

Assessment of cognitive function in patients with metastatic cancer: Are we using the right tools?

GEANA PAULA KURITA, M.S.N., PH.D.,^{1,2} MARLENE SANDVAD, RES. ASSIST.,¹
LENA LUNDORFF, M.D.,³ CIBELE ANDRUCIOLI DE MATTOS-PIMENTA, M.S.N., R.N., PH.D.,⁴
JETTE HØJSTED, M.D.,² AND PER SJØGREN, M.D., D.M.SC.^{1,5}

¹Palliative Research Group, Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

²Multidisciplinary Pain Centre, Department of Neuroanaesthesiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

³Department of Palliative Care, Uddevalla Hospital, Uddevalla, Sweden

⁴Department of Medical Surgical Nursing, School of Nursing, University of São Paulo, São Paulo, Brazil

⁵Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

(RECEIVED April 28, 2017; ACCEPTED July 1, 2017)

ABSTRACT

Objective: This study aimed at analyzing the validity and reliability of the continuous reaction time (CRT) test, the finger-tapping test (FTT), the Digit Span Test (DST), the Trail Making Test – part B (TMTB), and the Mini-Mental State Examination (MMSE) in patients with metastatic cancer.

Method: Eighty adult patients and 81 healthy controls were assessed between July of 2010 and November of 2015. The neuropsychological tests were analyzed regarding construct/discriminant/criterion validity and reliability.

Results: In terms of construct validity, it was not possible to estimate a model for the MMSE because of a skewed response distribution. For discriminant validity, patients were slower on two measures of the CRT ($p = 0.00483$, $p = 0.00030$) and FTT dominant hand ($p = 0.00306$). Regarding sensitivity and specificity, only the DST and TMTB seemed to predict cognitive deficit; however, the ROC curve areas were ≤ 0.73 . In terms of criterion validity, there were few significant correlations between the tests and the sociodemographic and clinical variables, and for the most part were very weak. Reliability was deemed to be adequate for the TMTB, DST, and FTT.

Significance of results: The findings of the full validation analyses were not clear-cut. However, CRT test, DST, FTT, and TMTB demonstrated partial positive results, indicating that these tests have good potential for use in clinical settings and require further study.

KEYWORDS: Cognition, Cognitive screening, Cancer, Neuropsychological assessment, Symptom validity

INTRODUCTION

Although cognitive complaints are frequent in cancer, systematic cognitive assessment is not routine in clinical practice, and mild alterations are seldom identified (Inouye et al., 2001; Pisani et al., 2003). These problems can be mild and subtle compared

to delirium and dementia and thus more difficult to detect, understand, and treat, but they may still be harmful to patients' ability to make decisions regarding their own treatment, as well as interfering with their roles within the family structure, at work, and in society, consequently reducing their quality of life. The lack of easily applied validated assessment tools probably contributes to the failure of identifying these problems. There are several neuropsychological tests, simple and complex, computerized and

Address correspondence and reprint requests to: Geana Kurita, Rigshospitalet, Blegdamsvej 9, Department 4111, 2100 Copenhagen, Denmark. E-mail: geana.kurita@regionh.dk.

not computerized, that have been developed to assess single or multiple cognitive domains; however, there is little information about their clinical validity and reliability in patients with cancer.

In addition, there is no consensus on which cognitive domains and tests are most clinically relevant and feasible to use (Kurita et al., 2009). Most of the tests do not have cutoff points based on specificity and sensitivity (Kurita et al., 2009) and may not be useful in a clinical context (dos Santos et al., 2014). Therefore, it seems justified to study the psychometric properties of a number of potentially relevant neuropsychological tests. Based on our experience and on studies regarding the cognitive effects of cancer, chronic pain, and opioids (Sjögren & Banning, 1989; Banning & Sjögren, 1990; Banning et al., 1992; Sjögren et al., 1994; Sjögren, 1997; Sjögren et al., 2000; Sjögren, 2006; Kurita & de Mattos-Pimenta, 2008; Kurita et al., 2011), we decided to assess the validity and reliability of the Continuous Reaction Time (CRT) test, the Finger Tapping Test (FTT), the Digit Span Test (DST), the Trail Making Test – part B (TMTB), and the Mini-Mental State Examination (MMSE) in patients with cancer.

METHODS

Design, Sample and Settings

Our sample was composed of 80 patients with metastatic cancer treated by specialized palliative care services at Rigshospitalet, Copenhagen University Hospital (66 outpatients) and at Herning Region Hospital (14 home-based patients) from July of 2010 to November of 2015. A control group of 81 healthy subjects (related or not related to patients) was assessed to allow for validity and reliability comparisons. Power calculations were performed based on the standard deviations with the CRT test and the FTT from previous studies (Sjögren & Banning, 1989; Banning & Sjögren, 1990; Sjögren et al., 1994; 2000). Assuming a power of 0.80 and a type I error of 0.05, the analyses indicated that a sample size of 50 patients would be sufficient to observe a mean difference of 4.0–8.5 ms on the CRT test and 5.1–5.7 taps on the FTT.

The inclusion criteria for our patients were: a diagnosis of metastatic cancer, aged ≥ 18 years, a Karnofsky Performance Status (KPS) index between 40 and 100%, at least 6 years of schooling, fluency in the Danish language, and stable medications for 4 days prior to first assessment (opioids, adjuvant analgesics, etc.). The exclusion criteria were: brain metastases, hepatic dysfunction, dementia, psychosis and/or delirium, misuse of drugs or alcohol, and last alcohol intake ≤ 24 hours.

