ORIGINAL ARTICLE

Selection Bias Introduced by Neuropsychological Assessments

Robert Olson, Maureen Parkinson, Michael McKenzie

ABSTRACT: ***Objective:*** Two prospective studies in patient with brain tumours were performed comparing the Mini Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA). The first assessed their feasibility and the second compared their diagnostic accuracy against a four-hour neuropsychological assessment (NPA). The introduction of the NPA decreased accrual and retention rates. We were therefore concerned regarding potential selection bias. ***Methods:*** Ninety-two patients were prospectively accrued and subsequently divided into three categories: a) no NPA required b) withdrew consent to NPA c) completed NPA. In order to quantify any potential bias introduced by the NPA, patient demographics and cognitive test scores were compared between the three groups. ***Results:*** There were significant differences in age ($p<0.001$), education ($p=0.034$), dexamethasone use ($p=0.002$), MMSE ($p=0.005$), and MoCA scores ($p<0.001$) across the different study groups. Furthermore, with increasing involvement of the NPA, patients' cognitive scores and educational status increased, while their age, dexamethasone use, and opioid use all decreased. Individuals who completed the NPA had higher MoCA scores than individuals who were not asked to complete the NPA (24.7 vs. 20.5; $p < 0.001$). In addition, this relationship held when restricting the analyses to individuals with brain metastases ($p < 0.001$). ***Conclusions:*** In this study, the lengthy NPA chosen introduced a statistically and clinically significant source of selection bias. These results highlight the importance of selecting brief and well tolerated assessments when possible. However, researchers are challenged by weighing the improved selection bias associated with brief assessments at the cost of reduced diagnostic accuracy.

RÉSUMÉ: Biais de sélection introduits par les évaluations neuropsychologiques. ***Objectif :*** Deux études prospectives ont comparé le Mini Mental State Examination (MMSE) et le Montreal Cognitive Assessment (MoCa) chez des patients atteints de tumeurs cérébrales. La première étude évaluait leur faisabilité et la seconde comparait leur précision diagnostique à une évaluation neuropsychologique de quatre heures (ENP). L'introduction de l'ENP a diminué le taux de recrutement et de rétention. La possibilité d'un biais de sélection nous a donc préoccupés. ***Méthodes :*** Quatre-vingt-douze patients ont été recrutés de façon prospective et ils ont été subséquemment divisés en trois catégories : a) aucune ENP requise; b) retrait du consentement à l'ENP; c) ENP complétée. Les données démographiques des patients et les scores aux tests cognitifs ont été comparés entre les trois groupes afin de quantifier tout biais potentiel introduit par l'ENP. ***Résultats :*** Nous avons observé les différences significatives suivantes entre les trois groupes : l'âge ($p < 0{,}001$), le niveau de scolarité ($p = 0{,}034$), la prise de dexaméthasone ($p = 0{,}002$), le score MMSE ($p = 0{,}005$) et le score MoCa ($p < 0{,}001$). En outre, plus l'ENP était complexe, plus les scores cognitifs et le niveau de scolarité des patients augmentaient et plus leur âge, leur prise de dexaméthasone et leur prise d'opiacés diminuaient. Les individus qui ont complété l'ENP avaient des scores MoCa plus élevés que les individus à qui on n'a pas demandé de compléter l'ENP (24,7 par opposition à 20,5; $p < 0{,}001$). De plus, cette relation était toujours présente quand les analyses portaient seulement sur les individus qui avaient des métastases cérébrales ($p < 0{,}001$). ***Conclusions :*** Dans cette étude, la longue ENP choisie a introduit un biais de sélection significatif au point de vue statistique et au point de vue clinique. Ces résultats soulignent l'importance de choisir autant que possible des évaluations qui sont brèves et bien acceptées par les patients. Cependant, ceci constitue un défi pour les chercheurs qui doivent choisir entre une réduction du biais de sélection avec des évaluations brèves au prix d'une précision diagnostique moindre.

