < 0.05 was used to define significance.

Results

The study was conducted from December 2014 to February 2015. Thirty-three mobile residents of a retirement home completed the monitoring protocol (26 female, 7 male), and their average age was 84.6 years (65–101 years). Common comorbidities were chronic heart disease ($n = 18$), chronic obstructive pulmonary disease ($n = 11$), diabetes mellitus ($n = 7$), rheumatic problems ($n = 7$), and others ($n = 3$). Residents with even mild signs of cognitive impairment were excluded from the study..

Adherence to the device was 100%. No change in position or movement was noticed during measurement time when fixed at two points, and participants did not have any complaints wearing it. The device is not in the market yet. PS assessment by three different methods is presented in Table 1. In general, 76% of residents estimated themselves to be in good performance status, e.g., PS ≤ 2. Nursing staff found 79% residents to have PS ≤ 2 while LDPS registered 88% residents to be PS ≤ 2. Residents and nursing staff also rated for PS 0, PS 1, and PS 2 (Table 3). The residents rated themselves better in comparison to nurses. Forty-four percent of residents rated themselves as PS 0 and PS 1 in comparison to 35% residents with PS 0 or PS 1 assessed by nurses. Two and three residents were classified as PS 0 by themselves and nursing staff respectively. There was no registration of PS 4 resident.

The LDPS registered four residents (12%) with ECOG ≥ 3 in comparison to 8 (24%) and 7 (21%) residents with self-estimation and nursing staff respectively. The difference is not statistically significant ($P = 0.42$) (Table 4).

Agreement of the three methods is presented in Table 5. The greatest agreement is between resident and nursing staff and minimal agreement is between resident and LDPS. The difference is not statistically significant ($P = 0.32$).

Table 3 PS assessment by three different methods

ECOG rate	0	1	2	3	4
Resident n	2	9	14	8	0
Nursing staff *n*	3	6	17	7	0
LDPS n	29			4	

Table 4 Results for each of the raters adapted for comparison with the objective assessment

Assessment by	ECOG score ≤ 2	ECOG score ≥ 3
Resident	25	8
Nursing staff	26	7
LDPS	29	4

The degree of agreement for ECOG ≥ 3 was also calculated using Cohen's kappa test, which showed lower site of agreement: self-score and the nursing staff ($\kappa = 0.57$), the nursing staff and LDPS ($\kappa = 0.25$), and self-score and LPDS ($\kappa = 0.006$) respectively.

Discussion

Our goal was to develop a device which will be able to challenge the human factor in the assessment of important input data forpotentially life-changing decisions.

An average ECOG score of residents was 1.76, in comparison to 1.82 estimated by nursing staff. However 9, 7, and 4 residents were rated PS ≥ 3 according to themselves, nurses, and LDPS, respectively. It means that residents rated themselves more critically. It was reported that patients tend to score themselves lower than professionals [12]. Our results suggest that 4 (residents' assessment) and 3 (nurses' assessment) residents were graded with worse PS (PS ≥ 3 instead of PS < 3) compared to LDPS. In real life, it means that 12 and 9% of residents were rated worse by themselves and nurses, respectively, in comparison to LDPS measurement. If a subjective grade is accepted in the therapy decision process, it theoretically means that cca. 10% of individuals receive palliative treatment instead of potentially curative treatment or that cca. 10% of individuals have less chance for specific treatment. However, treatment decisions are not based only on PS, but other factors are also considered, such as age and disease characteristics.

However, we do not have information about the amount of time spent in a recumbent or upright position but percentage

Table 5 Agreement in PS assessment between three methods

	Agreement in PS		Different PS	
	n	%	*n*	%
R: NS	28	85	5	15
R: LDPS	23	70	10	30
NS: LDPS	26	79	7	21

PS performance status, *R* resident, *NS* nursing staff, *LDPS* Logger data performance status

of time being up or recumbent as percentage of time around 50% which was chosen as cut-off value between PS 2 and PS 3 is critical as minimal difference in percentage around 50% has important impact on PS 2 or PS 3 classification. This is one of the weaknesses of the study.

A high degree of agreement for PS ≤ 2 between residents and nursing staff was evidenced. These can be attributed to the relatively low average grade PS score in the study, as various studies have shown that the agreement is higher in the lower grades (PS 0–2) than in the higher ones (PS 3–4). The existing literature cites both good and bad inter-rater reliability of PS scores. It is difficult to conclude which health care professionals PS assessments are more accurate [5, 13–15].

