Visual disability assessment: valid measurement of activity limitation and mobility in cataract patients

Konrad Pesudovs,[1] Thomas A Wright,[1] Vijaya K Gothwal[1,2]

[1]Department of Optometry and Vision Science, NH&MRC Centre for Clinical Eye Research, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia, Australia
[2]Meera and L B Deshpande Centre for Sight Enhancement, Vision Rehabilitation Centres, LV Prasad Eye Institute, Hyderabad, India

Correspondence to
Dr Konrad Pesudovs, Department of Optometry and Vision Science, NH&MRC Centre for Clinical Eye Research, Flinders Medical Centre, Bedford Park, South Australia 5042, Australia; konrad.pesudovs@flinders.edu.au

Accepted 1 November 2009
Published Online First 3 December 2009

ABSTRACT

Aim The Visual Disability Assessment (VDA), a questionnaire for measuring the impact of cataract on visual functioning, was developed using classical test theory. Since this approach is limited, our aim was to further investigate the psychometric properties of the VDA using Rasch analysis.

Methods 613 patients from the Flinders Medical Centre cataract surgery self-administered the VDA. Psychometric properties investigated for the overall VDA and each subscale included: measurement of a single construct (unidimensionality), item fit to the construct, reliable discrimination between strata of patient ability (person separation) and targeting of item difficulty to person ability.

Results The VDA discriminated five strata of patient ability. However, seven mobility items constituted a second dimension and formed a valid separate scale. Sequestration of these items resulted in a unidimensional 11-item measure of activity limitation. Both the mobility and activity limitation scales had acceptable person separation and neither contained misfitting items. Targeting was suboptimal for mobility (−2.12 logits) but good for activity limitation (−0.72). The subscales also satisfied the requirements of the Rasch measurement model.

Conclusions The Rasch-scaled VDA effectively measures two separate constructs: mobility and activity limitation (with two subscales). Its good psychometric properties make it suitable for measuring cataract surgery outcomes.

INTRODUCTION

Patient-reported outcomes (or questionnaires) are increasingly being used to track the outcomes of cataract patients.[1 2] In recent decades, a number of questionnaires,[3] including the Visual Disability Assessment (VDA),[4] have been developed specifically for cataract populations and applied in cataract outcomes research.[5–7]

Like most questionnaires (instruments), the VDA was developed using classical test theory (CTT),[4] and while highly validated with this approach,[3] CTT has several drawbacks.[8] Chiefly, this approach allows limited insight into the psychometric properties of the instrument, and scoring does not provide for interval-level measurement. Scores are simple sums of ordinal values (1, 2, 3, 4) applied to response categories (not at all, a little, quite a bit, a lot) across all questions. This assumes that the quantitative difference between each response category is equal and that each item has the same value. However, neither assumption is valid, which makes the scoring non-linear. This problem can be solved with Rasch analysis, a modern psychometric approach, which transforms ordinal scores into estimates of interval-level measurement.[9] This is important, as it allows parametric statistics to be computed from the questionnaire scores facilitating outcome measurement. Furthermore, Rasch analysis provides superior information about questionnaire performance. In particular, Rasch analysis informs whether all items measure a single construct (unidimensionality); this is an essential property of a measurement instrument. If items measure two different constructs, then the score derived from these items is confused by incompatible quantities: imagine a clinical instrument which measured corneal power and intraocular pressure and output these measures as a single number. Rasch analysis also gives insight into the measurement precision, how well the difficulties of the items target the abilities of the patients and performance of items across subgroups of patients. Such insight into instrument performance coupled with the scoring benefits of Rasch analysis have led to it being applied to a number of ophthalmic questionnaires originally developed using CTT.[10–13] The results from these studies indicate that Rasch analysis can be applied to examine and then optimise the measurement qualities of such questionnaires.

The VDA[4] has not been examined using Rasch analysis. The aim of this study was therefore to use Rasch analysis to re-examine the measurement properties of the VDA. If flawed we aimed to re-engineer the VDA to improve its measurement performance. A secondary aim was to provide spreadsheets that convert raw VDA scores to Rasch-scaled scores, to enable clinicians and researchers unfamiliar with Rasch analysis to utilise its benefits.