Can. J. Neurol. Sci. 2010; 37: 264-268

The majority of patients with primary or secondary brain tumours are cognitively impaired, as a combined result of their tumour, treatment, medical co-morbidities, and psychosocial factors[1-4]. Detection of cognitive impairment is clinically important as it has been shown to be an independent predictor of survival and disease recurrence, and a means of detecting patients who may benefit from neuropsychological rehabilitation[5-9]. Furthermore, in clinical trials, cognitive assessments are essential to define differing effects of treatments on cognitive function, either positively, through neuroprotection and better tumour control, or negatively, through treatment induced neurotoxicity[10,11].

Currently, the most commonly chosen tool to screen for cognitive impairment in patients with brain tumours is the Mini Mental State Examination (MMSE), despite its lack of validation in this setting. The MMSE was developed to detect gross cognitive impairment and dementia, and has been shown to be relatively insensitive to the cognitive changes that occur in

From the BC Cancer Agency, Vancouver Cancer Centre (RO, MM), Radiation Therapy Program; University of British Columbia, Division of Radiation Oncology and Developmental Radiotherapeutics (RO, MM), Vancouver; BC Cancer Agency (MP), Fraser Valley Cancer Centre, Department of Patient and Family Counseling, Surrey, British Columbia, Canada; Harvard School of Public Health (RO), Boston, MA, USA.

Received July 7, 2009. Final Revisions Submitted September 3, 2009.

Correspondence to: Robert Olson, 600 West 10th Avenue, Vancouver, British Columbia, V5Z 4E6, Canada. E-mail: rolson2@bccancer.bc.ca.

brain tumour patients, especially in detecting impairments in abstract reasoning, executive functioning, and visual perception[12-15]. As a result, formal neuropsychological assessments (NPAs) have been advocated[10]. However, these assessments are not available to the majority of cancer patients due to their cost and the limited availability of neuropsychologists.

Given the reported insensitivity of the MMSE, and the lack of availability of NPAs, we have investigated an alternative cognitive screening measure, the Montreal Cognitive Assessment (MoCA). Initially, we performed a feasibility study that demonstrated that both the MoCA and MMSE are well tolerated, the results of which are reported elsewhere[16]. Secondly, we performed a diagnostic study, in which the sensitivity and specificity of the MoCA and MMSE were compared through the use of a gold standard neuropsychologist administered NPA, the results of which have also been presented elsewhere[17]. In brief, neither the MoCA nor MMSE had good sensitivity and specificity at a single cut-off value; however the MoCA had superior diagnostic accuracy and better correlation with quality of life[17]. Therefore, focus has been placed on the MoCA over the MMSE in this analysis.

During the accrual process of the diagnostic study, we observed that approximately three quarters of our patients declined to participate, the majority of which cited the lengthy NPA for their decision. Furthermore, many of our patients who consented to the study (including NPA) subsequently declined or failed to complete the NPA after it was scheduled. Therefore, we were concerned regarding potential selection bias introduced by the NPA, which could potentially complicate the interpretation of our results. This retrospective analysis compares the demographics and cognitive screening test results of three patient groups in order to quantify this potential bias (Figure 1). Our primary objective was to compare differences in patient demographics and cognitive screening test scores. We hypothesize that patients who consented to, and completed the NPA are younger, more fit, and score higher on the cognitive screening tests.