The inclusion criteria were the same for healthy subjects with the exception of cancer diagnosis. Exclusion criteria were: treatment with medications with psychotropic properties, history of mental or physical disease or other chronic disease that could interfere with cognitive function, misuse of drugs or alcohol, and last alcohol intake ≤ 24 hours before testing.

Our study was submitted to the Danish Region Midtjylland Ethics Committee (no. M-20090221) and approved by the Danish Data Protection Agency for Health Science Research (no. 2007-58-0010/2012-58-0006). It was also conducted in accordance with the principles of the Declaration of Helsinki.

Neuropsychological Testing

Patients filled out and answered an identification form (sociodemographic data, cancer disease, KPS score, and treatments); two questions about sleep (“Do you feel rested?” and “How many hours of sleep last night?”); a visual numeric scale to assess “pain now” (0–10); the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983); and the European Organization for Research and Treatment of Cancer Quality of Life–Cancer 30 (EORTC QLQ–C30) (Aaronson et al., 1993). Healthy subjects filled in an identification form (socio-demographic data) and answered the two questions about sleep. Below are brief descriptions of each of the neuropsychological tests:

- The CRT test measures sustained attention. Through headphones, 100 auditory signals (500 Hz, 90 dB) were delivered to the patient at random intervals (2–5 s) over a period of 10 min. Subjects were instructed to press a button as soon as they heard the sound. A computer registered the time from emission of the sound signal to pressing of the button. Scores were summarized in milliseconds using the 10th, 50th, and 90th percentiles, where the 10th percentile represented the fastest and the 90th the slowest values (Elsass, 1986). More prolonged times meant worse performance.
- The DST assesses attention, concentration, and working memory (Wechsler, 1981). Subjects were asked to repeat series of numbers of increasing length, both forward (direct) and backward (reverse). The score was calculated by the number of correct answers. Scores ranged from 0 to 14, with higher scores meaning better performance.
- The FTT measures psychomotor speed. Participants were asked to use the second finger of each hand to make five 10-second trials tapping a key attached to a device that recorded the

number of taps (Peters, 1976). The score was calculated by the number of taps, with higher scores meaning better performance.

- The TMTB has been utilized to assess visual scanning speed, motor function, attention, and mental flexibility. The test has two parts (A and B). We only applied part B in our study. Numbers and letters must be connected in a crescent and alternated sequence on a sheet of paper. The time spent to correctly complete the test was recorded. Shorter times meant better performance (Reitan, 1958).
- The MMSE measures orientation to time and place, registration of words, attention, calculation, and word recall, as well as language and visual construction (Folstein et al., 1975). Scores can range from 0 to 30, with scores below 27 considered indicative of cognitive impairment.

Assessment Procedure

Eligible patients were identified during a medical appointment at the palliative care service. Patients' relatives were recruited via indication by the patients, and people not related to the patients were recruited from inside and outside of the hospital environment. They all received information regarding the study and signed written informed consent forms at the scheduled assessment appointment.

Patients and controls on stable medications were assessed and scheduled for a second assessment within an interval of 2 to 7 days (test and retest). The HADS and EORTC QLQ-C30 questionnaires were applied only to patients and answered once within a period of 2 days before or after the first assessment. Each cognitive test was explained in the standardized manner by the research assistant, and subjects were briefly pretested to ensure that they understood the tests. The control group was assessed regarding sociodemographic data and tested in the same way as the patients.

Data Analysis

Analyses were performed using SAS software (v. 9.1; SAS Institute, Cary, NC). The significance level was set at $p < 0.01$. The psychometric properties of the neuropsychological tools were analyzed with respect to the following:

- The construct validity of the MMSE was planned to be investigated through factorial analysis to evaluate the instrument items and domains that should comprise the cognitive state in patients with metastatic cancer.

- Criterion validity was examined by Spearman's correlation test between patients' results on the neuropsychological tests and sociodemographics, hours of sleep, and results on the HADS and EORTC QLQ-C30, and pain, which were applied at the first assessment. Our hypothesis was based on the idea that certain variables would present a significant positive or negative correlation with test results. Further, we analysed if a subject feeling rested at the time of the test would be associated with cognitive performance (Wilcoxon's test).
- Discriminant validity was analyzed by a comparison between patients and controls in the first assessment (Wilcoxon's test). The hypothesis was that patients would have worse cognitive performance. The sensitivity and specificity of each instrument were also analyzed in order to determine the cutoff values for the instruments. The area under the ROC curve was produced for each cognitive test to evaluate an ability to predict MMSE score ≤ 26 (cognitive deficits) during the first assessment (test).
- Reliability was analyzed using the intraclass correlation coefficient (ICC) (a two-way random single measure) for all of the neuropsychological instruments. Values were interpreted in accordance with the criteria described by Landis and Koch (1977).

Finally, as power calculations have only been performed to determine the detectable differences between patients with advanced cancer and healthy controls based on standard deviations on the CTR test and the FTT in earlier studies, we decided to carry out a post-hoc analysis with a sample power of $\alpha = 0.05$ for each neuropsychological test.