The LDPS study can be compared to other studies that tested some tools for physical activity measurement. One of these tools is accelerometer-based monitoring that measures step count. It was reported that in frail patients with advanced cancer and reduced PS, the device was accurate in providing estimates of body positions and step count [8]. The other tool is the electronic activity monitoring (EAM) systems. They measure the number of steps taken and the amount of time spent performing activities at different intensities. The potential of EAM devices has been explored in various disease settings to date including gynecology patients with polycystic ovarian syndrome, obese patients in pre-bariatric surgery survey, patients with chronic obstructive pulmonary disease, with chronic kidney disease, and psychiatry patients with depression [9]. The accuracy of EAM devices has also been confirmed in patients with colorectal cancer [10]. The EAM ActivPAL™ monitor was used in patients with thoracic cancer. The mean time spent each day in a range of activities, e.g., standing or stepping, or their frequency, e.g., number of sit-to-stand transitions and steps taken, were calculated and compared according to ECOG PS. The results showed that physical activity levels decline across the range of ECOG PS categories studied [16]. The LDPS on the other hand measures time of being upright or supine, situation which is the base of ECOG score system. Potential advantage of LDPS tool is measurement % of time being up or supine and support ECOG grading which also distinguishes % of time being up or supine. LDPS ECOG scores is only part of ECOG score judgment and for adequate scoring other patients' ability (for example capable of self-care) has also to be taken into account. It is clear that a decision between curative and palliative treatment cannot be made only on the basis of a single ECOG assessment. We believe that an LDPS device could be an additional help for PS assessment together with the patient him/herself deciding PS, accompanying persons, physicians, and nursing staff, or it could be a further source of support in PS evaluation in inconsistent or doubtful grading.

LDPS can be also a useful tool for PS measurements to be added to different models for survival of terminally ill patients, as it can objectively identify ECOG ≥ 3 PS [17, 18]. LDPS could also have a place in helping to enhance validation of prognostic scores in patients undergoing chemotherapy and at candidate selection for clinical trials [19].

Technically, this attempt for objective assessment of the PS yielded promising results, since the LDPS device itself is solid, objective, practical, handy to use, and cheap. Adherence to the device was 100%. No change in position or movement was noticed during measurement time when fixed at two points, and participants did not have any complaints wearing it. The results are also promising as difference between ECOG ≤ 2 and ECOG 3 is directly and easily evidenced, although number of participants in the study is small and differences are not statistically significant.

Weaknesses of the study are small sample group, relatively senior group of residents, non-cancerous patient's population, limited time of observation, one measurement only, and problematic LDPS numbers around 50%. Unfortunately, the amount of time spent in a supine or upright position, which would make judgment better, was not registered, but just percentage of time spent in an upright position.

Conclusion

The results show that more patients were classified as having a better PS with the tool. However, the sample size is rather small and difference was not significant.

However, the sample size is rather small and difference was not significant. In 12% of patients, a better performance score was observed with objective measurement in comparison to subjective judgment. In combination with standard measurements of PS, LDPS may provide a more accurate assessment of functional status. A relatively simple, well-attested USB data logger like LPDS can work as support for health care professionals in treatment decision as well as an incentive device for the individuals and for self-control.