METHODS

Visual disability assessment

The VDA consists of 18 items that relate to how one's vision interferes with performance of visual tasks (table 1). The items are grouped into three subscales: mobility (seven items), distance/lighting/reading (eight items) and near and related tasks (five items).[4] Items 1 and 6 belong to both the distance/lighting/reading subscale and the near and related tasks subscale. Thus, the total number of items across subscales is actually greater than 18 (ie, 20).

Each item is scored on a four-category rating scale, which consists of not at all (1), a little (2), quite a bit (3) and a lot (4). A total score for the VDA is obtained by adding the scores obtained for each item and then dividing the result by the number of questions answered. Subscale scores are obtained in a similar manner. All items are scored in the same direction with higher scores indicating greater visual disability.

Table 1 Item content of the visual disability assessment

Framing question for each item is 'To what extent, if at all, does your vision interfere with your ability to'

Item no	Item description
1* ‡	Read?
2*	See in the distance?
3*	Recognise faces across the street?
4*	Watch TV?
5*	See in bright light/glare?
6* ‡	See in poor or dim light?
7‡	Appreciate colours?
8*	Drive a car by day?
9*	Drive a car by night?
10†	Walk around inside?
11†	Walk around outside?
12†	Use steps?
13†	Cross the road?
14†	Use public transport?
15†	Travel independently?
16†	Move in unfamiliar surroundings?
17‡	Do your employment/housework activities?
18‡	Do your hobbies/leisure activities?

Response options are: not at all, a little, quite a bit and a lot.
*Items belong to distance/lighting/reading subscale.
†Items belong to mobility subscale.
‡Items belong to near and related tasks subscale.

Table 2 Demographic characteristics of 613 participants who completed the visual disability assessment

Characteristic		n (%) or mean±SD
Age (years)		73.9±9.4
Gender		
Male		44.1
Female		55.9
Binocular visual acuity		
Mean±SD		
	LogMAR	0.22±0.20
	Snellen	$6/9.5^{-1}$
Range		
	LogMAR	−0.26 to 1.00
	Snellen	$6/3^{-2}$ to 6/60
Awaiting second-eye surgery		241 (39.3)
Ocular comorbidity*		
Present		282 (45.9)
Absent		309 (50.3)
Systemic comorbidity†		
Present		524 (85.3)
Absent		17 (2.8)

*Includes glaucoma, diabetic retinopathy, age-related macular degeneration etc and data were missing for 23 cases.
†Includes diabetes, hypertension, angina etc and data were missing for 73 cases.

Study population

Participants were 613 patients currently on the waiting list for cataract surgery at the Flinders Medical Centre, Adelaide, South Australia (average waiting period 3–4 months). Participants were mailed the VDA questionnaire for self-administration. Completed questionnaires were returned via a self-addressed and prepaid envelope.

All participants were aged 18 years or older and English-speaking, and had no severe cognitive impairment. The mean age was 73.9 years (SD=9.4), and 55.9% were female. The majority of the participants (60.7%) were awaiting their first cataract surgery. Participants also had coexisting ocular and systemic comorbidities representative of a typical cataract population in Australia. The clinical characteristics of participants are summarised in table 2.

Ethical approval was obtained from the Flinders Clinical Research Ethics Committee, and all participants provided informed consent. The study was conducted in accordance with the tenets of the Declaration of Helsinki.

Clinical assessment

Routine clinical assessments were performed prior to referral for cataract extraction surgery. Visual acuity was measured using computerised testing based on logMAR principles, and the screen luminance was 150 cd/m^2. As binocular visual acuity is considered representative of real-world ability,[14] this was assessed clinically and is reported here.

Rasch analysis

The data were analysed using the Andrich rating scale model[15] with Winsteps software (version 3.68).[16] The Rasch model is a probabilistic mathematical model which provides estimates of person ability and item difficulty along a common measurement continuum, expressed in log-odd units (logits). A logit is the natural log-odds of a participant being successful at a task versus being unsuccessful. For the VDA, a positive item logit indicates that the item requires a lower level of ability than the average (ie, the item is relatively easier). A positive logit for participants (person ability) suggests that the participant's visual disability is greater than the mean required level of ability for the items (ie, the overall ability required for the tasks is greater than the ability that the participant possesses).