RESULTS

Eighty patients were tested, of whom 48 did not have alterations in terms of medication use and were returned to be retested after a mean interval of 5.2 days ($SD = 1.4$, range = 2–7 days). A total of 81 subjects comprised the control group (41 patients' relatives and 40 people not related to the patients), of whom 71 were retested (32 and 39, respectively) after a mean interval of 4.7 days ($SD = 1.3$, range = 3–7). The majority of patients were treated with opioids ($n = 61$), anticonvulsants ($n = 45$), antidepressants ($n = 31$), corticosteroids ($n = 19$), and hypnotics ($n = 19$). Some received chemotherapy ($n = 16$) and radiotherapy ($n = 1$).

Patients and controls had similar sociodemographic characteristics but differed with regard to

Table 1. Comparison between patients and controls at the first assessment

Variable	Patients	Controls	Wilcoxon's test <i>p</i>	Fisher's test <i>p</i>
Sex				
Male/female	31/49	38/43		0.3405
Cohabitation				
Yes/no	54/26	58/23		0.6101
Age				
<i>n</i>	80	81	0.89234	
Mean (<i>SD</i>)	61.2 (10.5)	61.7 (8.8)		
Median (min–max)	63.5 (39–87)	62 (40–83)		
Schooling				
<i>n</i>	76	72	0.13608	
Mean (<i>SD</i>)	11.1 (3.4)	12.0 (3.5)		
Median (min–max)	10.0 (6–21)	12.0 (7–23)		
Income/year				
<i>n</i>	78	81	0.00019	
<200.000 DKK	36	16		
200.000–500.000 DKK	27	33		
>500.000 DKK	15	32		
Sleep hours				
<i>n</i>	79	81	0.13445	
Mean (<i>SD</i>)	7.1 (1.7)	6.9 (1.1)		
Median (min–max)	7.0 (2.0–10.5)	7.0 (4.0–10.0)		
Rest				
<i>n</i>	78	81		0.0054
Yes/no	46/32	65/16		
Karnofsky Performance Scale				
<i>n</i>	80			
Mean (<i>SD</i>)	70.8 (9.4)			
Median (min–max)	70.0 (50–100)			

1 DKK = 7.5 €.

income ($p = 0.00019$) and sensation of rest ($p = 0.0054$) (Table 1). These differences were expected since it is not uncommon that patients with cancer have difficulty to carry on professional activities, and, as well, cancer disease/comorbidities/treatments frequently cause fatigue and exhaustion. Therefore, no specific statistical adjustment for these variables was done.

Construct Validity

It was not possible to estimate a model for factorial analysis of the MMSE because of the skewed response distributions for many items.

Criterion Validity

The significant correlations followed the expected direction (positive or negative), but the majority of correlations were weak. Correlations were significant between MMSE, age, and income; CRT, nausea/vomiting, and KPS; DST, schooling, KPS score, and insomnia; and TMT, age, schooling, income, and KPS score (Table 2). In addition, men had better performance on the FTT (dominant $p = 0.00000$, nondomi-

nant $p = 0.00011$) than women, while no differences were observed regarding cohabitation and sensation of rest (data not shown).

Discriminant Validity

The cognitive test results for patients' relatives were compared to those for people not related to patients so as to ensure that they had similar performance to and could be included in control group. They were collapsed, as no significant differences were observed (Table 3). Patients had slower performance on the CRT 50th ($p = 0.00483$) and 90th ($p = 0.00030$) percentiles and FTT dominant hand ($p = 0.00306$) compared to controls (Table 3). Regarding sensitivity and specificity, only the DST and TMTB seemed to predict cognitive deficits in agreement with MMSE scores; however, our results were fair, at best (ROC curves ≤ 0.73).

Reliability

The patient group presented almost perfect and substantial conformity between the two cognitive assessments on 5 of the 10 measures: the TMTB, FTT

Table 2. Correlations between patients' cognitive results and variables that could interfere with cognitive function at first assessment (criterion validity)