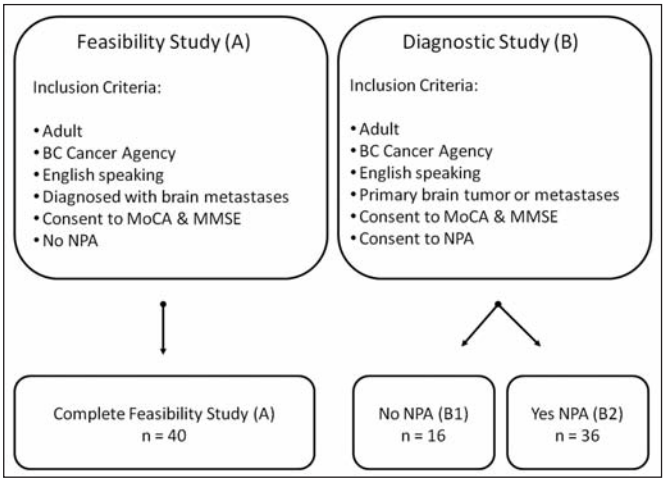

Figure 1: *Forty and fifty two patients were accrued to the feasibility (A) and diagnostic (B) studies, respectively. Inclusion Criteria are shown in the top two boxes. B1: Sixteen subjects who consented to the diagnostic study subsequently either did not initiate (n=15), or could not complete (n=1) the NPA. B2: Thirty six of the fifty two subjects who consented to NPA successfully completed the NPA. MoCA = Montreal Cognitive Assessment; MMSE = Mini Mental State Examination; NPA = Neuropsychological assessment.*

Methods

This study retrospectively compares patient demographics and cognitive screening test results from two prospective studies at the BC Cancer Agency (Figure 1). The methods of the first "feasibility" study are described in detail elsewhere[16]. In brief, patients with brain metastases were administered both the MoCA and MMSE in order to compare the feasibility of administering the MoCA against the MMSE. Patients were eligible if they were at least 18 years-of-age, English speaking, and were diagnosed with brain metastases. In a second "diagnostic" study, the diagnostic accuracy of the MoCA and MMSE were compared against a gold standard NPA. Entry criteria were similar for the diagnostic study, with the exception of a) diagnostic patients consented to NPA in addition to the MoCA and MMSE, and b) the diagnostic study included patients with primary brain tumours in addition to brain metastases (Figure 1)[17].

In order to explore the hypothesis that NPA are a source of selection bias in studies with brain tumour patients, we divided the patients into three categories: those who a) completed the feasibility study (no NPA), b) consented to the diagnostic study but did not complete the NPA, and c) consented and completed the diagnostic study (Figure 1). Differences in percentage of patients with brain metastases in Groups B1 and B2 were compared with Fisher's exact test. All other differences in patient demographics and cognitive screening test results between the three studies were assessed with one way analysis of variance (ANOVA). Subsequently, post hoc pair-wise comparisons of

Table 1: Differences in cognitive screening scores and demographic factors between the study groups

	Feasibility (n = 40)	Diagnostic: No NPA (n=16) [95% CI]	Diagnostic: Yes NPA (n = 36)	Difference between groups
mean MoCA	20.5 [18.8, 22.2]	22.6 [19.8, 25.4]	24.7 [23.7, 25.7]	p < 0.001†
mean MMSE	26.4 [25.2, 27.6]	27.8 [26.8, 28.8]	28.7 [28.0, 29.1]	p = 0.005†
mean Age	60.7 [57.4, 64.1]	53.2 [46.2, 60.2]	46.8 [42.5, 51.1]	p < 0.001†
high school completion	73% [58, 87]	87% [67, 100]	94% [87, 100]	p = 0.034†
brain metastases	N/A*	80% [57, 100]	39% [22, 56]	p = 0.013‡
dexamethasone use	75% [61, 89]	53% [27, 87]	31% [15, 46]	p = 0.002†
antiepileptic use	18% [5, 30]	27% [1, 52]	39% [22, 56]	p = 0.116†
opioid use	23% [9, 36]	20% [0, 46]	19% [6, 33]	p = 0.946†

* As per inclusion criteria, all subjects in the feasibility study had brain metastases; † ANOVA; ‡ Fisher's exact test

https://doi.org/10.1017/S0317167100010039 Published online by Cambridge University Press

MoCA scores in the different study groups were compared and statistical significance was determined after applying a Bonferroni Correction. Subset analyses of brain metastases patients were performed in order to eliminate potential confounding by diagnosis, given the different inclusion criteria of the two studies. Finally, linear regression modeling was used to assess differences in MoCA scores across the studies, controlling for all collected potential confounders. Analyses were conducted using SPSS Statistics 17.0 software and SAS for Windows version 9.2.