The first step was to examine the performance of the response categories.[17] Categories should follow the intended hierarchy; that is, they should demonstrate a stepwise change in ability level from category to category (a little difficulty should represent a higher level of ability than quite a bit of difficulty). If hierarchical ordering is not observed, then categories need to be combined to repair category performance and enable further analyses.

Rasch analysis provides summary statistics in the form of person separation that represents the extent to which the items distinguish between statistically different levels of participant ability. Larger person separation indicates higher precision, meaning more distinct levels of function can be distinguished.[18] Person separation >2.0 was considered acceptable, as it indicates that the questionnaire can differentiate between three distinct levels of participant ability (ie, high, medium and low ability).

Fundamental to measurement is the concept of unidimensionality, which occurs when only one construct is represented in the measurement score. Fit statistics are primarily used to indicate unidimensionality. Fit was investigated using the infit mean square (MnSq). The MnSq value is sensitive to unexpected behaviour that affects responses to items near the participant's ability level,[19] and represents the observed variance divided by the expected variance. Therefore, the desired MnSq value is 1.0. In the current study, misfitting items were indicated by infit MnSq values outside the range of 0.7 and 1.3 (30% more or less variance than expected).[19]

Fit statistics alone are inadequate to determine dimensionality. A principal-components analysis (PCA) of the residuals (observed minus expected scores) is performed as confirmation of dimensionality.[20] In the present study, we hypothesised that the VDA items represented a single factor, visual disability. We tested this hypothesis by examining the PCA eigenvalues and contrast loading statistics. In the PCA, a high level of variance (>60%) explained by the principal component indicates that there is a very low possibility of finding additional components.

Items were considered to load onto a contrast if loadings were higher than 0.3. A contrast was considered to be evidence of multidimensionality if it had the strength of at least two items (as measured by an eigenvalue >2.0), as this is greater than the magnitude seen with random data.[16 19]

The suitability of question difficulty for the cataract population was examined by inspection of targeting (ie, matching of item difficulty to participant ability). In a well-targeted measure (ie, with a balance of easy and difficult items) the mean locations of the items and persons will be in close proximity to each other (optimal targeting <0.5 logits, good targetting 0.5–1.0 logits). Mistargeting means that either the items are too easy or too difficult for the abilities of the participants.[18]

Ideally, item difficulty should be comparable across groups, and the scale should operate in the same way irrespective of the group assessed. Differences in item difficulty across participant groups is termed differential item functioning (DIF).[21] In the present study, DIF was evaluated for age based on median split (<76 years as younger, ≥76 years as older), gender (male, female), cataract status (first eye, second eye cataract surgery), and ocular, and systemic comorbidity (present, absent). We considered these DIF variables a priori in the present study. In some samples, and particularly large ones, DIF may occur (ie, be statistically significant) but have little practical significance. For these reasons, we defined DIF according to magnitude: <0.50 logits as inconsequential DIF, 0.50–1.0 logits as minimal (but probably inconsequential) DIF and >1.0 logits as notable DIF.[10]

The measurement properties of an overall (complete) questionnaire do not infer measurement properties of individual subscales. Therefore, the three subscales of the VDA were investigated separately, using the criteria described above. Descriptive analysis was performed using SPSS for Windows (version 15.0; SPSS, Chicago, Illinois), and statistical significance was set at $p<0.05$.

RESULTS

Analysis of response categories

Response categories were ordered. This indicated that the participants utilised the categories as intended.

Overall performance of the visual disability assessment

Person separation was above the acceptable limit, indicating that the VDA could distinguish reliably between several strata of participant ability (table 3). All items had infit MnSq values within the suggested limits. PCA analysis of the residuals (eigenvalue 2.6 (5.1%) for the first contrast) revealed that seven items loaded positively (>0.3) onto the first contrast, and all these items were related to mobility. This result demonstrated that there was breach of unidimensionality because an additional dimension (mobility) was also being measured by the VDA.

Based on the above findings, the 18-item VDA was segregated into a seven-item 'mobility' scale and 11-item 'activity limitation' scale. These scales were then analysed individually.