Variable	MMSE		CRT 10th		CRT 50th		CRT 90th		DST forward		DST backwards		TMTB		FTT dominant		FTT nondominant		FTT difference	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Age	-0.30675	0.0060	-0.11818	0.2996	-0.02738	0.8107	0.07823	0.4931	-0.11389	0.3144	-0.15636	0.1660	0.55467	<.0001	-0.21580	0.0545	-0.09695	0.3923	-0.22069	0.0492
Sleep	0.02383	0.8359	0.12139	0.2898	0.09906	0.3882	0.12705	0.2677	0.10153	0.3733	0.10615	0.3518	-0.08604	0.4724	0.02483	0.8280	0.04600	0.6873	-0.05201	0.6490
Schooling	0.25102	0.0298	-0.24945	0.0309	-0.26844	0.0199	-0.21716	0.0613	0.31984	0.0049	0.28540	0.0125	-0.42918	0.0002	-0.00290	0.9801	-0.00248	0.9830	0.09888	0.3954
Income	0.37110	0.0009	-0.12533	0.2775	-0.17136	0.1362	-0.18080	0.1156	0.13441	0.2407	0.18464	0.1056	-0.32338	0.0059	0.05506	0.6321	-0.00719	0.9502	0.17366	0.1284
KPS	0.01289	0.9102	-0.20808	0.0657	-0.24677	0.0284	-0.32367	0.0036	0.35226	0.0014	0.18070	0.1087	-0.38367	0.0008	-0.10723	0.3438	-0.20218	0.0721	0.17020	0.1312
Pain	0.01307	0.9096	0.07179	0.5322	0.03376	0.7692	-0.02593	0.8217	-0.23687	0.0356	-0.18418	0.1042	0.05496	0.6466	-0.12755	0.2626	-0.17298	0.1274	-0.01738	0.8792
Anxiety	-0.11032	0.3363	-0.06214	0.5889	0.05729	0.6183	0.09696	0.3984	-0.12928	0.2562	-0.22983	0.0416	0.10671	0.3723	-0.03210	0.7788	-0.12419	0.2755	0.18701	0.0989
Depression	-0.05659	0.6227	0.00796	0.9448	0.09275	0.4193	0.06722	0.5587	-0.26335	0.0190	-0.24042	0.0328	0.15116	0.2050	0.02794	0.8069	-0.08936	0.4336	0.11831	0.2990
PF	0.06726	0.5611	-0.02035	0.8606	-0.07003	0.5451	-0.09918	0.3908	0.19658	0.0845	0.20751	0.0683	-0.10833	0.3650	0.02514	0.8271	0.06634	0.5639	-0.06808	0.5537
RF	-0.16557	0.1501	-0.24688	0.0304	-0.22655	0.0476	-0.26729	0.0188	0.10063	0.3807	0.02856	0.8040	-0.09479	0.4283	-0.03088	0.7884	0.04836	0.6742	-0.18671	0.1017
EF	-0.10425	0.3734	-0.15250	0.1915	-0.20593	0.0763	-0.14662	0.2094	0.09729	0.4031	0.23638	0.0398	-0.06908	0.5727	-0.12892	0.2670	-0.06576	0.5725	-0.14420	0.2139
CF	-0.06515	0.5787	-0.07980	0.4961	-0.11803	0.3132	-0.18010	0.1221	0.09729	0.4031	0.27115	0.0178	-0.16743	0.1691	-0.05028	0.6662	-0.07228	0.5349	-0.01567	0.8931
SF	-0.05922	0.6137	-0.00970	0.9341	-0.03165	0.7875	-0.05199	0.6578	0.06040	0.6042	0.07466	0.5215	-0.14208	0.2442	-0.10371	0.3727	0.05485	0.6379	-0.24441	0.0334
QoL	0.13825	0.2369	-0.04062	0.7293	-0.07893	0.5009	-0.09350	0.4249	0.15850	0.1714	0.12808	0.2702	-0.05069	0.6792	-0.02573	0.8254	-0.00932	0.9363	-0.01886	0.8715
Fatigue	0.05682	0.6235	0.08853	0.4439	0.15908	0.1670	0.21875	0.0560	-0.19060	0.0946	-0.14460	0.2065	0.16149	0.1753	0.04987	0.6646	0.05911	0.6072	0.06823	0.5528
Nausea/ vomiting	0.09293	0.4215	0.31317	0.0055	0.35580	0.0015	0.34783	0.0019	-0.11111	0.3328	-0.15699	0.1699	0.21835	0.0654	0.09562	0.4050	0.12032	0.2940	0.00026	0.9982
Pain	-0.05087	0.6583	0.09010	0.4328	0.10003	0.3836	0.11088	0.3338	-0.11783	0.3010	-0.16305	0.1511	0.15175	0.2032	0.14220	0.2113	0.11817	0.2997	0.02950	0.7964
Dyspnea	0.00873	0.9404	0.05707	0.6244	0.07847	0.5005	0.05513	0.6362	-0.06149	0.5953	-0.06100	0.5982	0.08329	0.4899	-0.12054	0.2964	-0.14212	0.2176	-0.00784	0.9460
Insomnia	-0.01209	0.9175	-0.01809	0.8768	-0.03290	0.7778	-0.09346	0.4220	-0.05575	0.6301	-0.33776	0.0027	0.20162	0.0918	0.16531	0.1508	0.03998	0.7299	0.17783	0.1218
Appetite loss	0.03417	0.7695	0.07736	0.5065	0.01023	0.9301	-0.00262	0.9821	-0.05090	0.6602	-0.11439	0.3219	0.18483	0.1228	0.18280	0.1115	0.22385	0.0503	0.00255	0.9824
Constipation	0.28822	0.0122	0.00132	0.9910	0.01971	0.8667	0.00208	0.9859	-0.09180	0.4303	-0.21187	0.0662	0.12763	0.2960	0.22015	0.0560	0.19223	0.0962	0.03742	0.7483
Diarrhea	-0.05189	0.6606	0.12719	0.2802	0.10650	0.3664	0.09330	0.4291	-0.09020	0.4416	-0.25375	0.0280	0.08865	0.4689	0.24116	0.0371	0.15962	0.1713	0.09995	0.3936
FD	0.18616	0.1098	-0.05218	0.6566	-0.06699	0.5680	-0.09858	0.4001	0.16576	0.1524	-0.13725	0.2371	-0.09605	0.4324	-0.00506	0.9654	-0.05914	0.6118	0.11059	0.3416

Sleep in hours.

CF = cognitive functioning; CRT = Continuous Reaction Time; DST = Digit Span Test; EF = emotional functioning; FD = financial difficulties; FTT = Finger Tapping Test; MMSE = Mini-Mental State Examination; QoL = quality of life; PF = physical functioning; RF = role functioning; SF = social functioning; TMTB = Trail Making Test - part B.