Acknowledgments Residents and staff of DUC Tabor-Ljubljana, Ulica Janeza Pavla II/4, Ljubljana Retirement Home, its Manager Mrs. Silva Kučan, Head Nurse Mrs. Radmila Radanović, and Mr. Iztok Zupančič, Engineer.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Yates JW, Chalmer B, McKegney FP (1980) Evaluation of patients with advanced cancer using the Karnofsky performance status. Cancer 45(8):2220–2224
2. Evans C, McCarthy M (1985) Prognostic uncertainty in terminal care: can the Karnofsky index help? Lancet 1(8439):1204–1206
3. Hensing T, Cella D, Yount S (2005) The impact of ECOG performance status on quality of life symptoms in patients with advanced lung cancer. J Clin Oncol 23:8099
4. Maltoni M, Caraceni A, Brunelli C, Broeckaert B, Christakis N, Eychmueller S, Glare P, Nabal M, Viganò A, Larkin P, De Conno F, Hanks G, Kaasa S (2005) Prognostic factors in advanced cancer patients: evidence-based clinical recommendations—a study by the Steering Committee of the European Association for palliative care.; steering Committee of the European Association for palliative care. J Clin Oncol 23(25):6240–6248 Review
5. Chow R, Chiu N, Bruera E, Krishnan M, Chiu L, Lam H, DeAngelis C, Pulenzas N, Vuong S, Chow E (2016) Inter-rater reliability in performance status assessment among health care professionals: a systematic review. Ann Palliat Med 5(2):83–92
6. Sargent DJ, Köhne CH, Sanoff HK, Bot BM, Seymour MT, de Gramont A, Porschen R, Saltz LB, Rougier P, Tournigand C, Douillard JY, Stephens RJ, Grothey A, Goldberg RM (2009) Pooled safety and efficacy analysis examining the effect of performance status on outcomes in nine first-line treatment trials using individual data from patients with metastatic colorectal cancer. J Clin Oncol 27(12):1948–1955
7. Su C, Zhou F, Shen J, Zhao J, O'Brien M (2017) Treatment of elderly patients or patients who are performance status 2 (PS2) with advanced non-small cell lung cancer without epidermal growth factor receptor (EGFR) mutations and anaplastic lymphoma kinase (ALK) translocations—still a daily challenge. Eur J Cancer 83: 266–278
8. Skipworth RJ, Stene GB, Dahele M, Hendry PO, Small AC, Blum D, Kaasa S, Trottenberg P, Radbruch L, Strasser F, Preston T, Fearon KC, Helbostad JL (2011) Patient-focused endpoints in advanced cancer: criterion-based validation of accelerometer-based activity monitoring. Clin Nutr 30(6):812–821
9. Kelly CM, Shahrokni A (2016) Moving beyond Karnofsky and ECOG performance status assessments with new technologies. J Oncol 2016:6186543 Published online 2016 Mar 15
10. Kelly CM, Shahrokni A (2016) From shelf to bedside-wearable electronic activity monitoring technologies might assist oncologists in functional performance status assessment of older cancer patients. Clin Colorectal Cancer 30:S1533–0028(16)30256–0
11. Oken M, Creech R, Tormey D et al (1982) Toxicity and response criteria of the eastern cooperative oncology group. Am J Clin Oncol 5:649–655
12. Dajczman E, Kasymjanova G, Kreisman H, Swinton N, Pepe C, Small D (2008) Should patient-rated performance status affect treatment decisions in advanced lung cancer? J Thorac Oncol 3(10): 1133–1136
13. Abernethy AP, Shelby-James T, Fazekas BS, Woods D, Currow DC (2005) The Australia-modified Karnofsky performance status (AKPS) scale: a revised scale for contemporary palliative care clinical practice. BMC Palliat Care 4:7
14. Taylor AE, Olver IN, Sivanthan T, Chi M, Purnell C (1999) Observer error in grading performance status in cancer patients. Support Care Cancer 7(5):332–335
15. Blagden SP, Charman SC, Sharples LD, Magee LR, Gilligan D (2003) Performance status score: do patients and their oncologists agree? Br J Cancer 89(6):1022–1027
16. Maddocks M, Wilcock A (2012) Exploring physical activity level in patients with thoracic cancer: implications for use as an outcome measure. Support Care Cancer 20(5):1113–1116
17. Chow E, Abdolell M, Panzarella T, Harris K, Bezjak A, Warde P, Tannock I (2008) Predictive model for survival in patients with advanced cancer. J Clin Oncol 26(36):5863–5869
18. Suh SY, Choi YS, Shim JY, Kim YS, Yeom CH, Kim D, Park SA, Kim S, Seo JY, Kim SH, Kim D, Choi SE, Ahn HY (2010) Construction of a new, objective prognostic score for terminally ill cancer patients: a multicenter study. Support Care Cancer 18(2):151–157
19. Trédan O, Ray-Coquard I, Chvetzoff G, Rebattu P, Bajard A, Chabaud S, Pérol D, Saba C, Quiblier F, Blay JY, Bachelot T (2011) Validation of prognostic scores for survival in cancer patients beyond first-line therapy. BMC Cancer 11:95