Mobility scale

Similar to the full version, the seven-item mobility scale had satisfactory person separation (table 3). All items fitted the underlying construct, indicating the scale was unidimensional. PCA of the residuals further supported the unidimensionality. However, compared with the full version, the targeting of item difficulty to participant ability was poor; the ability (ie, level of mobility) of the participants was greater than the average item difficulty. No item showed DIF.

Table 3 Overall performance of the native version, 11-item activity limitation scale and seven-item mobility scale of the revised visual disability assessment

Versions	Visual disability assessment (native version)	Visual disability assessment (revised version): Activity limitation scale	Visual disability assessment (revised version): Mobility scale
No of items	18	11	7
No of misfitting items	0	0	0
Person separation	3.53	2.80	2.30
Mean item location	0	0	0
Mean person location	−1.36	−0.72	−2.12
Principal-components analysis—eigenvalue for first contrast (%)	2.6 (5.1%)	1.7 (5.9%)	1.7 (5.3%)
Differential item functioning (no of items)	5	2	0

Activity limitation scale

Person separation for the activity limitation scale was satisfactory (table 3). Unidimensionality was evidenced by a lack of item misfit and the PCA of the residuals. Targeting was satisfactory; items closely matched the abilities of the participants (figure 1). Two items demonstrated minimal DIF by gender.

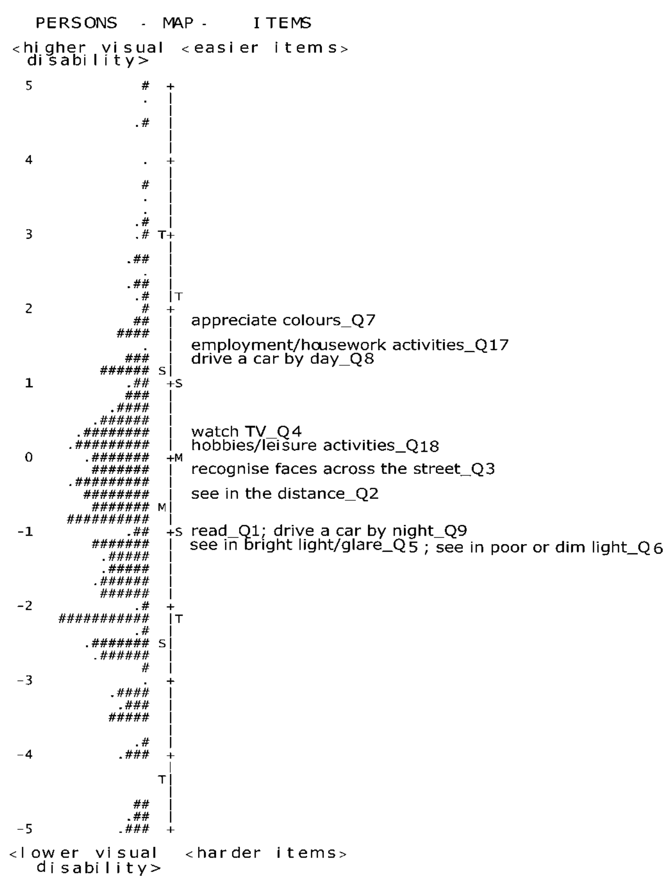

Figure 1 Person-item map of the 11-item activity limitation scale of the revised Visual Disability Assessment. The participants are mapped on the left of the dashed line and those with lower visual disability are located at the bottom of the map. Items are located on the right of the dashed line and harder items are also located at the bottom of the map. Each '#' represents three participants, and each '.' represents one to three.

Table 4 Results of subscale analysis of the visual disability assessment

Parameter	Subscale		
	Mobility scale	Distance/lighting/reading subscale	Near and related tasks subscale
No of items	7	8	5
No of misfitting items	0	0	0
Person separation	2.3	2.3	2.0
Mean item location	0	0	0
Mean person location	−2.12	−0.39	−1.16
Principal-components analysis—eigenvalue for first contrast (%)	1.7 (5.3%)	1.6 (5.1%)	1.6 (6.6%)
Differential item functioning (no of items)	0	2	2

Analysis of the subscales

Table 4 presents the analysis of the subscales from native version. All three subscales demonstrated adequate performance. The distance/lighting/reading subscale performs optimally, including ideal targeting of item difficulty to person ability. The mobility subscale is not so well targeted to person ability and is identical to the revised mobility scale above. The near and related tasks subscale possessed borderline person separation.