The study was approved by the Research Ethics Board at the BC Cancer Agency, and was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. All subjects gave their informed consent prior to their inclusion in the study. No potential conflicts of interest exist.

Results

Ninety two subjects were accrued to both studies (Figure 1). Of the 52 subjects who agreed to complete the NPA at accrual (B), 16 (31%) did not complete it (B1), the majority of whom (94%) did not even initiate NPA testing (Figure 1). The two most commonly cited reasons for not completing the NPA were decline in performance status or unwillingness to volunteer their time. Demographics of the individuals who maintained consent to the feasibility and diagnostic study are presented in Table 1.

Of note, at least 70% of patients approached for the diagnostic study outright declined or withdrew consent to the entire study, primarily because they were unwilling to complete a four-hour NPA. The majority of individuals who withdrew consent to the entire study (n = 24) had brain metastases (58%) or high grade gliomas (33%). The diagnoses of individuals who outright declined the study were not collected (n = 98). Since complete demographic data and cognitive screening results are not available for either of these groups, they are not included in the subsequent analyses.

Figure 2 graphically summarizes the differences in MoCA scores between the three study groups. In support of our hypothesis, mean MoCA scores were significantly different across the three study groups (p <0.001; Table 1). Furthermore, patients who consented to and completed the NPA scored higher on the MoCA than patients who were not asked to complete the NPA (p <0.001, A vs. B2; Table 2). In addition, individuals who consented to the NPA, yet did not complete it, had intermediate scores (B1; Figure 2). However, after correcting for post hoc multiple comparisons, individuals in Group B1 did not differ significantly from the other groups (Table 2).

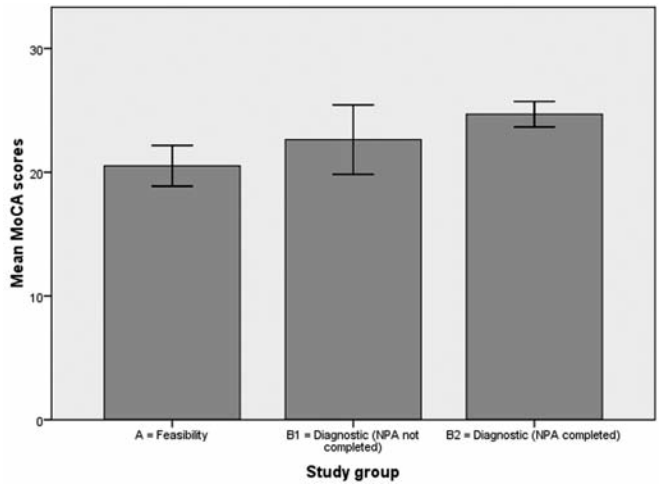

Figure 2: *Mean MoCA scores in the different study groups. Error bars represent the 95% confidence interval. MoCA = Montreal Cognitive Assessment*

As expected, the MoCA scores were consistently lower than the MMSE, since it is a more difficult and comprehensive cognitive screening test. Like the MoCA, the MMSE scores were significantly different between the study groups (p = 0.005; Table 1). Furthermore, despite the marked differences in MoCA and MMSE sensitivities[17], there are similar trends in MoCA and MMSE scores across the study groups.