Table 3. Cognitive performance of the groups (discriminant validity)

Tests	Patient		Relative		Not relative		<i>p</i> (relative vs. not relative)	<i>p</i> (patient vs. control)
	Test	Retest	Test	Retest	Test	Retest		
CRT 10th								
<i>n</i>	79	57	41	32	39	38	0.26201	0.03114
Mean (<i>SD</i>)	153.8 (32.1)	152.0 (38.3)	144.5 (22.8)	144.5 (22.2)	142.4 (30.6)	141.4 (28.9)		
Median (min–max)	147.0 (114.0–284.0)	141.0 (34.0–284.0)	145.0 (108.0–227.0)	142.5 (107.0–194.0)	135.0 (101.0–263.0)	135.5 (103.0–237.0)		
CRT 50th								
<i>n</i>	79	57	41	32	39	38	0.19043	0.00483
Mean (<i>SD</i>)	194.8 (59.4)	194.6 (866.8)	174.6 (32.2)	173.7 (31.1)	168.5 (37.2)	168.3 (40.0)		
Median (min–max)	179.0 (136.0–532.0)	178.0 (84.0–474.0)	176 (124.0–276.0)	177.0 (124.0–252.0)	162.0 (112.0–318.0)	161.0 (114.0–289.0)		
CRT 90th								
<i>n</i>	79	57	41	32	39	38	0.21604	0.00030
Mean (<i>SD</i>)	372.1 (93.2)	270.0 (129.2)	227.5 (65.6)	229.2 (50.1)	219.2 (61.9)	220.1 (57.7)		
Median (min–max)	237.0 (167.0–720.0)	227.0 (84.0–867.0)	219 (75.0–419.0)	225.5 (162.0–356.0)	207.0 (132.0–411.0)	209.0 (132.0–369.0)		
DST forward								
<i>n</i>	80	57	41	32	40	39	0.38589	0.35698
Mean (<i>SD</i>)	6.9 (5.9)	6.3 (2.3)	6.4 (2.0)	6.8 (2.4)	6.7 (2.0)	7.1 (2.3)		
Median (min–max)	6.0 (2.0–11.0)	6.0 (2.0–13.0)	6.0 (3.0–11.0)	6.0 (2.0–13.0)	6.0 (4.0–13.0)	6.0 (3.0–13.0)		
DST backward								
<i>n</i>	80	57	41	32	40	39	0.02027	0.06764
Mean (<i>SD</i>)	5.2 (1.9)	5.18 (2.3)	5.2 (1.9)	5.3 (2.3)	6.1 (1.7)	6.2 (2.1)		
Median (min–max)	5.0 (2.0–10.0)	5.0 (2.0–12.0)	5.0 (2.0–10.0)	5.0 (2.0–12.0)	6.0 (2.0–11.0)	6.0 (2.0–11.0)		
TMTB								
<i>n</i>	73	51	37	26	40	39	0.22299	0.01066
Mean (<i>SD</i>)	134.4 (112.6)	122.7 (128.4)	102.2 (47.9)	103.3 (58.1)	94.1 (50.9)	85.6 (40.2)		
Median (min–max)	102.0 (43.0–882.0)	92.0 (8.0–912)	89 (41.0–240.0)	92.0 (37.0–317.0)	83.0 (47.0–341.0)	77.0 (45.0–255.0)		
FTT dominant								
<i>n</i>	80	57	41	32	40	39	0.14569	0.00306
Mean (<i>SD</i>)	39.6 (10.2)	42.0 (10.9)	43.0 (9.3)	44.9 (9.5)	44.6 (6.6)	46.1 (6.5)		
Median (min–max)	39.6 (12.0–66.4)	43.2 (8.8–67.2)	42.0 (29.4–73.6)	44.7 (27.0–67.2)	45.4 (28.2–54.4)	46.0 (29.8–60.0)		
FTT nondominant								
<i>n</i>	80	57	41	32	40	39	0.86496	0.01450
Mean (<i>SD</i>)	36.9 (9.0)	38.8 (10.2)	41.1 (10.2)	41.4 (9.2)	39.7 (6.3)	40.6 (5.6)		
Median (min–max)	37.2 (11.0–57.6)	38.0 (9.4–58.6)	40.0 (22.6–76.6)	41.1 (25.8–65.0)	40.4 (25.4–53.0)	42.2 (25.6–49.8)		
FTT difference								
<i>n</i>	80	57	41	32	40	39	0.01538	0.43875
Mean (<i>SD</i>)	2.7 (5.9)	3.2 (6.0)	1.9 (5.3)	3.5 (5.1)	4.9 (5.6)	5.5 (5.6)		
Median (min–max)	2.9 (–11.8–17.6)	3.6 (–16.6–15.4)	2.4 (–10.0–14.0)	3.7 (–10.0–12.4)	4.4 (–7.8–18.2)	4.8 (–4.8–17.4)		
MMSE								
<i>n</i>	79	57	41	32	40	39	0.05179	0.24878
Mean (<i>SD</i>)	28.4 (1.8)	28.5 (1.9)	29.0 (1.5)	29.3 (0.8)	28.8 (1.1)	28.9 (1.6)		
Median (min–max)	29.0 (23.0–30.0)	29.0 (22.0–30.0)	30.0 (23.0–30.0)	30.0 (28.0–30.0)	29.0 (26.0–30.0)	30.0 (23.0–30.0)		

Control = relative + not relative, CRT = Continuous Reaction Time; DST = Digit Span Test; FTT = Finger Tapping Test; MMSE = Mini-Mental State Examination; TMTB = Trail Making Test – part B.