Conversion of raw scores to Rasch-scaled scores

Since populations vary, it is always best to implement Rasch measurement properties by actually performing Rasch analysis. However, other investigators may wish to use the VDA and also gain the interval scoring benefits of Rasch analysis, without performing Rasch analysis themselves. Therefore, we have provided a series of Excel (Microsoft, Seattle, Washington) spreadsheets which convert raw (ordinal) VDA scores to Rasch measurement estimates. These spreadsheets can be downloaded directly from the journal's website or obtained by contacting the corresponding author.

DISCUSSION

A fundamental problem with the VDA was the lack of unidimensionality that arose from the mobility items measuring a construct distinct from the remaining items. The absence of unidimensionality violates the fundamental requirement for summating a scale.[9] Nevertheless, unidimensionality could be restored, but this required that the mobility scale be separated from the activity limitation scale. This revision is consistent with the original aims of the VDA, which were to measure activity limitation and difficulty with mobility in cataract patients.[4] The formation of separate measurement of mobility and activity limitation was supported by the results of PCA of residuals, and the fit of items to each Rasch model. This finding was also consistent with previous studies which have found mobility items to form a construct separate from activity limitation in low-vision patients.[22]

The VDA was also designed to have three subscales, implying that there are three dimensions of visual functioning.[4] One of these subscales was identical to the mobility scale identified in this study. The other two subscales (distance/lighting/reading and near and related tasks) both measure within the activity limitation construct. However, it is unclear whether these subscales truly tap different aspects of functioning or simply predict the overall score from the activity limitation scale. What is clear is that all three subscales are valid unidimensional measures which can be used to report cataract surgery outcomes. This finding is important, as measurement validity of the entire questionnaire does not confer validity onto the subscales; this needs to be proven separately.

A number of other instruments designed for measuring visual disability in the cataract patient have been revised using Rasch analysis including the Catquest-9SF,[10] the Activities of Daily Vision Scale (ADVS),[11] the Cataract Symptom Scale (CSS)[12] and the Cataract TyPE Specification (TyPE).[23] In all cases, multidimensionality was a problem with the original instrument, but with misfitting items removed, valid measurement was possible except with the ADVS. These instruments are very similar, varying from nine to 12 items in length, and all measure activity limitation, plus the TyPE includes two subscales, and the CSS has a separate mobility scale like the VDA. The main performance difference is in terms of targeting of item difficulty to person ability. The ADVS, CSS and TyPE all suffer from poor targeting.[11 12 23] Prior to this study, the only questionnaire to exhibit good targeting was the Catquest-9SF.[10] For the VDA, targeting was good for the activity limitation scale but poor for the mobility scale. This probably reflects the fact that cataract patients experience little difficulty with mobility tasks. Therefore, in a cataract population, only the activity limitation scale may be required, unless one were particularly interested in mobility or were working with a less able population. Adding difficult items that target participants with less visual disability could improve the targeting of the 11-item activity limitation scale. However, redeveloping legacy questionnaires is tedious and potentially could fail. Therefore, a superior strategy may be the development of a visual disability item bank and computer-adaptive testing (CAT).[24] In CAT, Rasch calibrated items are presented based on the individual's response to previous items, which tailors the presentation of items to the ability of the participant. This approach has been used in other areas of healthcare[25] and should now be developed in Ophthalmology.

To conclude, Rasch analysis identified two unidimensional constructs in the VDA: mobility and activity limitation. The revised and improved Rasch-scaled VDA could be used to enhance our understanding of the outcomes of cataract surgery.

Funding This research was supported in part by National Health and Medical Research Council (Canberra, Australia) Centre of Clinical Research Excellence Grant 264620. Konrad Pesudovs is supported by National Health and Medical Research Council (Canberra, Australia) Career Development Award 426765.

Competing interests None.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the Flinders Clinical Research Ethics Committee, Flinders Medical Centre, South Australia.