In addition to differences in cognitive screening scores, we were interested in differences in patient demographic factors and other potential modifiers of cognitive function. As displayed in Table 1, there were significant differences; age (p < 0.001), education (p = 0.034), and dexamethasone use (p = 0.002) across

Table 2: Differences in MoCA results by study group

		Feasibility (A)	Diagnostic: No NPA (B1)	Diagnostic: Yes NPA (B2)	A vs. B1†	B1 vs. B2†	A vs. B2†
Mean MoCA [95% CI]	All subjects	20.5 [18.8, 22.2]	22.6 [19.8, 25.4]	24.7 [23.7, 25.7]	p = 0.45	p = 0.50	p < 0.001‡
	Brain metastases*	20.5 [18.8, 22.2]	22.0 [18.9, 25.1]	26.2 [24.8, 27.6]	p =1.00	p = 0.102	p < 0.001‡

* p = 0.001 for differences between A, B1 and B2 using ANOVA; † post hoc multiple pair-wise comparisons using ANOVA;
‡ Statistically significant after Bonferroni Correction

https://doi.org/10.1017/S0317167100010039 Published online by Cambridge University Press

the different study groups. In support of our hypothesis, with increasing involvement of the NPA (A < B1 < B2), patients' educational status increased, while their age, dexamethasone use, and opioid use all decreased (Table 1). In this sample, opioid and antiepileptic medication use was not significantly different between the three groups (Table 1).

Since the feasibility study only accrued patients with brain metastases, diagnosis could confound interpretation of the results above. Therefore, a subset analysis limited to individuals with brain metastases was performed. Within this subset, there remained a statistically significant difference in MoCA scores between the three study groups ($p < 0.001$; Table 2). In addition, individuals who completed the NPA had higher MoCA scores than individuals who were not asked to complete the NPA ($p < 0.001$, A vs. B2; Table 2). Furthermore, differences in MMSE scores ($p = 0.006$) and education ($p = 0.050$) remained significant, though dexamethasone use ($p = 0.430$) and age ($p = 0.220$) were no longer significantly different between the groups. Opioid ($p = 0.635$) and antiepileptic ($p = 0.341$) use remained similar between the three groups when restricting the analysis to the brain metastases subset.

Lastly, we assessed whether differences in MoCA scores across the three groups could be explained by the differences in demographics and previously reported modifiers of cognitive function (opioid, dexamethasone and antiepileptic medication). Using linear regression modelling, MoCA scores remained significantly different across the three study groups ($p = 0.002$), controlling for age, education, diagnosis (metastases, low grade glioma, high grade glioma), and use of opioids, antiepileptics, or dexamethasone. Furthermore, mean MoCA scores between Groups B2 and A, but not B2 and B1, remained significantly different, controlling for the above mentioned factors and after correcting for multiple comparisons using a Bonferroni Correction.

Discussion

Neuropsychological Assessments are a barrier to study participation. During the accrual process of the diagnostic study, at least 70% of patients declined to participate, the majority of whom because of unwillingness to complete a four-hour NPA. Furthermore, 29% of patients who consented to the NPA did not even initiate testing. Therefore, we were concerned that the inclusion of the lengthy NPA introduced selection bias into our study. In order to assess this potential bias, we compared patient demographics and cognitive screening scores across three groups with differing involvement of NPA (Figure 1).

As hypothesized, MoCA scores were highest for individuals who completed the NPA (Table 2). Furthermore, both MoCA and MMSE scores were inversely related with the degree of NPA involvement (Table 1). We propose that this relationship can be explained by selection bias, since older, less educated, more medicated patients, with worse prognosis tumours were preferentially deterred by the NPA (Table 1).

However, interpreting differences across the feasibility and diagnostic studies is potentially problematic, since the diagnostic study accrued individuals with primary brain tumours in addition to those with brain metastases (Figure 1). We argue that cognitive impairment from either primary or secondary brain tumours is comparable, and occurs through similar mechanisms, including radiotherapy, chemotherapy, supportive medications, psychosocial factors, and the tumours themselves[1-4]. However, there are potential differences in patient demographics, and therefore we performed a subset analysis of brain metastases patients to eliminate the possibility that such differences could confound interpretation. After restriction of the analysis to individuals with brain metastases, we found that differences in MoCA, MMSE, and education across the study groups remained (Tables 1 and 2). Furthermore, MoCA scores between the study groups remained significantly different after controlling for age, education, supportive medication, and diagnosis. Together, this suggests that patients with decreased cognitive abilities were preferentially deterred by the NPA.