Table 4. Agreement between cognitive results of test–retest in the two groups (reliability)

Test	ICC			
	Patient		Control	
TMTB	0.91759	almost perfect	0.84171	almost perfect
DST forward	0.65478	substantial	0.72442	substantial
DST backward	0.71656	substantial	0.73484	substantial
FTT non-dominant	0.87451	almost perfect	0.86782	almost perfect
FTT dominant	0.74396	substantial	0.82249	almost perfect
FTT difference	0.32525	fair	0.57629	moderate
CRT 10th	0.49155	moderate	0.88753	almost perfect
CRT 50th	0.59824	moderate	0.90408	almost perfect
CRT 90th	0.59824	moderate	0.82602	almost perfect
MMSE	0.50526	moderate	0.27453	fair

CRT = Continuous Reaction Time; DST = Digit Span Test; FTT = Finger Tapping Test; ICC = intraclass correlation coefficient; MMSE = Mini-Mental State Examination; TMTB = Trail Making Test – part B.

dominant and nondominant hand, DST forward, and DST backward. The control group demonstrated almost perfect and substantial consistency/conformity in 8 of the 10 measures (Table 4). Approximately 45% of the control group had higher MMSE scores at the second assessment, which may reflect the poor consistency for this test due to a learning effect.

Post-Hoc Power Calculation

The calculations based on the present results demonstrated the following power for each test: CRT 10th percentile = 0.59; CRT 50th = 0.85; CRT 90th = 1.00; FTT dominant = 0.82; FTT non-dominant = 0.72; FTT difference between hands = 0.12; TMTB = 0.72; MMSE = 0.52; DST forward = 0.09; and DST backward = 0.28.

DISCUSSION

Several factors may interfere with a routine assessment of symptoms in patients with advanced cancer, including barriers related to caregivers and patients (Ingham et al., 2015; Mitnick et al., 2010). Although assessment of cognitive function in patients with cancer is not standard-of-care in palliative care, the increase in life expectancy of patients with cancer due to improvement in diagnostics and treatments is beginning to draw the attention of clinicians to symptoms that may be controlled and/or prevented in order to preserve and improve quality of life, and even survival.

The tools that are being used to assess cognitive function in the cancer population were originally developed to screen or assess for a specific central nervous system dysfunction or disease, differentiating those with cognitive impairment from those without this disease/condition (a “normal population”)

(Harvey, 2012). For several years, our research group has used neuropsychological testing in patients with cancer, especially to assess cognitive status before/after medications and to compare groups of patients and analyze the factors that may cause deficits (Kurita et al., 2009; Sjøgren & Banning, 1989; Sjøgren et al., 1994; 2000; Kurita & de Mattos-Pimenta, 2008; Kurita et al., 2011; 2016). Due to the fact that many of the tools we have used for research were not designed to assess patients with cancer, a validation process is necessary to ensure their suitability and feasibility. Moreover, preestablished cutoff points may be difficult to apply to patients with cancer due to the fluctuating nature of their health condition.

In this present study, analysis of the 5 neuropsychological tests in 80 patients with metastatic cancer in specialized palliative care showed good reliability and some positive outcomes regarding examination of validity, indicating that the tests individually or combined have the potential to assess cognitive function in this population.

The attempt to examine the construct validity of the MMSE by confirmation of the different cognitive domains that comprise the instrument failed. There was a skewed response distribution for many of its items represented by high scores, indicating no cognitive deficits for most patients. This finding may indicate that the tool failed to detect milder cognitive deficits due to its lack of sensitivity, which is confirmed by the lack of statistical power found in the post-hoc power calculations. Considering the calculations, it would be necessary to involve a huge sample to demonstrate differences between groups that could capture mild cognitive deficits. The idea of validating this tool in patients with cancer is based on the premise that its original purpose was patient screening for dementia and other psychiatric conditions

that involved severe cognitive impairment (Folstein et al., 1975). However, it has been widely used to assess cognitive function in several other health/disease conditions, including in patients with cancer (Hjermstad et al., 2004; Kurita et al., 2011; Kurita et al., 2016). In addition, we found only one study regarding validation of the MMSE in patients with cancer that confirmed the good psychometric properties of the Greek version of the instrument by other validation methods (Mystakidou et al., 2007). Thus, the main criticisms of this instrument are that it is not sensitive enough to detect milder but clinically relevant cognitive deficits and that it is subject to significant learning effects (Lange et al., 2014), although its simple and brief application may seem to offer an attractive advantage.