Contributors Conception and Design (KP), Acquisition of data (TAW), Data analysis and interpretation (VKG, KP), Drafting the article/revising for intellectual content (KP, VKG, TAW), Final approval of manuscript (KP, VKG, TAW).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Black N,** Browne J, van der Meulen J, *et al*. Is there overutilisation of cataract surgery in England? *Br J Ophthalmol* 2009;**93**:13—17.
2. **Maki J,** Kusakul S, Morley K, *et al*. The effect of glasses on visual function following cataract surgery in a cataract camp. *Br J Ophthalmol* 2008;**92**:883—7.
3. **de Boer MR,** Moll AC, de Vet HC, *et al*. Psychometric properties of vision-related quality of life questionnaires: a systematic review. *Ophthalmic Physiol Opt* 2004;**24**:257—73.
4. **Pesudovs K,** Coster DJ. An instrument for assessment of subjective visual disability in cataract patients. *Br J Ophthalmol* 1998;**82**:617—24.
5. **Moradi P,** Shute N, Whitefield L. Can patients self-assess their requirement for cataract surgery? *J Cataract Refract Surg* 2006;**32**:692.
6. **Pesudovs K,** Coster DJ. Cataract surgery reduces subjective visual disability. *Aust N Z J Ophthalmol* 1997;25 Suppl 1:S3—5.
7. **Kirkwood BJ,** Pesudovs K, Latimer P, *et al*. The efficacy of a nurse-led preoperative cataract assessment and postoperative care clinic. *Med J Aust* 2006;**184**:278—81.
8. **Embretson SE,** Reise SP. *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum, 2000.
9. **Rasch G.** *Probabilistic models for some intelligence and attainment tests*. Chicago, IL, USA: MESA Press, 1980.
10. **Lundström M,** Pesudovs K. Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *J Cataract Refract Surg* 2009;**35**:504—13.
11. **Pesudovs K,** Garamendi E, Keeves JP, *et al*. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. *Invest Ophthalmol Vis Sci* 2003;**44**:2892—9.
12. **Gothwal VK,** Wright T, Lamoureux EL, *et al*. Cataract symptom scale: clarifying measurement. *Br J Ophthalmol* 2009;**93**:1652—6.
13. **Pesudovs K,** Burr JM, Harley C, *et al*. The development, assessment, and selection of questionnaires. *Optom Vis Sci* 2007;**84**:663—74.
14. **Rubin GS,** Bandeen-Roche K, Huang GH, *et al*. The association of multiple visual impairments with self-reported visual disability: SEE project. *Invest Ophthalmol Vis Sci* 2001;**42**:64—72.
15. **Andrich DA.** A rating scale formulation for ordered response categories. *Psychometrika* 1978;**43**:561—73.
16. **Linacre JM.** *A user's guide to WINSTEPS*. Chicago, IL: Winsteps.com, 2009.
17. **Linacre JM.** Investigating rating scale category utility. *J Outcome Meas* 1999;**3**:103—22.
18. **Mallinson T,** Stelmack J, Velozo C. A comparison of the separation ratio and coefficient alpha in the creation of minimum item sets. *Med Care* 2004;**42**:I17—24.
19. **Bond TG,** Fox CM. *Applying the Rasch model: fundamental measurement in the human sciences*. London: Lawrence Erlbaum Associates, 2001.
20. **Wright BD.** Local dependency, correlations and principal components. *Rasch Measurement Transactions* 1996;**10**:509—11.
21. **Holland PW,** Wainer H. *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum, 1993.
22. **Massof RW,** Hsu CT, Baker FH, *et al*. Visual disability variables. II: the difficulty of tasks for a sample of low-vision patients. *Arch Phys Med Rehabil* 2005;**86**:954—67.
23. **Gothwal VK,** Wright TA, Lamoureux EL, *et al*. Using Rasch analysis to revisit the validity of the Cataract TyPE Spec instrument for measuring cataract surgery outcomes. *J Cataract Refract Surg* 2009;**35**:1509—17.
24. **Wainer H.** *Computerized adaptive testing: a primer*. 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
25. **Haley SM,** Ni P, Hambleton RK, *et al*. Computer adaptive testing improved accuracy and precision of scores over random item selection in a physical functioning item bank. *J Clin Epidemiol* 2006;**59**:1174—82.