This potential NPA-induced selection bias has important implications when interpreting the results of our diagnostic study[17]. First, the reported low sensitivity and specificity of both tests likely only applies to the good prognosis patients who completed the NPA, and therefore it is difficult to generalize to a broader brain tumour population. For example, we hypothesize that our reported MMSE sensitivity of only 16.6% is largely due to a ceiling effect in this population, where the majority (92%) were classified intact by the MMSE[17]. Furthermore, the minimal range in MMSE scores (SD 1.7) compared to the MoCA (SD 3.4) may explain why only the MoCA was correlated with quality of life measures[17]. In other words, we hypothesize that if a poorer performing group was studied, in which the MMSE scores were more widely dispersed and less influenced by ceiling effects, the sensitivity and correlation with quality of life would have been improved[12,17].

We hypothesize that the magnitude of selection bias in our diagnostic study would have been reduced by selecting a shorter "gold standard" cognitive assessment. This is supported by the substantial number of patients who cited the length of the NPA as their rationale for declining or withdrawing from the study. However, there are likely many individual factors that influence this decision. For example, depressive symptoms, fatigue, poor prognosis, financial difficulty, and lack of social support could all act as barriers to NPA participation. While a shorter "gold standard" cognitive assessment is not a solution for all of these factors, it is also unlikely to worsen accrual.

Though choosing a shorter cognitive assessment may decrease selection bias, it comes at the potential cost of reduced diagnostic accuracy. Certainly, we would not advocate routinely using the MMSE to assess cognition, given its reported low sensitivity in multiple settings[12,17]. A potential compromise between a brief cognitive screen and an extensive NPA would therefore be preferred. A potential candidate is adaptive computerized testing (CAT), similar to what is commonly used in the educational setting[18].

The primary benefit to CAT is the dynamic nature of the software used in such testing, which allows for shorter test duration by adapting to the patient's ability level, while still maintaining accuracy[19]. However, like NPA, the set up cost is often prohibitive in many oncology clinics. Furthermore, CAT is limited by a lack of comprehensive normative data, and a potential bias in populations less familiar with computer use. Alternative approaches to reduce selection bias include providing incentive to patients (e.g. NPA feedback, financial compensation), breaking extensive NPA into shorter sessions,

https://doi.org/10.1017/S0317167100010039 Published online by Cambridge University Press

and organizing assessments in close temporal and spatial proximity to clinic appointments.

Conclusions

Though necessary in many clinical and research situations, extensive NPAs form barriers to study participation. Therefore, researchers must weigh the benefit of their improved diagnostic accuracy against the potential for introducing selection bias. This is a common research dilemma, in which any complex assessment or intervention has the potential to exclude poor performance patients. While most researchers and clinicians understand the threat to generalizability that selection bias introduces, internal validity is also jeopardized[20]. Therefore, reducing selection bias is of utmost importance in any clinical research. Fortunately, investigators have multiple approaches available to combat this common issue[21]. Computerized adaptive testing is an emerging tool that is especially well suited to minimize selection bias in research involving cognitive assessment.

Acknowledgements

Grant funding was provided by the Fraser Valley Interdisciplinary Research Grant and the Hershey & Yvette Porte Neuro-Oncology Endowment Fund.

The authors thank Dr. Hannah Carolan, Colleen McGahan, Dr. Ann-Louise Ellwood, Dr. Lee-Ann Martin, Dr. Brian Brooks, and Professor Grant Iverson for their input into the "diagnostic" study design. In addition, we would like to thank Professor Grant Iverson and Dr. Brian Brooks for their neuropsychological expertise. Lastly, I would like to thank Professor Garret Fitzmaurice for his instruction in linear regression modelling.