The criterion validity analyzed by correlations between the scores on the neuropsychological tests and variables related to patients' characteristics and symptoms demonstrated very few and weak relationships. Age, schooling, income, and KPS score seemed to be the main factors with some degree of relationship to at least some, but not all, of the tests. Similar to our study, a validation study with 90 Brazilian patients with cancer in palliative care observed a lack of significant correlations between the TMTB and such symptoms as pain, fatigue, depression, anxiety, and sensation of rest (dos Santos et al., 2014). On the other hand, a study analyzing the psychometric properties of the MMSE in 103 Greek patients with cancer found significant moderate correlations with age, schooling, marital status, metastasis, and treatments (Mystakidou et al., 2007). This is a complex area that is still under investigation, but in the present study the absence of strong significant correlations with such factors as age, schooling, and income, previously considered factors to be reliably associated with cognitive performance (Glisky, 2007; Tucker-Drob et al., 2009), was a surprising finding. It was also unexpected that no significant association between a sensation of rest and performance on the cognitive tests was found, since 41% of the patients felt unrested. Fatigue and exhaustion have formerly been described during progression of cancer and its treatments, and they have been associated with cognitive dysfunction (Janelins et al., 2014).

Concerning discriminant validity, some measures of the CRT test (50th and 90th percentiles) and the FTT (dominant hand) were able to discriminate patients from controls, where patients had worse cognitive performance. In previous studies regarding opioid treatment, worse performance of adult patients with cancer was registered on the CRT test and the FTT when compared with healthy controls or less sick patients or those on different treatments. These studies also showed that differences are not

always observed on all measures of the tests (Sjögren & Banning, 1989; Banning & Sjögren, 1990; Banning et al., 1992; Sjögren et al., 1994; 2000). Post-hoc power calculations demonstrated that the sample of the present study was too small to capture significant differences between groups on the MMSE, DST, and TMT tests, and on the 10th percentile of the CRT and FTT nondominant hand and the difference between hands. An attempt to determine cutoff values for each instrument did not exhibit good sensitivity and specificity in our sample. The literature regarding validation of these neuropsychological tests in patients with cancer is scarce, which makes comparisons across studies most difficult. We found a study regarding validation of a simple auditory CRT test and a battery for attentional performance with simple and complex visual continuous reaction time tests. The comparison between healthy subjects ($n = 74$) and patients with cancer from the surgical, gastroenterology, oncology, and internal medicine departments ($n = 70$) demonstrated that reaction time was longer in those patients (Jakobsen et al., 2011). Another study evaluated the psychometric properties of the MMSE in patients with cancer, and the instrument was able to discriminate between subgroups of patients with different degrees of disease severity (Mystakidou et al., 2007). Perhaps at the first assessment the patient sample in our present study did not present with a disease condition that would strongly interfere with cognitive function. Credence for this notion is supported by the mean KPS score being near 71% and the weak correlations with symptoms in our patient sample.

Reliability was almost perfect or substantial for the TMTB, DST, and FTT, but it varied between fair and moderate for the other tests. The disagreement between test and retest reliability may be explained by the vulnerable and unstable health condition in the advanced stages of cancer and the learning effect with some tests. Unfortunately, a second measure of quality of life and symptoms was not performed, so that we cannot confirm this hypothesis. In addition, as previously mentioned, the MMSE is highly criticized for its simplicity, easy memorization, and the possibility of learning effects.

This is one of the first and largest studies to investigate the validity of a battery of neuropsychological tests in patients with metastatic cancer. Considering that psychometric evaluation of assessment tools is an imperative step in order to select which measurement tools should be included in clinical as well as research endeavors in palliative cancer care, our study has produced new knowledge. The results from some analyses were less than definitive, but they offer information that should encourage further exploration of the area. We acknowledge that the abovementioned

limitations of the tool characteristics (e.g., insensitivity to detect milder cognitive deficits, the learning effect, and methodological variances), the patient characteristics (e.g., low variation on cognitive performance, reflecting possible good cognitive performance), and the underpowered sample size for some of the tests may have contributed to some of the tests' weak properties observed during this validation process.

The neuropsychological tests analyzed demonstrated good reliability, and some measures of the CRT test and the FTT could discriminate patients from controls, indicating that they have the potential to assess cognitive function in patients with metastatic cancer in clinical and research settings. The results of the present analysis provide a warning to avoid hasty conclusions about the patients' mental condition. Further evaluation of these neuropsychological tools is mandatory to ensure that their properties are fitted for this population.

Our study also offers insight for recommendation of future studies with larger samples and analysis of subgroups with pre-confirmed cognitive dysfunction to provide more consistent results, especially regarding establishment of cutoff points for clinical and research use.

DISCLOSURES AND ACKNOWLEDGEMENTS

This study received no funding, and the authors hereby state that they have no conflicts of interest to disclose. We thank biostatistician Morten Aagaard Petersen for his assistance with the statistical analysis.

Our study was developed in partnership with the School of Nursing at the University of São Paulo, where a similar parallel investigation was conducted among Brazilian patients with cancer.

REFERENCES

- Aaronson, N.K., Ahmedzai, S., Bergman, B., et al. (1993). The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85(5), 365–376.
- Banning, A. & Sjøgren, P. (1990). Cerebral effects of long-term oral opioids in cancer patients measured by continuous reaction time. *Clinical Journal of Pain*, 6, 91–95.
- Banning, A., Sjøgren, P. & Kaiser, F. (1992). Reaction time in cancer patients receiving peripherally acting analgesics alone or in combination with opioids. *Acta Anaesthetologica Scandinavica*, 36(5), 480–482.
- dos Santos, J., de Mattos-Pimenta, C.A., Kurita, G., et al. (2014). Validation of the Trail Making Test B for the cognitive assessment of patients with cancer in palliative care. *Open Journal of Statistics*, 4, 435–445. Available from https://file.scirp.org/pdf/OJS_2014082716264133.pdf.
- Elsass, P. (1986). Continuous reaction times in cerebral dysfunction. *Acta Neurologica Scandinavica*, 73(3), 1–22.
- Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.
- Glisky, E.L. (2007). Changes in cognitive function in human aging. In *Brain Aging: Models, Methods, and Mechanisms*. D.R. Riddle (ed.), chap. 1. Boca Raton, FL: CRC Press/Taylor & Francis. Available from <http://www.ncbi.nlm.nih.gov/books/NBK3885/>.
- Harvey, P.D. (2012). Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience*, 14(1), 91–99.
- Hjermstad, M.J., Loge, J.H. & Kaasa, S. (2004). Methods for assessment of cognitive failure and delirium in palliative care patients: Implications for practice and research. *Palliative Medicine*, 18(6), 494–506.
- Ingham, J., Moore, H., Phillips, J.L., et al. (2015). The measurement of, and tools for, pain and other symptoms. In *Oxford Textbook of Palliative Medicine*, 5th ed. N. Cherny et al. (eds.), pp. 377–390. Oxford: Oxford University Press.
- Inouye, S.K., Foreman, M.D., Mion, L.C., et al. (2001). Nurses' recognition of delirium and its symptoms: Comparison of nurse and researcher ratings. *Archives of Internal Medicine*, 161(20), 2467–2473.
- Jakobsen, L.H., Sorensen, J.M., Rask, I.K., et al. (2011). Validation of reaction time as a measure of cognitive function and quality of life in healthy subjects and patients. *Nutrition*, 27(5), 561–570.
- Janelins, M.C., Kesler, S.R., Ahles, T.A., et al. (2014). Prevalence, mechanisms, and management of cancer-related cognitive impairment. *International Review of Psychiatry*, 26(1), 102–113.
- Kurita, G.P. & de Mattos-Pimenta, C.A. (2008). Cognitive impairment in cancer pain patients receiving opioids: A pilot study. *Cancer Nursing*, 31(1), 49–57.
- Kurita, G.P., Lundorff, L., Pimenta, C.A., et al. (2009). The cognitive effects of opioids in cancer: A systematic review. *Supportive Care in Cancer*, 17(1), 11–21.
- Kurita, G.P., Sjøgren, P., Ekholm, O., et al. (2011). Prevalence and predictors of cognitive dysfunction in opioid treated cancer patients: A multinational study. *Journal of Clinical Oncology*, 29(10), 1297–1303.
- Kurita, G.P., Ekholm, O., Kaasa, S., et al. (2016). Genetic variation and cognitive dysfunction in opioid-treated patients with cancer. *Brain and Behavior*, 6(7), e00471. Available from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4864175/pdf/BRB3-6-e00471.pdf>.
- Landis, J. & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Lange, M., Rigal, O., Clarisse, B., et al. (2014). Cognitive dysfunctions in elderly cancer patients: A new challenge for oncologists. *Cancer Treatment Reviews*, 40(6), 810–817.
- Mitnick, S., Leffler, C., Hood, V.L., et al. (2010). Family caregivers, patients and physicians: Ethical guidance to optimize relationships. *Journal of General Internal Medicine*, 25(3), 255–260.
- Mystakidou, K., Tsilika, E., Parpa, E., et al. (2007). Brief cognitive assessment of cancer patients: Evaluation of the Mini-Mental State Examination (MMSE) psychometric properties. *Psycho-Oncology*, 16(4), 352–357.

- Peters, M. (1976). Prolonged practice of a simple motor task by preferred and non-preferred hands. *Perceptual and Motor Skills*, 43, 447–450.
- Pisani, M.A., Redlich, C., McNicoll, L., et al. (2003). Underrecognition of preexisting cognitive impairment by physicians in older ICU patients. *Chest*, 124(6), 2267–2274.
- Reitan, R.M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Sjøgren, P. (1997). Psychomotor and cognitive functioning in cancer patients. *Acta Anaesthesiologica Scandinavica*, 41(1, Pt. 2), 159–161.
- Sjøgren, P. (2006). *Cognitive Function in Cancer and Chronic Non-Malignant Pain*. Doctoral dissertation. Aarhus, Denmark: Faculty of Health Sciences, University of Aarhus.
- Sjøgren, P. & Banning, A. (1989). Pain, sedation and reaction time during long-term treatment of cancer patients with oral and epidural opioids. *Pain*, 39(1), 5–11.
- Sjøgren, P., Banning, A.M., Christensen, C.B., et al. (1994). Continuous reaction time after single dose, long-term oral and epidural opioid administration. *European Journal of Anaesthesiology*, 11(2), 95–100.
- Sjøgren, P., Olsen, A.K., Thomsen, A.B., et al. (2000). Neuropsychological performance in cancer patients: The role of oral opioids, pain and performance status. *Pain*, 86(3), 237–245.
- Tucker-Drob, E.M., Johnson, K.E. & Jones, R.N. (2009). The cognitive reserve hypothesis: A longitudinal examination of age-associated declines in reasoning and processing speed. *Developmental Psychology*, 45(2), 431–446.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised*. New York: The Psychological Corporation.
- Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370.