References

1. Chang EJ, Wefel JS, Maor MH, Hassenbush III SJ, Mahajan A, Lang FF, et al. A pilot study of neurocognitive function in patients with one to three new brain metastases initially treated with stereotactic radiosurgery alone. Neurosurg. 2007;60(2): 277-84.
2. Taphoorn MJB, Klein M. Cognitive deficits in adult patients with brain tumours. Lancet Neuro. 2004;3(3):159-68.
3. Herman MA, Tremont-Lukats I, Meyers CA, Trask DD, Froseth C, Renschler MF, et al. Neurocognitive and functional assessment of patients with brain metastases: a pilot study. Am J Clin Oncol. 2003;26(3):273-9.
4. Mehta MP, Shapiro WR, Glantz MJ, Patchell RA, Weitzner MA, Meyers CA, et al. Lead-in phase to randomized trial of motexafin gadolinium and whole-brain radiation for patients with brain metastases: centralized assessment of magnetic resonance imaging, neurocognitive, and neurologic end points. J Clin Oncol. 2002;20(16):3445-53.
5. Meyers CA, Hess KR, Yung WKA, Levin VA. Cognitive function as a predictor of survival in patients with recurrent malignant glioma. J Clin Oncol. 2000;18(3):646-50.
6. Meyers CA, Smith JA, Bezjak A, Mehta MP, Liebmann J, Illidge T, et al. Neurocognitive function and progression in patients with brain metastases treated with whole-brain radiation and motexafin gadolinium: results of a randomized phase III trial. J Clin Oncol. 2004;22(1):157-65.
7. Meyers CA, Hess KR. Multifaceted end points in brain tumor clinical trials: cognitive deterioration precedes MRI progression. Neuro-oncol. 2003;5(2):89-95.
8. Patel RR, Mehta MP. Targeted therapy for brain metastases: improving the therapeutic ratio. Clin Cancer Res. 2007;13(6): 1675-83.
9. Sherman AM, Jaeckle K, Meyers CA. Pretreatment cognitive performance predicts survival in patients with leptomeningeal disease. Cancer. 2002;95(6):1311-6.
10. Meyers CA, Brown PD. Role and relevance of neurocognitive assessment in clinical trials of patients with CNS tumors. J Clin Oncol. 2006;24(8):1305-9.
11. Sherman AM, Jaeckle K, Meyers CA. Pretreatment cognitive performance predicts survival in patients with leptomeningeal disease. Cancer. 2002;95(6):1311-6.
12. Meyers CA, Wefel JS. The use of the mini-mental state examination to assess cognitive functioning in cancer trials: no ifs, ands, buts, or sensitivity. J Clin Oncol. 2003;21(19):3557-8.
13. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Ger Soc. 2005;53(4):695-9.
14. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psych Res. 1975;12(3):189-98.
15. Smith T, Gildeh N, Holmes C. The Montreal cognitive assessment: Validity and utility in a memory clinic setting. Can J Psychiat. 2007;52(5):329-32.
16. Olson RA, Chhanabhai T, McKenzie M. Feasibility study of the Montreal cognitive assessment (MoCA) in patients with brain metastases. Support Care Cancer. 2008;16(11):1273-8.
17. Olson RA, Iverson G, Parkinson M, Carolan H, Ellwood A, McKenzie M. Investigation of cognitive screening measures in patients with brain tumors: diagnostic accuracy and correlation with quality of life. J Clin Oncol (Meeting Abstracts). 2009;27 (15S):e13000.
18. Weiss DJ, Kingsbury GG. Application of computerized adaptive testing to educational problems. J Educ Meas. 1984;21(4): 361-75.
19. Meijer RR, Nering ML. Computerized adaptive testing: overview and introduction. App Psych Meas. 1999;23(3):187-94.
20. Berk RA. An introduction to sample selection bias in sociological data. Am Sociol Rev. 1983;48(3):386-98.
21. Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ. 2001;323 (7303):42.